A Comparison of Eighth-Grade Mathematics Scores by State and by the Four Census-Defined Regions of National Assessment of Educational Progress (NAEP)

2014

Laurel Robinson
University of Central Florida

A COMPARISON OF EIGHTH-GRADE MATHEMATICS SCORES BY STATE AND BY THE FOUR CENSUS-DEFINED REGIONS OF NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP)

by

LAUREL J. ROBINSON
B. S. Stetson University 1989
M. Ed. University of Central Florida, 1993

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the School of Teaching, Learning, and Leadership
in the College of Education and Human Performance
at the University of Central Florida
Orlando, Florida

Spring Term
2014

Major Professor: Rosemarye Taylor

ABSTRACT

The purpose of this study was to investigate the information regarding the comparative relationship between the proficient mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment by state, census –defined regions and AYP subgroups. Analysis was completed and six research questions were used to guide the study. A multiple regression was used to assess the relationship between the percentage of eighth-grade students who were proficient in mathematics as assessed by the 2009 NAEP and those who were proficient in mathematics as assessed by their 2009 state assessment. A significant quadratic (non-linear) relationship between the state and NAEP levels of proficiency was determined. Several two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) were conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments for eighth grade students overall and in the following AYP subgroups : (a) low socioeconomic students, (b) white students, (c) black students and (d) Hispanic students. The within-subjects factor was type of test (NAEP or state), and the between-subjects factor was region (Midwest, Northeast, West, and South). Overall, the percentage proficient on state mathematical assessments was always higher than the percentage proficient on the NAEP mathematics assessments. The degree of discrepancy is discussed, as well as possible reasons for this divergence of scores.

These findings are consistent with other research. According to Lee (2007), "The percentages of students reaching the proficient level tend to be generally lower on NAEP than on state assessments" (p.172).

This dissertation is dedicated to my mother, Lucille Margaret Robinson.

“Always finish what you start.”

ACKNOWLEDGMENTS

I would like to begin by thanking my dissertation committee: Dr. Rosemarye Taylor, Dr. M. H. Clark, Dr. Mary Little and Dr. George Pawlas for their support and direction in this venture. My chair, Dr. Rosemarye Taylor for reviewing my writing and for pushing me to do my very best, while always ready with words of encouragement. A special thanks to Dr. M.H. Clark for her expertise in the analysis of data and to Dr. Pawlas for his caring spirit. Additionally, I need to thank Dr. Mary Ann Lynn for her assistance in editing and formatting questions. I would not have been able to accomplish this dissertation without all of you. Also, much gratitude to Leah Mitchell for her various counseling sessions and assistance with scheduling and deadlines

My family and friends have always been a source of strength and support in my life. Cindy Gardner was constantly at my side prompting me to continue and encouraging me every step of the way. A big thanks to my siblings Dennis, Leslie, Chris, Tom, Mike, Scott, Nate, Lenore, Roy, Judy and Terry for their constant words of wisdom. A special thank you to my sister, Leslie, and her husband, Doug, for being my rock. Your relentless support and encouragement was extremely important in this endeavor. Janet Kearney paved the way by being the first to finish the dissertation process which brought out my competitive spirit. Joelle Mills and Colleen King were great stress relievers who played a part in motivating me to push on through adversity and complete the task at hand. Also a special thanks to Lauren and Jackie for their contribution, and last, but not least, my colleagues at work who kept me afloat. It takes a village. Thank you all.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
THE PROBLEM AND ITS CLARIFYING COMPONENTS

Background of the Study

Throughout the history of the United States, education has been at the center of creating a stronger, better country and society for the next generation. The political, religious, economic, and industrial changes that have occurred since the earliest days of American history have affected the educational system in a variety of ways (Sadker & Sadker, 2005). The ideas about who should receive an education, what should be taught, and the overall purpose of education has been debated since the inception of the education system in the United States (Sadker & Sadker, 2005).

Sadker and Sadker (2005) summarized two schools of thought indicating "two fundamental, often opposing, purposes of schools, are (1) to transmit society's knowledge and values, passing on the cultural baton, and (2) to reconstruct society, empowering to promote social progress" (p. 159). Either view sparks great debate over the intended purpose of education. Sadker and Sadker agreed that school is the one institution that is common to all Americans and has, therefore, been an appropriate tool to make societal changes. The influence of politics, economics, religion, and industrial advances has shifted public school policy from equity to access to excellence and to the current era of high stakes testing and accountability (Sadker & Sadker, 2005). Within all of these phases of educational reform and policy shifting, accountability has become an increasingly important component.

Assessment and accountability have, in fact, been major factors in many of the education reform efforts of the past 50 years (Linn, 2005). Hansen (1993) summarized the emphasis on accountability beginning in the early 1900s and continuing through the 20th century as follows:

> The current accountability movement has historical antecedents that date back to the early 1900s and beyond. Educational accountability languished during the 1930s and 1940s, but enjoyed a minor reawakening in the late 1950s, during the Sputnik reform movement. The late 1960s marked the beginning of mandated accountability in federal programs, while accountability in the 1970s was characterized by applications of systems models and complex technical accounting systems. Now, in the 1990s we see accountability being employed as a tool for educational reform on a national scale. Historically, the accountability movement reflects continuing trends that have shaped American education and indeed, our whole society. (p. 11)

Conceptual Framework

Educational policy in regard to high stakes testing, standards based reform, and accountability relating to student outcomes in the latter part of the 20th century resulted in the largest federal involvement in education in U.S. history (Linn, Baker, & Betebenner, 2002). The signing into law of Public Law 107-110, the No Child Left Behind Act (NCLB) of 2001 in January 2002 made accountability the centerpiece of an educational agenda which already had an overarching theme of state policies to improve education

(Linn et al., 2002). This new legislation addressed concerns about the quality of education in the United States and required annual testing and reporting on the progress of all students. Student test performance has come to be an integral part of statewide accountability systems in an attempt to improve student learning and to close the achievement gap with accountability, flexibility, and choice, so that no child is left behind (No Child Left Behind Act [NCLB], 2001). Through this legislation, educational policymakers increased accountability for instruction, student outcomes, and the assignment of responsibility for the improvement of the educational system at the state and local levels.

Many researchers have investigated National Assessment of Educational Progress (NAEP) and state standardized test data results as a means to evaluate the stringency of the state standardized tests (Bandeira de Mello, Blankenship, & McLaughlin, 2009; Carnoy & Loeb, 2002; Ercikan, 1997; Gordon, 2009; Kolen & Brennan, 2004; Lee, 2007; Linn & Kiplinger, 1995; McLaughlin et al., 2008a, b; Peterson & Lastra-Anadon, 2010; Prowker & Camilli, 2007; Taylor & Gordon, 2010; Waltman, 1997). In this study, the researcher sought to provide information that would be useful to policy makers and educational stakeholders in understanding the differences in the rigor of state assessments and the implications of using a common national assessment to make accountability decisions. To accomplish this, the researcher utilized public data from the National Center for Education Statistics (NCES) to compare the 2009 mathematics eighth-grade state assessment results to the 2009 mathematics eighth-grade National Assessment of Educational Progress (NAEP), scores of proficient and above.

Statement of the Problem

The passage of NCLB increased accountability and sanctions for schools in the United States. However, differences in the rigor of examinations among the states and adequate yearly progress (AYP) calculations have impeded the ability to make state to state comparisons of school systems. "States still control many important system characteristics in complying with NCLB, such as the specification of content standards, the choice of assessments, and the setting of academic achievement standards" (Linn, 2005). Also, the difference between the states as to the number of students needed to make up a sub-group can have an effect on comparing data from one state to another. According to Taylor and Gordon (2011),

> The number of students it takes to create an AYP subgroup varies state to state from as low as 30 students to as high as 100. Schools with large enrollments or great diversity in states where 30 students make a subgroup have an increased possibility of sanctions by the federal government. (p. 29)

The number of obstacles for meeting AYP increases for large schools with diverse student populations because of the disaggregation requirements of NCLB (Linn, 2005). Given the larger number of hurdles to be cleared by more diverse schools, it is not surprising that Novak and Fuller (2003) found that schools serving more diverse student bodies were less likely to meet AYP requirements than schools serving less diverse student bodies.

By comparing the percentage of students designated as proficient on the state assessments to the percentage of students designated as proficient on the NAEP

assessment, it was anticipated that a relationship could be found between the two sets of data that could be discussed. Such a discussion based on a single uniform measure would allow for a more accurate description of the condition of education in various states. Even though the NCLB does not advocate the intention of the National Assessment of Educational Progress (NAEP) for state assessments, the expected use of the NAEP was to confirm state test results under the purview of the U.S. Department of Education, evaluating the rigor of state standards, growth in student achievement, and the reduction of achievement gaps among subgroups of students (Ad Hoc Committee on Confirming Test Results, 2002).

Purpose of the Study

The purpose of this study was to provide all stakeholders and policy makers of the United States with information regarding the relationship of state assessments and NAEP and their results. The mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment were compared by state and by census regions.

Definition of Terms

The following definitions are provided for terms particularly relevant to the study:

Adequate yearly progress (AYP). Under the Elementary and Secondary Education Act (ESEA), reauthorized as No Child Left Behind in 2002, each state has developed and implemented measurements for determining whether or not its schools and local

educational agencies (LEAs) are making adequate yearly progress (AYP). AYP is an individual state's measure of progress toward the goal of 100 percentage of students achieving to state academic standards in at least reading/language arts and mathematics. It sets the minimum level of proficiency that the state, its school districts, and schools must achieve each year on annual tests and related academic indicators. Parents whose children are attending Title I (low-income) schools that do not make AYP over a period of years are given options to transfer their child to another school or to obtain free tutoring (supplemental educational services) (U. S. Department of Education, 2010).

Adequate yearly progress (AYP) subgroup. States have the flexibility to determine what constitutes a major racial or ethnic subgroup. NCLB did not identify the major racial or ethnic groups for states but instead called upon states to make this determination based upon demographic factors within their state borders; economically disadvantaged students (students receiving free or reduced price lunch), students who are limited English proficient (LEP), and students with disabilities (SWD) (U.S. Department of Education, 2004b).

National Assessment of Educational Progress (NAEP). "The National Assessment of Educational Progress (NAEP) is the largest nationally representative and continuing assessment of what America's students know and can do in various subject areas. Assessments are conducted periodically in mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history. Since NAEP assessments are administered uniformly using the same sets of test booklets across the nation, NAEP results serve as a common metric for all states and selected urban districts. The

assessment stays essentially the same from year to year, with only carefully documented changes. This permits NAEP to provide a clear picture of student academic progress over time "(National Center for Educational Statistics, 2010b, p.2).

National Assessment of Educational Progress Census Defined Regions (Midwest). Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota and Wisconsin. (National Center for Educational Statistics, 2010b).

National Assessment of Educational Progress Census Defined Regions (Northeast). Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island and Vermont. (National Center for Educational Statistics, 2010b).

National Assessment of Educational Progress Census Defined Regions (South). Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Texas, Virginia, and West Virginia (National Center for Educational Statistics, 2010b).

National Assessment of Educational Progress Census Defined Regions (West). Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington and Wyoming(National Center for Educational Statistics, 2010b).

National Center for Education Statistics (NCES). One of the principal federal statistical agencies, is the primary federal entity for collecting and analyzing data related

to education in the United States and other nations. It provides statistical services for educators and education officials at the federal, state, and local levels; Congress; researchers; students; parents; and the media and the general public. NCES is located within the Institute of Education Sciences (IES), the research arm of the U.S. Department of Education (U.S. Department of Education, 2010).

No Child Left Behind (NCLB). "The Elementary and Secondary Education Act (ESEA), reauthorized as the No Child Left Behind Act of 2002, is the main federal law affecting education from kindergarten through high school. ESEA is built on four principles: accountability for results, more choices for parents, greater local control and flexibility, and an emphasis on doing what works based on scientific research" (U.S. Department of Education, "No Child Left Behind," 2009, p. 1).

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 (National Center for Educational Statistics, 2003).

Range of mean scores with descriptors on state assessments: (39%-57%) Low, (58%-76%) Moderate and (77%-92%) High.

Range of mean scores with descriptors on the NAEP assessments: (16%-28%) Very Low, (29%-41%) Moderately Low and (42%-51%) Low.

Research Questions

The following research questions were formulated to guide the research in this study.

1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?
2. What comparisons can be made between the percentage of eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by the NAEP in 2009?
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?
5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

Methodology

This study was a quantitative research study that investigated the relationship among the mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment by state and by census regions. Even though the NCLB does not advocate the purpose of the National Assessment of Educational Progress (NAEP) for state assessments, the expected use of the NAEP as a measure was to confirm state test results under the purview of the U.S. Department of Education, evaluating the rigor of state standards, growth in student achievement, and the reduction of achievement gaps among concerned subgroups of students (Ad Hoc Committee on Confirming Test Results, 2002).

Population and Sample

The population utilized for this study was comprised of eighth-grade students who participated in the 2009 NAEP and state mathematics assessments. The following

subgroups were represented as subpopulations: low socioeconomic students, white students, black students, and Hispanic students.

Sources of Data

The 2009 NAEP mathematics exam data were collected using the National Center for Education Statistics database. State mathematics assessment data were collected through the Department of Education's Consolidated State Performance Reports (U.S. Department of Education, 2010). For the purpose of the study, data collected were analyzed through a series of statistical procedures by the researcher utilizing the Statistical Package for Social Science (SPSS) software. No research was initiated prior to approval of the Institutional Review Board of the University of Central Florida (Appendix A).

Data Collection and Analysis

The data analyzed in the study were obtained from two sources. The 2009 eighth-grade NAEP mathematics data were retrieved using the National Center for Education Statistics database. The 2009 eighth-grade state mathematics test data were retrieved from the U.S. Department of Education's Consolidated State Performance Reports. The data retrieved from these agencies were entered into the SPSS data spreadsheet under the following categories: (a) state, (b) the four census-defined regions corresponding to the state identified by NAEP, (c) the percentage of low socioeconomic eighth-grade students identified as proficient and above in mathematics on the NAEP, (d) the percentage of low

socioeconomic eighth-grade students identified as proficient in mathematics on state assessments, (e) the percentage of white eighth-grade students identified as proficient and above in mathematics on the NAEP, (f) the percentage of white eighth students identified as proficient and above in mathematics on state assessments, (g) the percentage of black eighth-grade students identified as proficient and above in mathematics on the NAEP, (h) the percentage of black eighth-grade students identified as proficient and above in mathematics on state assessments, (i) the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on the NAEP, and (j) the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on state assessments. Data were analyzed using a multiple regression, a split-plot two-way analysis of variance (ANOVA), and paired t-tests. The research questions and sources of data are displayed in Table 1.

Table 1

Research Questions and Sources of Data[a]

Research Questions
1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?
2. What comparisons can be made between the percentage of eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by the NAEP in 2009?
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South)

[a]Sources of data for all research questions were (a) U.S. Department of Education (SY 2008-2009) Consolidated State Performance Report (CSPR) and (b) National Center for Education Statistics (NCES).

Limitations

This research was limited to some extent by the data itself. Data were collected from an outside agency. Also, the state standards and assessments of proficiency in eighth-grade mathematics were developed independently by each state and with various

degrees of stringency. The National Center for Educational Statistics (NCES) concluded that though all states used annual standardized tests to report on the performance of students on their specific curriculum objectives, state assessments varied substantially from state to state, making it impossible to make state by state comparisons (National Center for Educational Statistics, 2010) According to Linn (2002),

> State content standards, the rigor of their tests, and the stringency of their performance standards vary greatly; consequently, the percentage of students who score at the proficient level or higher on the state assessments varies radically from state to state (p. 1).

Lee (2007) identified several reasons for observed differences of state and NAEP assessment results which could be considered limitations related to the data used in conducting the study. They are as follows:

> (1) Despite the same or similar labels of standards, states' own performance standards were set at different levels than the NAEP performance standard. (2) NAEP and state assessments use different approaches to setting cut points for achievement levels. (3) States with high stakes testing could exert greater pressure and result in possible inflation of proficiency level. (4) State assessments had narrower distributions of item difficulties and thus less discriminating power than the NAEP. (p. 197)

Another limitation was related to the difference in the motivational aspects of each test. The state assessments have been considered high stakes tests for students and student performance impacts satisfactory progress leading to graduation from high

school. In contrast, no such stakes are associated with NAEP performance (Decker & Bolt, 2008).

Assumptions

It was assumed that the 2009 eighth-grade mathematics scores collected from the National Center of Educational Statistics were accurate and reliable. Similarly, it was assumed that the data collected from the United States Department of Education SY2008-2009 Consolidated State Performance Report were accurate and reliable.

Significance of the Study

This study may provide information to policy makers and educational stakeholders who are involved in the decision making process of school assessment and accountability. It may also provide a measuring stick to allow for state to state comparisons based on the ability to compare the stringency of state assessments by comparing them to a "common yardstick", the NAEP mathematical examinations. By comparing the percentage of students who were proficient on the various state mathematical assessments to the percentage of students who were proficient on the NAEP mathematics examination, stakeholders can examine differences in the percentages of a state's schools making AYP due to the varying difficulty of the state assessments. This study may further bolster the need to move towards a common national assessment of academic progress.

Summary

This chapter has provided an introduction to the study. Included were the problem, purpose of the study, and definition of terms. Also presented were the conceptual framework, research design, data collection and analysis procedures. Chapter 2 contains a review of the literature. Chapter 3 contains a description of the methods and procedures used in conducting the research. Chapter 4 contains a presentation of the analysis of the data, and Chapter 5 contains a summary of the findings, implications of the results, and recommendations for future study.

CHAPTER 2
REVIEW OF LITERATURE

Introduction

Reform is not a new idea in the education realm. Historically, the roots of educational reform in the United States can be traced to the formation of the country. The historical background of accountability began with legislation in the 1960s, and continued with the publication of "A Nation at Risk" (Hansen, 1993). The focus on accountability in the United States school system has become increasingly prevalent since the inception of the No Child Left Behind Act of 2001.

The idea of accountability also brings competition. Research has attempted to investigate ways to compare states to one another in order to inform policy makers and stake holders on the condition of education in their respective states. The National Assessment of Educational Progress (NAEP) assesses what students can do in various subject areas and has been viewed as a 'common yardstick' measuring academic proficiency (NCES, 2010). This national assessment has enabled the comparison of state proficiency levels on current state assessments.

In order to appreciate the direction that the United States has taken in regard to education, it is important to review the history of the country's educational development and the foundations that led to the educational system in existence at the time of this study. In the early years of its development, education policy was considered a matter for consideration by individual states (Sadker & Sadker). As the nation grew and

progressed, however, the federal government assumed an expanded role in public schools (Sadker & Sadker, 2005).

This chapter has been organized to provide a historical perspective of educational reform and pertinent legislation from the Colonial Period to the present day in an attempt to provide the context for the current era of high stakes testing and accountability. Sadker and Sadker (2005) stated, "The complex network of expectations surrounding today's schools is the product of a society that has been evolving for over three centuries. Individuals, groups and the government all have contributed to making public schools more accessible" (p. 284). Literature has been reviewed with a specific focus on access, equity, excellence, and accountability, and the historical legislation that shaped education in the United States (Murray & Murray, 2007).

In reviewing the literature, the researcher searched in many sources. These included books, periodicals, articles, dissertations, technical briefs and government documents. Indexing services utilized were the Educational Resources Information Center (ERIC) online database, ProQuest Dissertations and Theses, Sage, Wilson Web, EBSCO and the National Center for Education Statistics and the US Department of Education websites. The search included peer reviewed articles published after 1966. Key words searched included, accountability, NAEP/State assessment, proficiency, student evaluation and measurement, NCLB, Federal legislation, National Standards and student achievement. The chapter has been organized to address the components which emerged as important to the research in the literature review: (a) historical perspectives

of educational reform, (b) accountability, and (c) assessment in the mathematics content area with special attention to the linking of NAEP and state assessments.

Historical Perspective of Educational Reform

To understand the direction of education policy, one must first investigate the history of educational policy. Many present-day educational issues can be traced to the origins of the U.S. educational system. According to Butts and Cremin (1953),

> The study of the history of education will not solve our present problems nor will it dictate the roads to the future, but intelligent decisions cannot be reached without it. We believe, therefore, that the study of the history of education is one of the ways in which the profession and the public together should prepare themselves for making better judgments about American education. (pp. v-vi)

By understanding the past struggles and pitfalls of the American educational system, educators and stakeholders are better armed to deal with the future challenges (Butts & Cremin, 1953).

Education in the Colonial Period

Many of the current issues related to the American educational system had their roots in the nation's colonial period. As noted by Butts and Cremin (1953), there were three aspects in particular that began taking shape in the colonial period that remained issues of concern at the time of the present study: (a) the proper role of education in relation to the state; (b) the proper role of religion in education; and (c) the merit of

equality of educational opportunity. During the colonial period, debates on the responsibility to educate the nation's children began, and over time, the basic patterns of economic, class, and sectional distinctions were established within the educational process (Butts & Cremin, 1953).

During the early colonial period, the education of children followed a Christian religious based system and was understood to be the responsibility of parents. Though the Massachusetts Bay Colony had established education agencies by the 1630s, parents were responsible for teaching children to read and write. Schooling was patterned after English cultural norms (Urban & Wagoner, 2009).

The beginning of the path to mandatory education in America was established by the statute known as the Massachusetts Act of 1642. Barnard and Burner (1975) quoted Cremin in explaining this first piece of educational legislation:

> This statute empowered the selectmen of each town to take account from time to time of all parents and masters, and of their children, concerning their calling and employment of their children, especially of their ability to read and understand the principles of religion and the capital laws of this country and authorizing them to put forth apprentices the children of such as they shall not be able and fit to employ and bring them up. (p. 4)

It should be noted, however, that education itself was not first and foremost in the thoughts of the founding fathers. They came to America in order to escape the religious oppression they had to contend with in England (Butts & Cremin, 1953). The founding fathers carried with them to America a longing to generate a world where they could have

religious freedom (Butts & Cremin, 1953). The idea of compulsory education was needed to further their societal aspirations by enabling the citizens to understand the written codes, both religious and secular, under which the colonies were living (Butts & Cremin, 1953). Without some sort of education, such understanding would not have been possible.

The integral aspects of the Law of 1642 was that it had nothing to do with the establishment of schools. It stated that parents and masters of those children who had been apprenticed to them were responsible for their basic education and literacy. All children, and servants as well, should be able to demonstrate competency in reading and writing as outlined by the governing officials (Butts & Cremin, 1953). In this early period of the 1600s, there was no idea of a formal school. It was understood that individuals would be sufficiently educated to meet the individual needs of their stations in life and social harmony would be that much closer. Parents were assumed to be the best persons to educate their children. The law did state, however, that if parents and masters grew lax in their responsibility and their children were not able to meet basic criteria, the government could step in. In such cases, it was the government's right to remove children from the home and place them in a place where they could receive adequate instruction (Butts & Cremin, 1953).

The Massachusetts Law of 1647, also known as the Old Deluder Satan Act, was born out of this supposed parental negligence, and formal schooling became more important. The Law of 1647 required towns of 50 families to hire a schoolmaster who would instruct children to read and write. (Sadker & Sadker, 2005).

During this period, teachers were hired by individual parents and guardians as well as government entities. Once teachers began to work educating the youth of the new nation education became more of a social responsibility (Sadker & Sadker, 2005). School, during this period, became a higher priority than it had previously been. Dame schools first appeared at this time which allowed women to tutor students in their home for a small salary. (Sadker & Sadker, 2005). School masters travelled from town to town to further the societal vision through religion and literacy (Sadker & Sadker, 2005). In Pennsylvania, William Penn, a Quaker, spearheaded efforts to pass a 1683 law requiring all children in Pennsylvania, including women, blacks, and Indians, to read and write (Butts & Cremin, 1953). Butts and Cremin described the changes that occurred during the colonial period as follows:

> At the beginning of the colonial period educational thought was dominated by theological, philosophical, political, and social orthodoxies; by the end of the colonial period more and more voices were being heard proposing an education that would be more liberal, more secular, more scientific, more utilitarian, more humanitarian and more democratic. (p. 65)

Education in the 18th Century

According to Bracey (2009), in 1782 Thomas Jefferson created an education system for America that would be a great sorting machine based on his belief that though all men might be equal in some moral or legal sense, they were not equal in the intellectual sense. Jefferson went beyond educating just a small elite class of students or

using only religious instruction to provide education to children from all economic and social classes (Sadker & Sadker, 2005). His plan did not, however, allow access to the educational system for women past the primary grades or for the children of slaves. Clearly, the controversy over school reform and federal influence over that reform has been a topic of debate in the United States for many years (Barnard & Burner, 1975).

In that same era, Benjamin Franklin created a new kind of secondary school in Pennsylvania called the Franklin Academy. These secondary schools or academies were free from religious influence and offered various practical subjects such as, mathematics, astronomy, athletics, navigation, drama, and bookkeeping. The Franklin Academy later became the University of Pennsylvania (Sadker & Sadker, 2005).

Education in the 19th Century

In the beginning of the 19th century, schools were seen as a luxury because there was such disparity between the rich and the poor. There were very few schools in the south and rural areas (Sadker & Sadker, 2005). The quality of the schools was in question. Poor white people, immigrants, and urban laborers also known as the "common people" demanded greater participation in the democracy and greater access to education. Horace Mann, also known as the father of the public school (Sadker & Sadker, 2005, p. 209) created a common school which evolved into the present day public elementary school. These schools were only open to white children (Sadker & Sadker, 2005).

The Supreme Court decision in *Plessy v. Ferguson* (1896) cited the 14th amendment in a decision challenging racial segregation in schools. It was not until

Brown v. Board of Education (1954) that the separate but equal practice of public school systems was ruled unconstitutional and thereby illegal. Access became a priority in public education as a result of the Brown decision (Murray & Murray, 2007).

Education in the 20th Century

Many researchers have discussed the various conflicts in history that occurred in the 1900s that influenced the educational policy of the 20th century (Bracey, 2003; Gordon, 2009; Linn, 2006). World War I, The Great Depression, World War II, the Korean War, the Vietnam War, and the Cold War all played parts in the various legislation that occurred in this span of time (Bracey, 2003). The beginning of the 20th century was characterized by the progressive education movement which called for a more student-centered educational approach. John Dewey was the most notable person attached to this movement because of several essays he wrote on experiential learning in which he challenged the traditional education of the early 1900s (Sadker & Sadker, 2005).

The launching of Sputnik by the Soviet Union in 1957 propelled the country onto a traditional educational path with an emphasis on mathematics and science. The United States found itself in a space race with the Soviet Union to accomplish a moon landing (Bracey, 2007). In 1958, soon after the launching of Sputnik, the U.S. Congress enacted the National Defense Education Act (NDEA). Schools were considered by some to be at the heart of the issue, and the NDEA sought to improve instruction and curriculum through training of teachers and providing scholarships to students in subjects that would

help to secure the national defense (Sadker & Sadker, 2005). This was, in effect, the beginning of higher accountability placed on schools and states by the federal government. Project Talent was undertaken from 1957-60 as "the first scientifically planned inventory of human talents" (Passow, 1960, p. 147). It was conducted by the U.S. Office of Education in conjunction with The National Science Foundation and the Office of Naval Research. This study was the first to focus on subgroups of low socio-economic student achievement.

The largest source of federal support for K-12 education to date has been the Elementary and Secondary Education Act of 1965 (ESEA) (National Center for Educational Statistics, 2003). This act authorized grants for (a) elementary and secondary school programs for children of low-income families; (b) school library resources, textbooks, and other instructional materials for school children; (c) supplementary educational centers and services; (d) strengthening state education agencies; and (e) educational research and research training (National Center for Educational Statistics, 2003).

In the 1960s, advocates sought a federal role in providing leadership and funding for efforts to provide a free appropriate public education (FAPE) to children with disabilities. Congress took a step toward this in 1966 when it established the Bureau for Education of the Handicapped under Title VI of the Elementary and Secondary Schools Act (ESEA). Public Law 94-142, passed in 1975, required public schools to include children with disabilities and give them equal access to programs and curricula. P. L. 94-142 grew out of the courts, namely the *Pennsylvania Association of Retarded Citizens*

(*PARC)* and *Mills* cases (Murray & Murray, 2007). In both court cases, parents of children with disabilities challenged the school systems in an effort to gain access to public education for their children. Until the mid-1970s, schools could deny education to children with disabilities. Students with disabilities were systematically refused access to education for a variety of reasons (Itkonen, 2007). In 1991, PL 94-142 was renamed Individuals with Disabilities Education Act (IDEA) (Itkonen, 2007).

The overarching themes of access and equity were addressed during much of the 20th century (Murray & Murray, 2007). Given the enhanced role of the federal government in education, there was a need to construct an accountability system that could give the government a view of the condition of the educational system in the United States. In 1981, then Secretary of Education Terrell Bell created the National Commission on Excellence (U.S. Department of Education, 1983). He assigned the commission the task of investigating the quality of education in the United States and reporting to him within 18 months. The Commission's report, *A Nation at Risk,* called for a wide range of reforms that would improve the quality of education (U.S. Department of Education, 1983). According to Peterson and West (2003), "*A Nation at Risk* pushed the nation further toward accountability, principally by raising educational issues higher on state political agendas" (p. 6), thereby defining excellence of the country's education system as an important goal in its effort to compete with the world's advanced societies. This resulted in Goals 2000: Educate America Act (P.L. 103-227) which was signed into law on March 31, 1994 and established excellence in education as a goal for the future (Goals, 2000).

Education in the 21st Century

In the era of high stakes testing, the most prominent federal legislation of the first decade of the 21st century was the No Child Left Behind (NCLB) Act of 2001. According to one political analyst (Broder, 2001), NCLB "may well be the most important piece of federal legislation in thirty-five years" (p. A1). The NCLB Act revised the Elementary and Secondary Education Act (ESEA) of 1965 that provided for revisions every five years.

NCLB was built on four pillars: (a) stronger accountability for results, (b) freedom for states and communities, (c) proven education methods, and (d) more choices for parents. Under NCLB, states were challenged to work to close achievement gaps and ensure that all students achieved academic proficiency by 2014 (U.S. Department of Education, NCLB, 2004b). The policies of NCLB were stated in the following 10 titles of NCLB:" (a) Title I, Improving the Academic Achievement of the Disadvantaged; (b) Title II, Preparing, Training, and Recruiting High Quality Teachers and Principals; (c) Title III, Language Instruction for LEP and Immigrant Students; (d) Title IV, 21st Century Schools; (e) Title V, Promoting Informed Parental Choice; (f) Title VI, Flexibility and Accountability; (g)Title VII, Indian, Native Hawaiian, and Alaska Native Education; (h) Title VIII, Impact Aid Program; (i) Title IX, General Provisions; and (j) Title X, Repeals, Re-designations, and Amendments to Other Statutes" (107th Congress, 2002, p.1).

The American Recovery and Reinvestment Act of 2009 was signed into law by President Barack Obama on February 17, 2009. In speaking about this legislation, the U.S. Department of Education (2009) indicated:

> It is an unprecedented effort to jumpstart our economy, create or save millions of jobs, and put a down payment on addressing long-neglected challenges so our country can thrive in the 21st century. The Act is an extraordinary response to a crisis unlike any since the Great Depression, and includes measures to modernize our nation's infrastructure, enhance energy independence, expand educational opportunities, preserve and improve affordable health care, provide tax relief, and protect those in greatest need. (U.S. Department of Education, 2009, p. x)

At the time of the present study, work was ongoing to reauthorize the Elementary and Secondary Education Act of 1965, known as the blueprint to reform. This reform was initially focused on four areas: (a) improving teacher and principal effectiveness, (b) providing information to families to help them evaluate and improve their children's school, (c) implementing college- and career-ready standards, and (d) improving student learning and achievement in America's lowest-performing schools by providing intensive support and effective interventions (U.S. Department of Education, 2009).

Accountability

In October of 1969, the National Assessment of Educational Progress (NAEP) emerged as a federally supported program under the U.S. Office of Education (USOE) (Bourque, 2009). In 1971, the USOE assigned administrative responsibility for the NAEP testing and data reporting to the National Center for Educational Statistics (NCES) NCES, in developing NAEP, has been responsible for the development of the largest nationally representative sample and continuing assessment of what America's students

know. NAEP assesses what students can do in various subject areas and has been viewed as a 'common yardstick' measuring academic proficiency. Assessments are conducted periodically in mathematics, reading, science, writing, the arts, civics, economics, geography, and U.S. history (NCES, 2010).

The uses for NAEP data have changed over the years. The early years of NAEP score reporting were more general in nature and described trends, demographics, and regions of the country (Bourque, 2009). By 1986, NAEP started to use the report card method of reporting data to the nation which made it easier for stakeholders to use and understand. NCLB required states to participate in the main NAEP for mathematics and reading at Grades 4 and 8 every two years (NCES, 2010). Lane et al. (2009) wrote that NAEP data had evolved to the point of being useful in state by state and international comparisons.

With passage of the No Child Left Behind Act of 2001, state and federal policymakers have looked to the National Assessment of Educational Progress (NAEP) to perform a new role. This role involves providing information about student achievement in Grades 4 and 8 in reading and mathematics that can be used by the U.S. Department of Education as confirmatory evidence about student achievement on state tests (NCES, 2010). In the parent guide that explains NCLB's many facets, the use of the national test data were discussed as follows:

> NAEP data will highlight the rigor of standards and tests for individual states: If there is a large discrepancy between children's proficiency on a state's test and their performance on NAEP that would suggest that the state needs to take a

closer look at its standards and assessments and consider making improvements. (U.S. Department of Education, 2003, p. 14)

No Child Left Behind (NCLB)

The No Child Left Behind Act of 2001(NCLB) required each state to test students in Grades 3-8 in mathematics and language arts starting no later than the 2005-2006 school year. In accordance with NCLB,

> Each State plan shall demonstrate that the State has adopted challenging academic content standards and challenging student academic achievement standards that will be used by the State, its local educational agencies, and its schools to carry out this part, except that a State shall not be required to submit such standards to the Secretary. (P.L. 107-110, Section 1111(b) (1) (A)

The fact that every state constructed their own content standards with discrepancies in the rigor of those standards has made it impossible to use state assessments in comparing educational progress. NCLB further detailed that:

> . . . each State plan shall demonstrate that the State has developed and is implementing a single, statewide State accountability system that will be effective in ensuring that all local educational agencies, public elementary schools, and public secondary schools make adequate yearly progress as defined under this paragraph (P.L. 107-110, Section 1111(b) (2) (A)

Linn (2005) observed, ". . . substantial differences between the accountability requirements of many state systems and NCLB still have resulted in mixed messages

regarding the performance of schools" (p. 1). In general, the diversity of state accountability systems can produce a result that has nothing to do with student performance and everything to do with the stringency of the standards in place. Therefore, comparison to a common measurement is paramount to evaluating various state school systems (Linn, 2005).

The Ad Hoc Committee, assisted by the Planning Work Group, studied NAEP's capacity to serve as a source of confirmatory evidence for state test results. Through the examination of state test results in eight states, the preparation of "arguments" about performance in three of those states, and the use of relevant NAEP data, the Ad Hoc Committee concluded that the National Assessment of Educational Progress could serve this role effectively. The Committee identified factors that could limit this role and made recommendations to address these factors. The Committee also recommended new ways of representing achievement gains and achievement gaps and encouraged further work to provide such information in formats accessible to the general public (Ad Hoc Committee on Confirming Test Results, 2002).

Adequate Yearly Progress (AYP)

NCLB requires measurable adequate yearly progress (AYP) objectives. These must be measured for all students. In addition subgroups of pupils must be measured. These subgroups include socio-economic background, English language proficiency, race-ethnicity, and disabilities. NCLB requires that states develop AYP objectives accordingly:

1. States must develop AYP statewide measurable objectives for improved achievements by all students and for specific groups: economically disadvantaged students, students from major racial and ethnic groups, students with disabilities, and students with limited English proficiency.
2. The objectives must be set with the goal of having all students at the proficient level or above within 12 years (i.e., by the end of the 2013-2014 school year).
3. AYP must be based primarily on state assessments, but must include one additional academic indicator.
4. The AYP objectives must be assessed at the school level. Schools that have failed to meet their AYP objective for 2 consecutive years will be identified for improvement.
5. School AYP results must be reported separately for each group of students identified above so that it can be determined whether each student group met the AYP objective.
6. At least 95% of each group must participate in state assessments.
7. States may aggregate up to 3 years of data in making AYP determinations. (Linn et al., 2002, p. x)

As shown in these guidelines, states set their own AYP targets. This aspect of NCLB leads to some discrepancies when reporting state data on levels of proficiency.

> While each state has constructed its own definition of Adequate Yearly Progress (AYP) requirements within the confines of NCLB, substantial differences between the accountability requirements of many state systems and NCLB still

have resulted in mixed messages regarding the performance of schools. (Linn, 2005, p. x)

Schools may meet goals according to NCLB standards but not reach goals of their own state accountability systems. The reverse may also occur.

In order to achieve AYP, a minimum percentage of students must score at the proficient level or higher on the state assessment. The percentage encompasses the overall student population as well as students who fit in specifically identified subgroups. The subgroups look more closely at students who reside in low income families, have limited English proficiency, have disabilities, or are considered to be racial or ethnic minorities (Olson, 2005).

Olson (2002b) reported on the work of a committee made up primarily of testing experts who studied data from eight states to determine the viability of comparing the outcome of individual state assessments with the state's results on NAEP. The group determined that NAEP outcomes were in effect mirroring the movement of state assessment outcomes. The committee's report also noted the complexity of such analysis. The report supported states and NAEP in defining student subgroups in a similar fashion whenever possible (Olson, 2002b).

Content Area Assessment: Mathematics

Since its first mathematics assessments in the early 1970s and early 1980s, NAEP has regularly gathered data on students' understanding of mathematical content (U.S. Department of Education, 2007). Although the names of the content areas in the

frameworks and some of the topics in those areas may change somewhat from one assessment to the next, a constant effort toward accumulating information on student performance in five key areas remains. The framework for the 2009 Mathematics Assessment was anchored in these same five broad areas of mathematical content: (a) number properties; (b) measurement; (c) geometry; (d) data analysis, statistics, and probability; and (e) algebra (U.S. Department of Education, 2007).

Linking Statewide Tests to the National Assessment of Educational Progress (NAEP)

Although both NAEP and state assessment results may be used as a tool to compare state-level performance, previous appraisals of the NAEP and state assessment results showed significant discrepancies in the level of student achievement as well as the size of statewide achievement gains. The percentages of students reaching the proficient level tended to be generally lower on NAEP assessments than on the state assessments (Klein, Hamilton, McCaffrey, & Stecher, 2000; Lee & McIntire, 2002; Linn et al., 2002).

Linn and Kiplinger (1995) discussed the problems associated with comparisons of test results. They indicated that there were many statistical aspects that must be satisfied in order to compare different assessments:

> It has long been a common practice to equate results of different forms of a test then treat the results as interchangeable and, the validity of comparisons across tests or assessments may depend on the context of the assessments, the groups used to calculate statistics, and the time of administration. (p. 136)

In accordance with the multiple measures of accountability imposed by the No Child Left Behind Act (NCLB), Lee (2007) inspected similarities and differences between the National Assessment of Educational Progress (NAEP) and math assessment results in Kentucky and Maine in 1996 and in 2003. In 1996, Lee found that 31% of eighth graders in Maine met the NAEP's proficiency level in mathematics, but only 9% of the students met the Maine Educational Assessment (MEA) advanced level. In comparing 1996 NAEP and Instructional Results Information System (KIRIS) in Kentucky, results revealed inconsistent performance results. The meta-analysis of the NAEP, MEA, and KIRIS assessment results identified inconsistent percentages of proficient students from the two states. The results of the scores from 1996 indicated large effect sizes and also large discrepancies between the two assessments. (Lee, 2007).

One of the integral issues with comparing state and NAEP assessments was the definition of proficiency and the cut score that represented this term. Bracey (2008) stated, "On virtually all tests these days, there is a score that determines whether a student passes or fails, is proficient or not or is being educated or left behind. This is the cut score" (p. 20). States have been allowed to define proficient within their borders which has led to widely different definitions of what counts as proficient. This discrepancy has been discussed as a problem since at least 1996. The classification of proficient or not has been dependent on geography (Olson, 2002a).

Summary

This chapter presented a review of the literature and research related to (a) historical perspectives of educational reform, (b) accountability, (c) mathematics content area assessment including the linkage of NAEP and state assessments. The methodology used for the study, instrumentation, and the statistical procedures employed to analyze the research questions to determine if differences in the proficiency levels exist in the 2009 NAEP and state eighth grade mathematical assessments are outlined in Chapter 3. Chapter 4 includes the results of the statistical analysis of the data. Chapter 5 provides a summary and discussion of the findings of the study, implications for practice, and recommendations for future research.

CHAPTER 3
METHODOLOGY

Introduction

This chapter contains an explanation of the methods and procedures used to conduct the study which involved a comparison of quantitative data from the eighth-grade 2009 National Assessment of Educational Progress (NAEP) mathematics assessment with the 2009 state mathematics assessment. The 2009 NAEP mathematics assessment was accessed using the National Center for Education Statistics database. State mathematics assessment data were collected using the U.S. Department of Education's Consolidated State Performance Reports. The sections found within this chapter are: (a) statement of the problem, (b) population and sample, (c) research questions, (d) instrumentation, (e) data collection, (f) data analysis, and (g) statistical procedures

Statement of the Problem

The passage of NCLB increased accountability and sanctions for schools in the United States. However, differences in the rigor of examinations among the states and adequate yearly progress (AYP) calculations have impeded the ability to make state-to-state comparisons of school systems. Also, the difference between the states as to the number of students needed to make up a sub-group can have an effect on comparing data from one state to another (Taylor & Gordon, 2011).

By comparing the percentage of students designated as proficient on the state assessments to the percentage of students designated as proficient on the NAEP assessment, a relationship is found between the two tests that can be discussed. Such a discussion based on a single uniform measure allows for a more accurate description of the condition of education in various states assessments (Klein et al., 2000; Lee & McIntire, 2002; Linn et al., 2002).

The purpose of this study was to provide all stakeholders and policy makers of the United States public school system with information regarding the comparative relationship between the mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment by state and by census regions.

Population and Sample

The population utilized for this study was comprised of eighth-grade students who participated in the 2009 NAEP mathematics assessment and the 2009 state mathematics assessment. The eighth grade population was examined and entered in the following categories (a) state, (b) the four census-defined regions corresponding to the state identified by NAEP, (c) the percentage of low socioeconomic eighth-grade students identified as proficient a in mathematics on the NAEP, (d) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the state assessment, (e) the percentage of white eighth-grade students identified as proficient a in mathematics on the NAEP, (f) the percentage of white eighth students identified as

proficient a in mathematics on the state assessment, (g) the percentage of black eighth-grade students identified as proficient a in mathematics on the NAEP, (h) the percentage of black eighth-grade students identified as proficient in mathematics on state assessments, (i) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on the NAEP, and (j) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments.

Instrumentation

The data from the eighth grade 2009 NAEP mathematics assessment and the 2009 state mathematics assessments were analyzed in this study. NAEP is a congressionally authorized project of the National Center for Education Statistics (NCES) within the Institute of Education Sciences of the U.S. Department of Education. The Commissioner of Education Statistics is responsible for carrying out the NAEP project. The National Assessment Governing Board oversees and sets policy for NAEP (U.S. Department of Education, 2010).

The mathematics framework classifies assessment questions that are used to guide the assessment in two dimensions: content area and mathematical complexity. Each question is designed to measure one of the five mathematics content areas: (a) number properties and operations; (b) measurement; (c) geometry; (d) data analysis, statistics, and probability; (e) and algebra (U.S. Department of Education, 2010).

Research Questions

The following research questions were formulated to guide the research in this study.

1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?
2. What comparisons can be made between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by the NAEP in 2009?
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient a in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?

6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?

Sources of Data

All of the data utilized in this study were archival data. NAEP mathematics assessment data were collected through the National Center for Education Statistics database. State mathematics assessment data were acquired from the Department of Education's Consolidated State Performance Reports (U.S. Department of Education, 2010). Eighth-grade mathematic performance data were collected from the NCES 2009 State Snapshot Report (Appendix B) and from the SY 2008-2009 Consolidated State Performance Reports (Appendix C). Collected data were organized in an SPSS worksheet and disaggregated by the four census-defined regions identified by NAEP as shown in Appendix D. Upon completion of regional disaggregation, data were further disaggregated by low socioeconomic eighth-grade students, white eighth-grade students, black eighth-grade students and Hispanic eighth-grade students. Appendix E contains the NAEP mathematics average report scores for 2009. Appendix F contains the definitions for NAEP achievement levels by grade.

Data Analysis

Six research questions were used to guide the study. Data gathered to answer the questions were analyzed by the researcher using a series of statistical procedures and the Statistical Package for Social Science (SPSS) software. The data were entered in the following categories (a) state, (b) the four census-defined regions corresponding to the state identified by NAEP, (c) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the NAEP, (d) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the state assessment, (e) the percentage of white eighth-grade students identified as proficient in mathematics on the NAEP, (f) the percentage of white eighth students identified as proficient in mathematics on the state assessment, (g) the percentage of black eighth-grade students identified as proficient in mathematics on the NAEP, (h) the percentage of black eighth-grade students identified as proficient in mathematics on state assessments, (i) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on the NAEP, and (j) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments. The following is a detailed explanation of the statistical procedures that were used in responding to each of the six questions.

To respond to Research Question 1 as to the relationship between state and NAEP mathematics assessments in terms of the percentage of students who were proficient in 2009, a multiple regression was used to assess the relationship between the percentage of

eighth-grade students who were proficient in mathematics as assessed by the 2009 NAEP and those who were proficient in mathematics as assessed by their 2009 state assessment.

Research Question 2 sought to determine whether NAEP census-defined regions (Midwest, Northeast, West, and South) moderated the difference between state and NAEP mathematics assessments in terms of the percentage of students who were proficient in 2009. A two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments. Follow up t-tests were also used to further describe the difference in the scores.

Research Questions 3, 4, 5 and 6 focused on whether NAEP census-defined regions (Midwest, Northeast, West, and South) moderated the difference between state and NAEP mathematics assessments in terms of the percentage of students who were proficient in 2009 among four different groups of eighth-grade students: low socioeconomic students, white, black and Hispanic. A two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments in the various AYP subgroups. Follow up t-tests were also used to further describe the difference in the scores.

For the purpose of clarity, the following descriptors were used to further describe the range of scores on each assessment. The state mathematic assessment mean scores ranged from 39% proficient to 92% proficient. The range for low was considered 39%-

57% proficient, the range for moderate was considered 58%-76% proficient and the range for high was considered 77%-92% proficient.

The range of mean scores for proficiency on the NAEP mathematics assessment ranged from 16% to 51% proficient. The range that depicted the very low proficiency levels was 16%-28%. The range that depicted the moderately low proficiency levels was 29%-41%, and the range that depicted the low proficiency levels was 42%-51%. These levels were created in order to further interpret the results.

Table 2

Research Questions, Sources of Data and Analyses

Research Questions	Data Source	Data Analyses
1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?	CSPR NCES	Multiple Regression
2. What comparisons can be made between the percentage of eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?	CSPR NCES	Split-Plot ANOVA
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?	CSPR NCES	Split-Plot ANOVA
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?	CSPR NCES	Split-Plot ANOVA
5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?	CSPR NCES	Split-Plot ANOVA
6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?	CSPR NCES	Split-Plot ANOVA

Note. CSPR = U.S. Department of Education (SY 2008-2009) Consolidated State Performance Report. NCES = National Center for Education Statistics.

Summary

The research design and methodology used in the study have been explained in this chapter. The sources that were utilized to secure the data for this study were accessed through the National Center for Education Statistics database. State mathematics assessment data were collected through the Department of Education's Consolidated State Performance Reports (U.S. Department of Education, 2010). Data analysis and statistical procedures were also presented in this chapter. The results of the data analysis for each of the six research questions are presented in Chapter 4. Chapter 5 is the concluding chapter in the dissertation and provides a summary and discussion of the findings, implications for policymakers and stakeholders, and recommendations for further research.

CHAPTER 4
PRESENTATION AND ANALYSIS OF DATA

Introduction

The purpose of this study was to investigate the information regarding the comparative relationship between the mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment by state and by census regions. Analysis was completed and six research questions were used to guide the study. Data gathered to answer the research questions were analyzed by the researcher using a series of statistical procedures and the Statistical Package for Social Science (SPSS) software. The data were entered in the following categories (a) state, (b) the four census-defined regions corresponding to the state identified by NAEP, (c) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the NAEP, (d) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the state assessment, (e) the percentage of white eighth-grade students identified as proficient in mathematics on the NAEP, (f) the percentage of white eighth-grade students identified as proficient in mathematics on the state assessment, (g) the percentage of black eighth-grade students identified as proficient in mathematics on the NAEP, (h) the percentage of black eighth-grade students identified as proficient in mathematics on state assessments, (i) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on the NAEP, and (j) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments.

Following are the detailed results of the statistical procedures that were used in responding to each of the six questions.

The results of the data analysis including descriptive statistics are offered in this chapter. Each of the six research questions was analyzed using the previously mentioned statistical procedures.

The following Research Questions were formulated to guide the research in this study.

1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?
2. What comparisons can be made between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by the NAEP in 2009?
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?

5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?
6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?

Descriptive Statistics

Data were collected from two sources. State data were gathered from the United States Department of Education's SY 2008-2009 Consolidated State Performance Report (2008-2009 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, 10 variables were considered integral to answering the six research questions. The variables were: (a) state, (b) the four census-defined regions corresponding to the state identified by NAEP, (c) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the NAEP, (d) the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on the state assessment, (e) the percentage of white eighth-grade students identified as proficient in mathematics on the NAEP, (f) the percentage of white eighth-grade students identified as proficient in mathematics on the state

assessment, (g) the percentage of black eighth-grade students identified as proficient in mathematics on the NAEP, (h) the percentage of black eighth-grade students identified as proficient in mathematics on state assessments, (i) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on the NAEP, and (j) the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments.

An example of the data sources were: (a) The National Center for Educational Statistics State Snapshot Report, found in Appendix B, (b), The Consolidated State Performance Report for the State of Alabama's 2009 eighth-grade students' mathematics performance found in Appendix C, and (c) The National Center for Education Statistic's Four Census-defined Regions of NAEP, found in Appendix D. Alabama was used as an example because it is the first state alphabetically named in each report.

Testing the Research Questions

Data Analysis for Research Question 1

What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?

A scatterplot and multiple regression were used to test the linear assumption, or more specifically that there was not a non-linear relationship between state and national levels of proficiency (Lomax & Hahs-Vaughn, 2012). Based on the visual evidence of the scatter plot, a weak positive linear relationship existed between the two variables.

However on further investigation, a test of non-linear trends indicated that there was a significant quadratic (non-linear) relationship between the state and NAEP levels of proficiency. The multiple regression summary statistics relating state and NAEP proficiency rates are reported in Table 3.

Table 3

Multiple Regression Summary Statistics Relating State and NAEP Proficiency Rates

Curve Estimates for State Proficiency Rates	Unstandardized Coefficients		Standardized Coefficients		
	B	*SE*	β	*t*	*p*
Linear	.400	.228	.245	1.752	.086
Quadratic	-.075	.022	-3.035	-3.384	.001
Cubic	-.003	.002	-6.139	-1.303	.199

Note. NAEP = National Assessment of Educational Progress.

Because the two proficiency rates were not linearly related to each other, a multiple regression was used to test the relationship between the two assessments. As indicated by the statistics in Table 3, even after accounting for the linear relationship, there was a significant quadratic relationship between the percentage of students who were proficient on the state examinations and those who were proficient on the NAEP examination. As illustrated in Figure 1, the majority of states with very low (16%-28%) proficiency rates on the NAEP had moderate (58%-76%) proficiency rates on the state

assessments. The states with moderately low (29%-41%) proficiency rates on the NAEP had moderate (58%-75%) and high (77%-92%) proficiency rates on the state assessments. The seven states with low (42%-51%) proficiency rates on the NAEP ranged from low (39%-57%) to moderate (58%-76%) proficiency rates on the state assessments, with no scores in the high (77%-92%) proficiency rates on the state assessments .

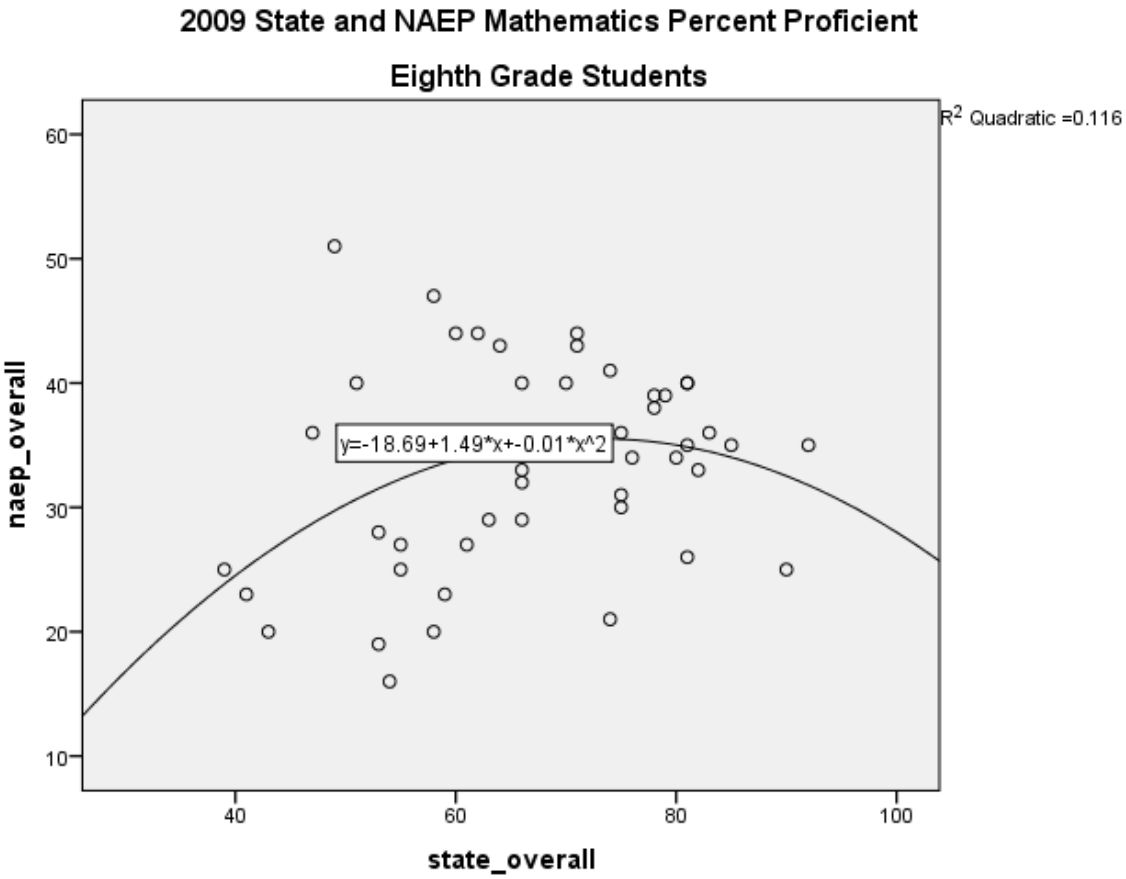


Figure 1. Scatter Plot relating the Percentage of Students Proficient on State and NAEP Assessments.

Data Analysis for Research Question 2

What comparisons can be made between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, and South) identified by the NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage of students who were proficient on the state and NAEP assessments. The within-subjects factor was type of test (NAEP or state), and the between-subjects factor was region (Midwest, Northeast, West, and South). There was a significant interaction between type of test and region ($F(3,46) = 5.3$, $p = .002$), indicating that the difference between state and NAEP proficiency percentages depended on the region. Therefore, a series of paired samples *t*-tests were conducted to compare the rates of proficiency in mathematics state assessments and the NAEP for eighth-grade students in each of the four census defined regions (Midwest, Northeast, West, and South). The variations in the means of each test ranged from slightly (0%-20%) different to moderately (21%-40%) different to significantly (41% and above) different. The effect sizes in all of the regions were large. The results of the paired samples t-test are synthesized in Table 4.

Table 4

Means, Standard Deviations, t-scores, CI and Effect Sizes for Overall Students' Percentages Proficient

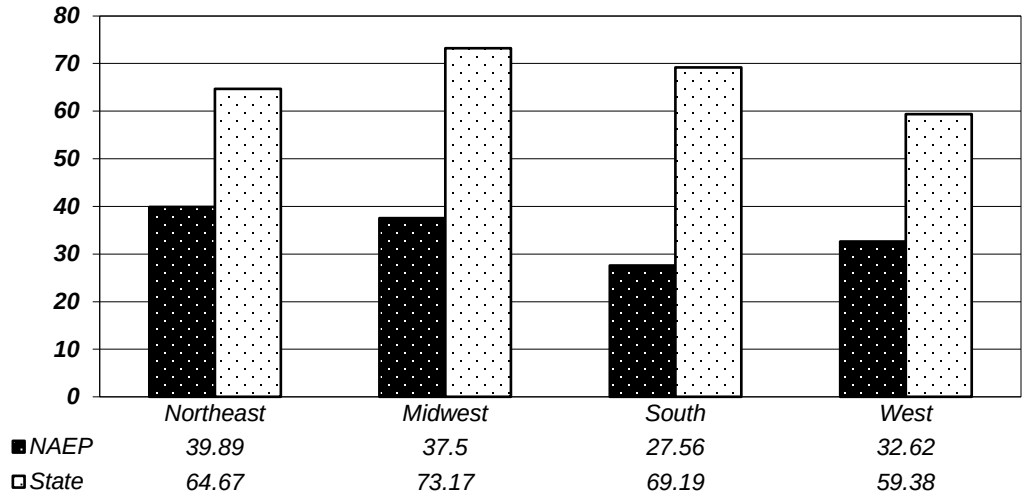
	State		NAEP				95% CI		Cohen's
Region	*M*	*SD*	*M*	*SD*	*T*	*p*	*LL*	*UL*	*d*
Midwest	73.17	11.40	37.5	4.52	9.01	<.001	26.96	44.38	2.50
Northeast	64.67	11.83	39.89	6.77	5.28	.001	13.96	35.59	3.05
South	69.19	12.19	27.56	6.84	15.79	<.001	36.00	47.24	3.95
West	59.38	13.33	32.62	7.51	9.79	<.001	20.81	32.73	2.71

Note. NAEP = National Assessment of Educational Progress; CI = confidence interval; LL = lower limit; UL = upper limit.

The results indicated that among students in the Midwest region, the mean state proficiency rate (*M* = 73.17, *SD* = 11.4) was almost twice the mean for NAEP proficiency rates (*M* = 37.5, *SD* = 4.52) with a large effect size. Between students in the Northeast region, the mean state proficiency rate (*M* = 64.67, *SD* = 11.83) was 24.78 percentage points greater than the mean for NAEP proficiency rates (*M* = 39.89, *SD* = 6.77) with a large effect size. The mean state proficiency rate (*M* = 69.19, *SD* = 12.19) for students in the Southern region was more than twice the mean for NAEP proficiency rates (*M* = 27.56, *SD* = 6.84) and the largest effect size of all the regions. Among students in the Western region, the mean state proficiency rate (*M* = 59.38, *SD* = 13.33)

was significantly greater than the mean for NAEP proficiency rates (M = 32.62, SD = 7.51) and the effect size was large once again.

The Southern region had the greatest difference between the means of the two assessments at 41.6%. The Midwest region had the next highest mean difference (35.7%), followed by the Western region with a mean difference of 26.8 %; and finally the Northeastern region had the smallest mean difference of 24.8%. Figure 2 illustrates the differences in mathematics proficiency rates between each region for the state and NAEP assessments. (Midwest, Northeast, West, and South).



Note. Percentages identified as proficient and above in mathematics on state assessments and on the NAEP, disaggregated by the four census-defined regions.

Figure 2. Mean Percentages of Proficient Eighth-grade Students by Census Region

Data Analysis for Research Question 3

What relationship exists, if any, between the percentage of low socioeconomic (SES) eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census defined regions (Midwest, Northeast, West, and South) identified by NAEP in 2009?

A two-factor split plot ANOVA was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among low socioeconomic students. There was a significant interaction between type of test and region (F (3,46)=5.06, p = .004), indicating that among low socioeconomic students the difference between state and NAEP proficiency percentages depended on the region. Hence, several paired samples t-tests were performed to further investigate the difference between the two assessments for low socioeconomic students in the four regions. The results of the analyses are shown in Table 5.

Table 5

Means, Standard Deviations, t-scores, CI and Effect Sizes for Low SES Students' Percentages Proficient

	State		NAEP				95% CI		Cohen's
Region	*M*	*SD*	*M*	*SD*	*t*	*p*	*LL*	*UL*	*d*
Midwest	58.8	14.4	19.60	4.00	8.8	<.001	29.1	49.2	2.54
Northeast	46.0	14.2	20.11	5.42	4.7	<.05	13.2	38.6	1.57
South	58.4	14.1	14.60	3.70	13.1	<.001	36.7	50.8	3.39
West	44.6	14.3	18.2	4.80	7.5	<.001	18.7	34.1	2.71

Note. SES = Socioeconomic status; CI = confidence interval; LL = lower limit; UL = upper limit.

The low socioeconomic students in the Midwest region had a mean state proficiency rate (*M* = 58.8, *SD* = 14.4) that was nearly three times greater than the mean for NAEP proficiency rates (*M* = 19.6, *SD* = 4.1) with a large effect size.

Among low socioeconomic students in the Northeast region, the mean state proficiency rate (*M* = 46, *SD* = 14.2) was greater than the mean for NAEP proficiency rates *(M* = 20.11, *SD* = 5.42) with a large effect size.

In the Southern region, the mean state proficiency rate (*M* = 58.4, *SD* = 14.1) was four times greater than the mean for NAEP proficiency rates (*M* = 14.6, *SD* = 3.7) with the largest effect size of the four regions.

The outcome for low socioeconomic students in the Western region, was a mean state proficiency rate (M = 44.6, SD = 14.3) that was more than two times greater than the mean for NAEP proficiency rates (M = 18.2, SD = 4.8) and a large effect size. In all of the regions the disparity between the two assessments and the effect sizes were large for low socioeconomic students.

Data Analysis for Research Question 4

What relationship exists, if any, between the percentage of white, eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot ANOVA was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among white students. There was a significant interaction between type of test and region ($F(3,46)=5.81$, p = .002). A series of paired samples t-tests were conducted to examine the differences between the results of the tests for white students based on region. The results of the t-tests are shown in Table 6.

The results showed that for white students in the Midwest region, the mean state proficiency rate (M = 78.3, SD = 10.8) was higher than the mean for NAEP proficiency rates (M = 49.9, SD = 4.5) and the effect size was large.

For white students in the Northeast region, the mean state proficiency rate (M = 70.8, SD = 13.9) was higher than the mean for NAEP proficiency rates (M = 45.6, SD = 7.7) and the effect size was large.

Table 6

Means, Standard Deviations, t-scores, CI and Effect Sizes for White Students' Percentages Proficient

	State		NAEP				95% CI		Cohen's
Region	*M*	*SD*	*M*	*SD*	*t*	*p*	*LL*	*UL*	*d*
Midwest	78.3	10.8	49.9	4.5	10.1	<.01	27.7	43.1	2.92
Northeast	70.8	13.9	45.6	7.7	5.3	<.01	14.3	36.1	1.77
South	77.6	12.00	37.0	10.6	17.4	<.01	35.6	45.5	4.35
West	67.5	11.9	41.3	5.1	10.0	<.01	20.5	31.9	2.77

Note. NAEP = National Assessment of Educational Progress; CI = confidence interval; LL = lower limit; UL = upper limit.

In the Southern region, the mean for state proficiency rate (M = 77.6, SD = 12) of white students was 40.6 percentage points more than the mean NAEP proficiency rates (M = 37, SD = 10.6) with the largest effect size of all four regions.

According to the statistics, white students in the Western region had a mean state proficiency rate (M = 67.5, SD = 11.9) that was greater than the mean NAEP proficiency rate (M = 41.3, SD = 5.1) and the effect size was large.

Data Analysis for Research Question 5

What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analyses of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among black students. The within-subjects factor was type of test (NAEP or state), and the between-subjects factor was region (Midwest, Northeast, West, and South). There was a significant interaction between type of test and region ($F(3,38)=2.93$, $p = .046$), which purported that the difference between state and NAEP proficiency percentages depended on the region.

Four paired samples *t*-tests were conducted to compare the rates of proficiency in mathematics state assessments and the NAEP for black, eighth-grade students for each of the four census defined regions (Midwest, Northeast, West, and South). Table 7 contains the results of the paired samples *t*-tests.

Table 7

Means, Standard Deviations, t-scores, CI and Effect Sizes for Black Students' Percentages Proficient

	State		NAEP				95% CI		Cohen's
Region	*M*	*SD*	*M*	*SD*	*t*	*p*	*LL*	*UL*	*d*
Midwest	48.6	18.4	10.8	2.9	6.2	<.01	24.1	51.5	1.96
Northeast	42.4	15.6	14.0	4.9	4.2	<.05	11.8	45.1	1.58
South	54.0	16.3	10.8	3.3	11.4	<.01	35.2	51.3	2.85
West	40.7	12.8	15.3	4.6	6.2	<.01	15.9	34.8	2.07

Note. NAEP = National Assessment of Educational Progress; CI = confidence interval; LL = lower limit; UL = upper limit.

Black students in the Midwest region had mean state proficiency rates (M = 48.6, SD = 18.4) that were significantly greater than the mean for NAEP proficient data (M = 10.8, SD = 2.9) and there was a large effect size.

The results indicated that among black students in the Northeast region, the mean state proficiency rate (M = 42.4, SD = 15.6) was three times greater than the mean for NAEP proficient data (M = 14, SD = 4.9) with a large effect size.

The results also indicated that among black students in the Southern region, the mean state proficiency rates (M = 54, SD = 16.3) was five times greater than the mean for

NAEP proficiency rates (M = 10.8, SD = 3.3) with the largest effect size of the four regions.

In the Western region, the mean state proficiency rate data (M = 40.7, SD = 12.8) for black students was two times larger than the mean for NAEP proficient data (M = 15.3, SD = 4.6) with a large effect size.

Data Analysis for Research Question 6

What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analyses of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among Hispanic students. The within-subjects factor was type of test (NAEP or state), and the between-subjects factor was region (Midwest, Northeast, West, and South). There was a significant interaction between type of test and region (F (3, 40) =3.16, p = .035), indicating that the difference between state and NAEP proficiency percentages depended on the region.

A series of paired samples t-tests were conducted to compare the rates of proficiency in mathematics state assessments and the NAEP for Hispanic, eighth-grade

students for each of the four census defined regions (Midwest, Northeast, West, and South). The statistics are reported in Table 8.

Table 8

Means, Standard Deviations, t-scores, CI and Effect Sizes for Hispanic Students' Percentages Proficient

	State		NAEP				95% CI		Cohen's
Region	*M*	*SD*	*M*	*SD*	*t*	*p*	*LL*	*UL*	*d*
Midwest	58.1	16.1	18.8	7.0	6.1	<.01	25.0	53.6	1.84
Northeast	46.4	17.5	17.1	5.2	4.3	<.01	2.6	46.0	1.62
South	65.7	12.7	19.5	5.0	13.1	<.01	38.5	53.8	3.63
West	45.5	12.7	16.5	5.5	8.7	<.01	21.7	36.2	2.41

Note. NAEP = National Assessment of Educational Progress; CI = confidence interval; LL = lower limit; UL = upper limit.

The results indicated that among Hispanic students in the Midwest region, the mean state proficiency rates (M = 58.1, SD = 16.1) was three times greater than the mean for NAEP proficiency rates (M = 18.8, SD = 7) with a large effect size.

The results also indicated that among Hispanic students in the Northeast region, the mean state proficiency rates (M = 46.4, SD = 17.5) was more than 2.5 times greater than the mean for NAEP proficiency rates (M = 17.1, SD = 5.2) and a large effect size.

Among Hispanic students in the Southern region, the mean state proficiency rates (M = 65.7, SD = 12.7) was significantly greater than the mean for NAEP proficiency rates (M = 19.5, SD = 5) with the largest effect size of all four regions.

In the Western region, the mean state proficiency rate (M = 45.5, SD = 12.7) for Hispanic students was more than three times higher than the mean for NAEP proficiency rates (M = 16.5, SD = 5.5) with a large effect size.

Summary

This chapter contained an introduction, descriptive statistics and an analysis of the data organized around each of the six research questions. The research questions were tested and the statistical outcomes were presented.

The results of Research Question 1 revealed a large difference between the means of the two tests in a majority of the states. There was a non-linear relationship between the NAEP and state assessments. Chapter 5 delves deeper into the data in the summary and discussion of findings followed by educational significance and recommendations for future research.

CHAPTER 5
SUMMARY, DISCUSSION, AND RECOMMENDATIONS

Introduction

Educational reform through accountability is not a new concept in the education field. However the idea of sanctions for schools that do not meet the predetermined percentage proficient on high stakes tests is a new concept that has led to controversy over the rigor of some states' standards. The intent of this study was to compare and analyze the percentage proficient and above results of the eighth-grade mathematics state assessments against proficiency levels of the eighth-grade mathematics National Assessment of Education Progress (NAEP) assessment. Because each state has been able to select its own testing system and set its own passing scores, there was no direct way to compare the proficiency levels established by one state against the other. The NAEP is a common measurement used by all states. Thus, by comparing the percentage of students achieving proficiency on state tests with the percentage achieving proficiency on the NAEP, one can determine the rigor of each state's tests and standards (Peterson & Hess, 2005). This research study was also intended to add to the body of knowledge that existed concerning the evolution of accountability and high stakes testing.

In the previous chapter, the presentation and analysis of data were reported. Chapter 5 consists of a summary of the study, discussion of the findings, educational significance, recommendations for further research, and a summary. The chapter presents a succinct review of the topics included in this study organized by the six research questions and provides further commentary on the statistical results discovered.

Summary of the Study

The purpose of this study was to provide all stakeholders and policy makers of the United States with information regarding the relationship between state mathematics assessments and NAEP mathematic assessments for eighth-grade students in 2009. The mathematics scores of eighth-grade students on the 2009 state mathematics assessments and the 2009 National Assessment of Educational Progress (NAEP) mathematics assessment were compared by state, AYP sub groups, and census regions.

In 2009, the National Center for Educational Statistics (NCES), a division of the United States Department of Education (USDOE), administered and scored the NAEP assessment in mathematics by randomly selecting eighth-grade students in all 50 states to achieve a typical population of the country including approximately 30 students per school (National Center for Educational Statistics, 2010b). The data were reported as part of the Nation's Report Card accessed on the NCES website. Also, during the 2008-2009 school year, in compliance with NCLB, all 50 states conducted eighth grade mathematics assessments and reported data to the USDOE. The state assessment data for this study were collected through the Department of Education's Consolidated State Performance Reports (U.S. Department of Education, 2010).

This study included six research questions:

1. What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?

2. What comparisons can be made between the percentage of eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by the NAEP in 2009?
3. What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?
4. What relationship exists, if any, between the percentage of white eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?
5. What relationship exists, if any, between the percentage of black eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?
6. What relationship exists, if any, between the percentage of Hispanic eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

Research Question 1 was answered using the results of a multiple regression to determine the type of relationship that existed between the overall percentages of

proficient scores on the two assessments. Research Questions 2 through 6 were answered using a two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments overall and by the various AYP subgroups. The within-subjects factor was type of test (NAEP or state), and the between-subjects factor was region (Midwest, Northeast, West, and South).

Discussion of Findings

In March of 2009, President Barack Obama in his remarks to the Hispanic Chamber of Commerce speech, put the issue of state standards on the national agenda by stating,

> That's why I'm calling on states that are setting their standards far below where they ought to be to stop low-balling expectations for our kids. The solution to low test scores is not lowering standards, it's tougher, clearer standards. Standards like those in Massachusetts, where 8th graders are now tying for first, first in the whole world in science (para. 21).

The President was bringing into the discussion the discrepancy between the assessments, standards and proficiency levels in the 50 states. He accused states of having seemingly low expectations for students, and he called for tougher, clearer standards (Peterson & Lastra-Anadon, 2010). This was a push towards the desire for national standards by many educational policy makers and stakeholders. In March 2010, Secretary of Education Duncan added to the political banter and blamed educators for lowering the

bar for proficiency in order to meet the requirements set by the federal education law, No Child Left Behind (NCLB). NCLB required that all students be proficient in reading and math by the year 2014 (Peterson & Lastra-Anadon, 2010). This preliminary information provides a context for the following summary and discussion of the findings, organized around the six research questions, in the present study.

Research Question 1

What relationship exists, if any, between the percentage of eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP in 2009?

Fourteen states had very low (16%-28%) proficiency rates on the NAEP assessment. Seven of those states in the very low (16%-28%) NAEP proficiency rating continued the pattern with low (39%-57%) state assessment proficiency rates, and five states in the very low group had moderate (58%-76%) proficiency rates on the state assessments. These results were not surprising. However, two states with very low (16%-28%) NAEP proficiency rates had high (77%-92%) proficiency rates on the state assessments. The two states with the largest difference between the proficiency levels of the two assessments were those of Georgia and Tennessee. Georgia had a difference of 55 percentage points and Tennessee had a difference of 65 percentage points, the largest discrepancy of all fifty states. There could be several reasons for the enormity of the difference between the two assessments, (a) poor alignment of state standards with the NAEP assessment, (b) low cut scores for proficiency on the state assessments and (c) low stakes vs. high stakes testing motivating factors.

The vast majority of the states fell in the moderately low (29%-41%) proficiency rates range on the NAEP assessment. The breakdown of those 29 states in the moderately low (29%-41%) NAEP range were: three states with low (39%-57%), 15 states in the moderate (58%-75%) state proficiency range and eleven in the high (77%-92%) proficiency rates on the state assessments.

The final seven states were in the low (42%-51%) proficiency range on the NAEP assessment. Six of those states fell in the moderate (58%-76%) proficiency range on the state assessments. The state of Massachusetts had the smallest difference between the proficiency levels of the two assessment with a difference of 2 percentage points. The proficiency level on the NAEP was 51% proficient, the highest of all the 50 states, and the proficiency level on their state assessment was 49%. According to the data, Massachusetts has the most stringent state standards and were closely aligned with the NAEP assessment. Peterson and Lastra-Anadon (2010) reported that in 2009, five states (Hawaii, Massachusetts, Missouri, New Mexico and Washington) had set their standards at or close to world-class levels despite the incentive to lower expectations to avoid sanctions and meet the goal of all children proficient by 2014.

Research Question 2

What comparisons can be made between the percentage of eighth-grade students identified as proficient and above in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by the NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analysis of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments. The results indicated that region moderates the difference in proficiency percentages on the two assessments and a comparison was made between the average performances of students on the two mathematics assessments.

The results showed that the sixteen states from the Southern region had the greatest difference between the means of the two assessments at 41.6%. The 12 states that comprised the Midwest region had the next highest mean difference of 35.7 %. The Western region, of 13 states, had a mean difference of 26.8 %. The nine states that make up the Northeastern region had the smallest mean difference of 24.8%.

Research Question 3

What relationship exists, if any, between the percentage of low socioeconomic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by the NAEP in 2009?

A two-factor split plot ANOVA was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among low socioeconomic students. There was a significant interaction between type of test and region. The difference between the two assessments for low socioeconomic students in the four census-defined regions was explored further.

In the nine states that make up the Northeast region, among low socioeconomic students, the difference in the mean proficiency rates of the two assessments was 25.9%. The 14 states in the Southern region that reported on the low socioeconomic subgroup saw a mean difference between the two assessments of 43.8%. The outcome for low socioeconomic students in the 13 states that comprised the Western region, was a 26.4% difference in the means of both assessments. The Midwest region had a mean difference of 39.2% between the two tests among low socioeconomic students. The Northeast region once again had the least discrepancy between the two tests, followed closely by the Western region, then the Midwest region and lastly the Southern region.

Research Question 4

What is the difference between the percentage of white eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot ANOVA was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among white students. There was a significant interaction between type of test and region. Upon further investigation of the mean differences of the two assessments among white students, once again the Northeast region had the smallest difference between the two assessments at 25.2%, followed by the Western region with a difference of 26.2%. The Midwest region was third with a mean difference of 28.4%. White students in the Midwest, had the greatest proficient percentage on the NAEP and state assessment. This was the only subgroup in the Midwest region to have a higher percentage of proficient scores than any other region.

Research Question 5

What is the difference between the percentage of black eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analyses of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments

among black students. There was a significant interaction between region and type of test. The subgroup was analyzed further.

The Western region lead the way with only a 25.4% difference in the means of the two assessments and also the highest percentage proficient on the NAEP. The Northeast region was not far behind with an average difference of 28.4%, followed by the Midwest region reporting a 37.8% mean difference. The Southern states had the largest discrepancy between the two assessments with 43.2%; however, the mean for the NAEP of 10.8% proficient was the same as the Midwest region.

Research Question 6

What is the difference between the percentage of Hispanic eighth-grade students identified as proficient in mathematics on state assessments and on the NAEP, in the four census-defined regions (Midwest, Northeast, West, South) identified by NAEP in 2009?

A two-factor split plot (one within-subjects factor and one between-subjects factor) analyses of variance (ANOVA) was conducted to determine if region moderated the difference between the percentage proficient on the state and NAEP assessments among Hispanic students. Region moderated the difference between the two assessments. Therefore further analysis was used to describe the results.

The smallest difference between the means of the two assessments occurred in the Western region at 29%. This was followed by the Northeast region with a difference of 29.3 %, and the Midwest region with a variance of 39.3% between the two assessments in the AYP subgroup of Hispanic students. The Southern region had the largest difference

(46.2%) between the two assessments. It should be noted that for the first time in this study the Southern region had the highest percentage proficient on both assessments for Hispanic students.

Educational Significance

The research upon which this study was based purported that there was a statistically significant discrepancy in the determination of proficient across the United States on accountability assessments. President Obama (2009) stated,

> Let's challenge our states to adopt world-class standards that will bring our curriculums to the 21st century. Today's system of 50 different sets of benchmarks for academic success means 4th grade readers in Mississippi are scoring nearly 70 points lower than students in Wyoming, and they're getting the same grade. Eight of our states are setting their standards so low that their students may end up on par with roughly the bottom 40 percent of the world. (para. 20)

There has been a push to make each state hold their students to more rigorous standards, align content standards and set similar proficiency levels on state assessments.

The findings indicated that there needed to be a way to compare results in each state to a uniform measure administered to all states. At the time of the study, the NAEP assessment was the one assessment common to all states and, therefore, was used. Findings allow state policy makers and stakeholders to make decisions as to the standards of the standards in mathematics and language arts (Linn, 2005). The implication is that

the difference between percentage proficient between the random sample of NAEP test takers and those taking state assessments calls for a common assessment that reflects the same standards nationally and measures students at the same level of rigor.

The review of literature revealed that many policy makers have taken the difference between the two assessments and used it to condemn some states for setting the proficiency bar too low and challenging these states to increase the rigor of their standards (Peterson & Lastra-Anadon, 2010). Policy makers, administrators and educators should have a solid grasp on statistics and the content of the assessments, which are different in some states, to make well-informed decisions pertaining to the use or abuse of high stakes testing data.

Be cautious of inferring too much from the results of this study. The results may mean that some states need to set higher standards for their students, but it could also mean that the policy makers are muddying the waters by tying funding to the backs of students, teachers and administrators. Perhaps the decision makers should retreat from the sanctions and punishments to get a clearer picture of the state of education in American schools.

An interesting way to look at the disparity in percentage proficient between the two assessments is to investigate the political backdrop from which the two sets of achievement levels were developed. The NAEP was developed as a low stakes test used to gauge what American school children know and can do, whereas the state achievement levels were developed for high stakes testing with an unrealistic goal of 100% proficient by 2014. Though the NAEP achievement levels have been seen as visionary, the state

achievement goals are more realistic and attainable in the specified time limit. "The evaluation of whether NAEP's achievement levels are too stringent should take into account the policy context in which NAEP's achievement levels were set relative to the NCLB policy environment in which achievement levels were set for state assessments" (Lane et al., 2009, p.313).

Unfortunately, the NCLB policy has had some unintended negative side effects for students in various subgroups due to the AYP accountability measure. "Representation of subgroups across states varies considerably as well as the inclusion and exclusion rates for students with disabilities, impacting the validity of the use of NAEP results for state by state comparisons and to verify state assessment results" (Lane et al., p. 316).

Many new topics in education have come about since 2001, e.g., NCLB, Race to the Top (RttT) and common core state standards. Many times these efforts have negative effects on true academic achievement. If anything has been learned throughout the years of constant reform efforts, it is that slow and steady wins the race, not knee jerk reactions to political pressures. Educational leaders should not waiver in their goals enveloped in political currents that are constantly bashing the state of education. Set high attainable standards for all children based on research, and most importantly provide the necessary resources needed for student learning.

Educators have been asked to meet the students where they are, to develop trusting relationships, and improve student achievement. Education has been moving away from that type of fostering or collegial environment to one of win at all costs. The

amount of mistrust at each level of the educational pyramid is an issue that needs to be addressed to improve student achievement. Schools are more than just test scores, and leaders should demand that policy makers understand the research for all students to achieve at their optimum level. Policy makers and educational leaders should set high and attainable standards for all students without losing sight of what is best for the students overall.

Recommendations for Future Research

The following are topics to consider for future research related to this dissertation and accountability in the field of education:

1. A quantitative analysis of professional development opportunities and best practices used in various states with high and low proficiency rates.
2. A quantitative study of the impact of high stakes assessments on improving the quality of education and/or learning outcomes.
3. A qualitative study of creative ways to close the achievement gaps that exist for various subgroups.
4. A quantitative study to find a common meaning of the term proficiency among the states through alignment of content standards and cut scores that would allow for a more meaningful comparison of student achievement.
5. A quantitative study to determine the extent that the Common Core State Standards movement results in less disparity among states for student achievement.

Summary

The findings of this study add to the body of knowledge and research in the area of educational reform and accountability through state to state comparisons of the rigor of the 2009 standards in 8th grade mathematics. Through the use of NAEP, a common measurement to all states, a comparison of the stringency of state standards was ascertained for eighth-grade students in mathematics in 2009. The final chapter of this study has included a brief synopsis of the various components of the research, discussion of the findings, educational significance and recommendations for future research.

APPENDIX A
INSTITUTIONAL REVIEW BOARD APPROVAL

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

From : **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To : **Laurel J. Robinson**

Date : **June 22, 2011**

Dear Researcher:

On 6/22/2011 the IRB determined that the following proposed activity is not human research as defined by DHHS regulations at 45 CFR 46 or FDA regulations at 21 CFR 50/56:

Type of Review:	Not Human Research Determination
Project Title:	A COMPARISON OF EIGHTH GRADE STATE MATHEMATICS SCORES BY STATE AND BY THE FOUR CENSUS-DEFINED REGIONS TO NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS (NAEP) SCORES FOR MATHEMATICS
Investigator:	Laurel J Robinson
IRB ID:	SBE-11-07721
Funding Agency:	
Grant Title:	
Research ID:	NA

University of Central Florida IRB review and approval is not required. This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are to be made and there are questions about whether these activities are research involving human subjects, please contact the IRB office to discuss the proposed changes.

On behalf of Kendra Dimond Campbell, MA, JD, UCF IRB Interim Chair, this letter is signed by:

Signature applied by Joanne Muratori on 06/22/2011 11:33:43 AM EDT

Joanne Muratori

IRB Coordinator

APPENDIX B
NCES 2009 STATE SNAPSHOT REPORT (ALABAMA)

ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

The Nation's Report Card Mathematics 2009

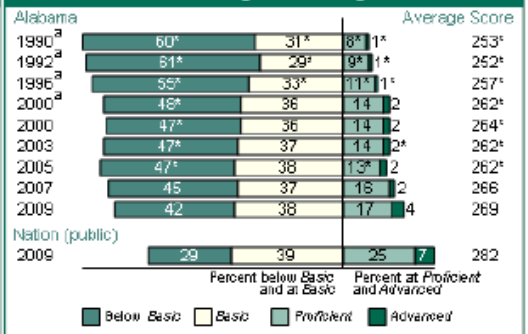
Snapshot State Report

Alabama
Grade 8
Public Schools

Overall Results

- In 2009, the average score of eighth-grade students in Alabama was 269. This was lower than the average score of 282 for public school students in the nation.
- The average score for students in Alabama in 2009 (269) was not significantly different from their average score in 2007 (266) and was higher than their average score in 1990 (253).
- In 2009, the score gap between students in Alabama at the 75th percentile and students at the 25th percentile was 49 points. This performance gap was not significantly different from that of 1990 (48 points).
- The percentage of students in Alabama who performed at or above the NAEP *Proficient* level was 20 percent in 2009. This percentage was not significantly different from that in 2007 (18 percent) and was greater than that in 1990 (9 percent).
- The percentage of students in Alabama who performed at or above the NAEP *Basic* level was 58 percent in 2009. This percentage was not significantly different from that in 2007 (55 percent) and was greater than that in 1990 (40 percent).

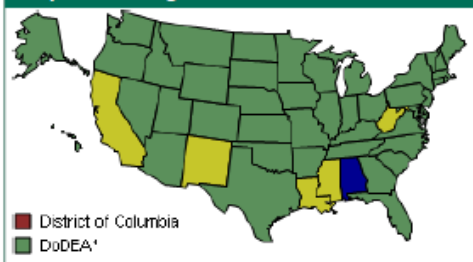
Achievement-Level Percentages and Average Score Results



* Significantly different ($p < .05$) from state's results in 2009.
a Accommodations not permitted.

NOTE: Detail may not sum to totals because of rounding.

Compare the Average Score in 2009 to Other States/ urisdictions

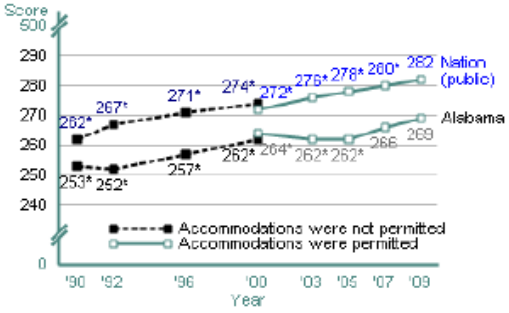


* Department of Defense Education Activity schools (domestic and overseas).

In 2009, the average score in **Alabama** was

- lower than those in 45 states/jurisdictions
- higher than that in 1 state/jurisdiction
- not significantly different from that in 5 states/jurisdictions

Compare the Average Score to Nation (public)



* Significantly different ($p < .05$) from 2009.

Results for Student Groups in 2009

Reporting Groups	Percent of students	Avg. score	Percentages at or above *Basic*	Percentages at or above *Proficient*	Percent at *Advanced*
Gender *					
Male 51		268	58	21	4
Female	49	269	58	20	3
Race/Ethnicity					
White 60		280	72	29	5
Black 35		248	34	6	1
Hispanic 3		260	49	10	#
Asian/Pacific Islander	1	‡	‡	‡	‡
American Indian/Alaska Native	1	‡	‡	‡	‡
National School Lunch Program *					
Eligible 50		255	44	10	1
Not eligible	50	282	73	31	6

Rounds to zero. ‡ Reporting standards not met.

NOTE: Detail may not sum to totals because of rounding, and because the "Information not available" category for the National School Lunch Program, which provides free/reduced-price lunches, and the "Unclassified" category for race/ethnicity are not displayed.

Score Gaps for Student Groups

- In 2009, female students in Alabama had an average score that was not significantly different from that of male students. This performance gap was not significantly different from that in 1990 (2 points). '
- In 2009, Black students had an average score that was 32 points lower than that of White students. This performance gap was not significantly different from that in 1990 (30 points).
- In 2009, Hispanic students had an average score that was 20 points lower than that of White students. Data are not ' reported for Hispanic students in 1990, because reporting standards were not met.
- In 2009, students who were eligible for free/reduced-price school lunch, an indicator of poverty, had an average score that was 27 points lower than that of students who were not eligible for free/reduced-price school lunch. This performance gap was not significantly different from that in 1996 (33 points).

NOTE: Statistical comparisons are calculated on the basis of unrounded scale scores or percentages.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), various years, 1990–2009 Mathematics Assessments.

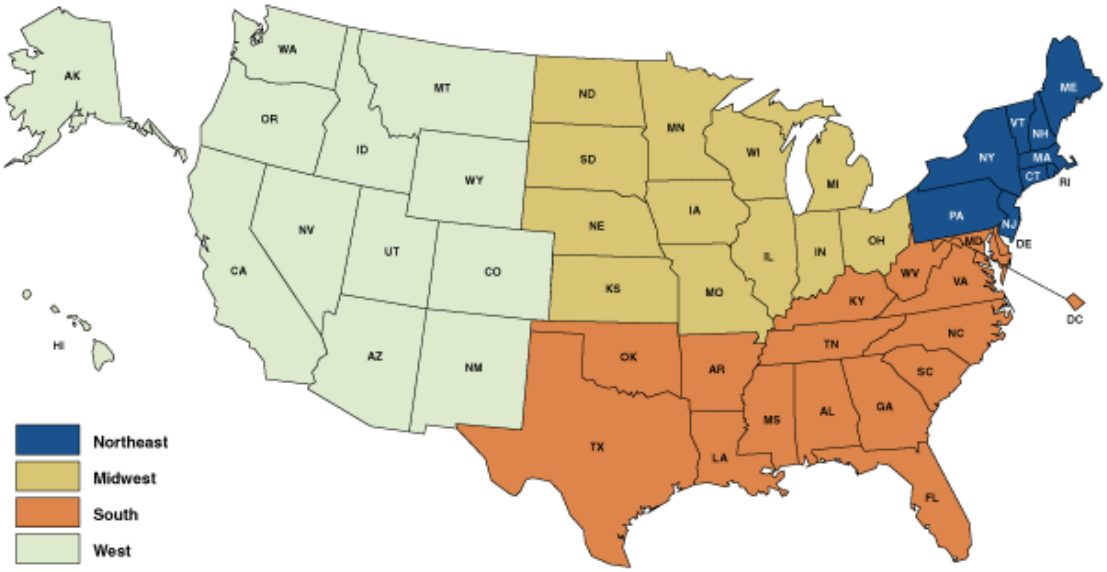
APPENDIX C
SY 2008-2009 CSPR REPORT (ALABAMA)

1.3.1.6 Student Academic Achievement in Mathematics -Grade 8

Grade 8	# Students Who Received a Valid Score and for Whom a Proficiency Level Was Assigned	# Students Scoring at or Above Proficient	Percentage of Students Scoring at or Above Proficient
All students	57,055	41,939	73.5
American Indian or Alaska Native	562	429	76.3
Asian or Pacific Islander	649	598	92.1
Black, non-Hispanic	20,282	12,131	59.8
Hispanic	1,894	1,264	66.7
White, non-Hispanic	33,526	27,413	81.8
Children with disabilities (IDEA)	6,339	1,954	30.8
Limited English proficient (LEP) students	954	491	51.5
Economically disadvantaged students	30,163	19,021	63.1
Migratory students	74	48	64.9
Male	29,374	20,644	70.3
Female	27,680	21,294	76.9
Comments: In addition with respect to special education students when their assessment data was entered for 003, 004, 093 all students who took the alternate assessment were included. In the case of 075, 078 and 079 all sub populations are run through an automated query based on the student demographics in the enrollment record, therefore fewer students were counted in the special ed subpopulation who took the alternate assessment becuase their enrollment record did not notate them as being special ed.			

Source – Initially populated from EDFacts. See Attachment D: CSPR & EDFacts Data Crosswalk. 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 collection tool.

APPENDIX D
NAEP CENSUS-DEFINED REGIONS

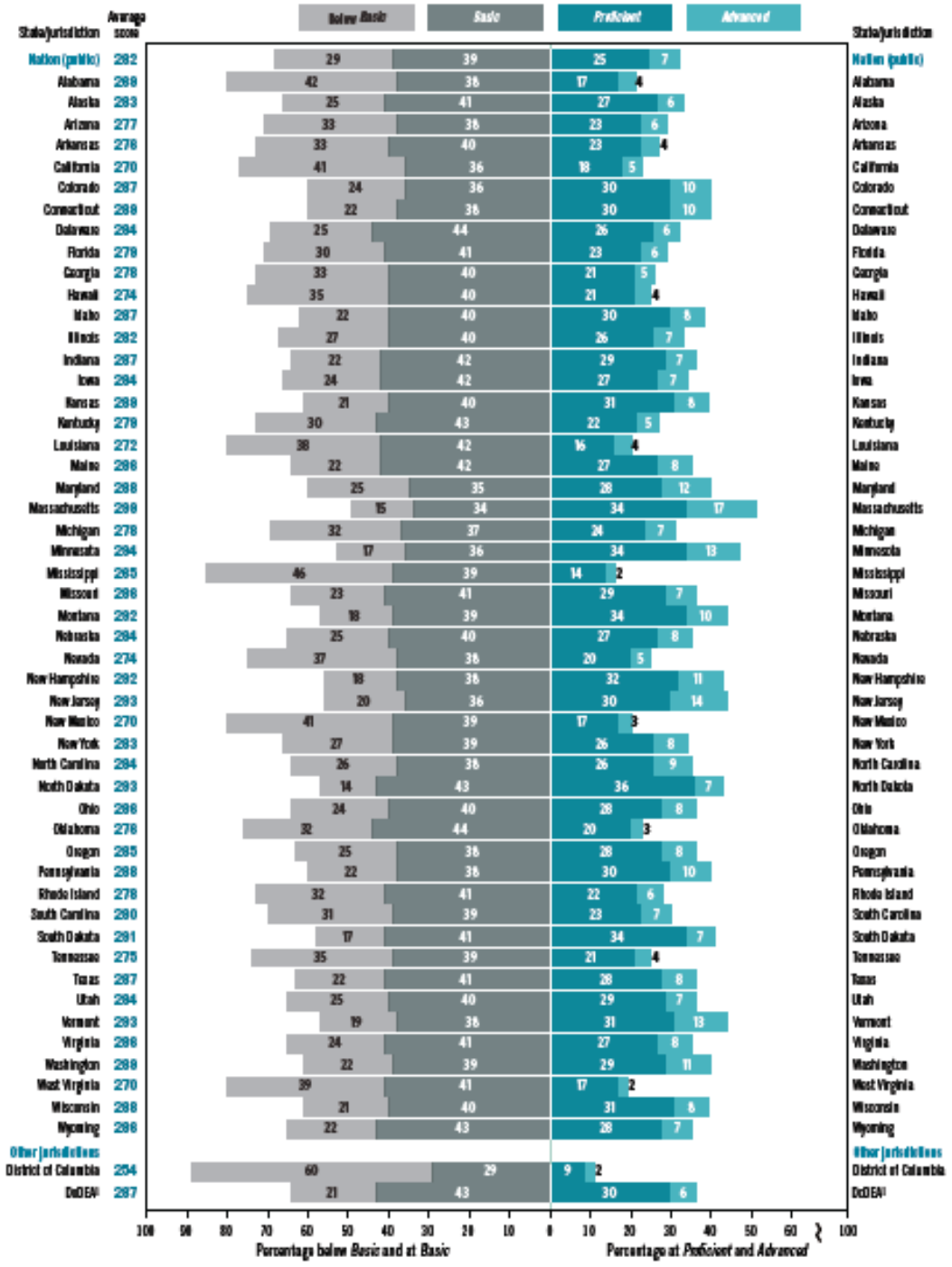


- *Northeast*: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont;
- *South*: Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, 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.

APPENDIX E
2009 NAEP MATHEMATICS AVERAGE SCORES AND ACHIEVEMENT-LEVEL RESULTS BY STATE/JURISDICTION

Figure 23. Average scores and achievement-level results in NAEP mathematics for eighth-grade public school students, by state/jurisdiction: 2009



[1] Department of Defense Education Activity (overseas and domestic schools).
NOTE: The shaded bars are graphed using unrounded numbers. Detail may not sum to totals because of rounding.
SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Mathematics Assessment.

APPENDIX F
THE NAEP MATHEMATICS ACHIEVEMENT LEVELS

***Basic* (262)**

Eighth-grade students performing at the *Basic* level should exhibit evidence of conceptual and procedural understanding in the five NAEP content areas. This level of performance signifies an understanding of arithmetic operations—including estimation—on whole numbers, decimals, fractions, and percents.

Eighth-graders performing at the *Basic* level should complete problems correctly with the help of structural prompts such as diagrams, charts, and graphs. They should be able to solve problems in all NAEP content areas through the appropriate selection and use of strategies and technological tools—including calculators, computers, and geometric shapes. Students at this level also should be able to use fundamental algebraic and informal geometric concepts in problem solving.

As they approach the *Proficient* level, students at the *Basic* level should be able to determine which of the available data are necessary and sufficient for correct solutions and use them in problem solving. However, these eighth-graders show limited skill in communicating mathematically.

***Proficient* (299)**

Eighth-grade students performing at the *Proficient* level should apply mathematical concepts and procedures consistently to complex problems in the five NAEP content areas.

Eighth-graders performing at the *Proficient* level should be able to conjecture, defend their ideas, and give supporting examples. They should understand the connections between fractions, percents, decimals, and other mathematical topics such as algebra and functions. Students at this level are expected to have a thorough understanding of *Basic* level arithmetic operations—an understanding sufficient for problem solving in practical situations.

Quantity and spatial relationships in problem solving and reasoning should be familiar to them, and they should be able to convey underlying reasoning skills beyond the level of arithmetic. They should be able to compare and contrast mathematical ideas and generate their own examples. These students should make inferences from data and graphs, apply properties of informal geometry, and accurately use the tools of technology. Students at this level should understand the process of gathering and organizing data and be able to calculate, evaluate, and communicate results within the domain of statistics and probability.

***Advanced* (333)**

Eighth-grade students performing at the *Advanced* level should be able to reach beyond the recognition, identification, and application of mathematical rules in order to generalize and synthesize concepts and principles in the five NAEP content areas.

Eighth-graders performing at the *Advanced* level should be able to probe examples and counterexamples in order to shape generalizations from which they can develop models. Eighth-graders performing at the *Advanced* level should use number sense and geometric awareness to consider the reasonableness of an answer. They are expected to use abstract thinking to create unique problem-solving techniques and explain the reasoning processes underlying their conclusions

Source: (National Center For Education Statistics, 2010b).

LIST OF REFERENCES

Ad Hoc Committee on Confirming Test Results. (2002, March 1). *Using the National Assessment of Education Progress to confirm state test results: A report of the Ad Hoc Committee on Confirming Test Results*. Washington, DC: National Assessment Governing Board. Retrieved from http://nagb.org

Baker, E. L., Linn, R. L., Herman, J. L., & Koretz, D. (2002, Winter). From the directors: Standards for educational accountability systems. *The CRESST Line*, 1–4.

Bandeira de Mello, V., Blankenship, C., & McLaughlin, D. H. (2009). *Mapping state proficiency standards onto NAEP scales: 2005-2007 (NCES 2010-456).* Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.

Barnard, J., & Burner, D. (1975). *The American experience in education*. New York, NY: New Viewpoints.

Bourque, M.L. (2009). *A history of NAEP achievement levels: Issues, implementation, and impact 1989-2009.* Retrieved from http://www.nagb.org/publications/reports-papers.htm

Bracey, G. (2003). *What you should know about the war against America's public schools.* Boston, MA: Allyn and Bacon.

Bracey, G. (2007). The first time "everything changed": The 17th Bracey report on the condition of education. *Phi Delta Kappan*, *89*(2), 119-136.

Bracey, G. W. (2008). Cut Scores, NAEP Achievement Levels and Their Discontents. *School Administrator*, *65*(6), 20-23.

Bracey, G. (2009). *Education hell: Rhetoric vs. reality, transforming the fire consuming America's schools*. Alexandria, VA: Educational Research Service.

Broder, D.S. (2001, December 17). *Long road to reform: Negotiators forge education legislation*. Washington Post. p. A1.

Brown v. Topeka Board of Education, 347 U.S. 483 (1954).

Butts, R. F., & Cremin, L. A. (1953). *A history of education in American culture*. New York, NY: Holt.

Carnoy, M., & Loeb, S. (2002). Does external accountability affect student outcomes: A cross-state analysis. *Educational Evaluation and Policy Analysis*, *24*(4), 305-331.Decker, D. M., & Bolt, S. E. (2008). Challenges and opportunities for promoting student achievement through large-scale assessment results: Research, reflections, and future directions. *Assessment for Effective Intervention*, *34*(1), 43-51. Retrieved from EBSCO*host*.

Ercikan, K. (1997). Linking statewide tests to the National Assessment of Educational Progress: Accuracy of combining test results across states. *Applied Measurement in Education*, *10*(2), 145-159.

Goals 2000: Educate America Act, Pub. L No. 103-227 §. (1994). Retrieved from http://lamar.colostate.edu/~hillger/laws/goals-2000.html

Gordon, W. R. (2009). A comparison of eighth grade reading scores by state and by the four census-defined regions identified by NAEP. Orlando, FL: University of Central Florida. Retrieved from Dissertation & Theses: Full Text. (CFE0002536.)

Hambleton, R. K., Brennan, R. L., Brown, W., Dodd, B., Forsyth, R. A., Mehrens. . . Zwick, R. (2000). A response to "Setting reasonable and useful performance standards" in the National Academy of Sciences' Grading the Nation's Report Card. *Educational Measurement: Issues and Practice, 19*(2), 5–13.

Hansen, J. (1993). Is educational reform through mandated accountability an oxymoron? *Measurement and Evaluation in Counseling and Development, 26*(1), 11-21.

Itkonen, T. (2007). PL 94-142: Policy, evolution, and landscape shift. *Issues in Teacher Education*, *16*(2), 7-17.

Klein, S. P., Hamilton, L. S., McCaffrey, D. F., & Stecher, B. M. (2000). *What do test scores in Texas tell us?* Santa Monica, CA: Rand.

Kolen, M. J., & Brenan, R. L. (2004). *Test equating, scaling, and linking,* (2nd ed.). New York, NY: Springer.

Lane, S., Zumbo, B., Abedi, J. Benson, J., Dossey, J., Elliott, S., . . . Willhoft, J. (2009). Prologue: An introduction to the evaluation of NAEP, *Applied Measurement in Education*, *22*(4), 309-316. Retrieved from http://dx.doi.org/10.1080/08957340903221436

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. Retrieved from EBSCO*host*.

Lee, J., & McIntire, W. (2002). *Using national and state assessments to evaluate the performance of state education systems: Learning from the cases of Kentucky and Maine*. Orono, ME: University of Maine.

Linn, R. L. (2005). Conflicting demands of No Child Left Behind and state systems: Mixed messages about school performance. *Education Policy Analysis Archives, 13*(33). Retrieved from http://epaa.asu.edu/epaa/v13n33/

Linn, R. L. (2006). Educational accountability systems. CSE Technical Report 687. *National Center for Research on evaluation, standards, and student testing (CRESST)*. Retrieved from EBSCOhost.

Linn, R. L., Baker, E. L., & Betebenner, D. W. (2002). *Accountability systems: Implications of requirements of the No Child Left Behind Act of 2001*. CSE Technical Report. Retrieved from EBSCO*host*.

Linn, R. L., & Kiplinger, V. L. (1995). Linking statewide tests to the National Assessment of Educational Progress: Stability of results. *Applied Measurement in Education*, *8*(2), 135-155.

Lomax, R.G., & Hahs-Vaughn, D. L. (2012). *An introduction to statistical concepts* (3rd. ed.). New York, NY: Routledge/Taylor & Francis.

McLaughlin, D., Bandeira de Mello, V., Blankenship, C., Chaney, K., Esra, P., & Hikawa, H. (2008). *Comparison between NAEP and state mathematics assessment results: 2003. Vol. 1*. Research and Development Report. NCES 2008-475. National Center for Education Statistics, Retrieved from EBSCOhost.

McLaughlin, D., Bandeira de Mello, V., Blankenship, C., Chaney, K., Esra, P., & Hikawa, H. (2008). *Comparison between NAEP and state mathematics assessment results: 2003. Vol. 2*. Research and Development Report. NCES 2008-475. National Center for Education Statistics, Retrieved from EBSCOhost.

Murray K. T., & Murray B. A. (2007). *School law for the Florida educator* (3rd ed.). Merritt Island, FL: IntraCoastal.

National Center for Educational Statistics. (2003). *Federal programs for education and related activities*. Retrieved from http://nces.ed.gov/pubs2003/2003060d.pdf

National Center for Education Statistics. (2010a). *The NAEP mathematics achievement levels by grade.* Retrieved from http://nces.ed.gov/nationsreportcard/mathematics/achieveall.asp

National Center for Educational Statistics (2010b). *NAEP overview.* Retrieved from http://nces.ed.gov/nationsreportcard/about/

No Child Left Behind Act of 2001 [NCLB], Pub. L. No. 107-110, §. Retrieved from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Novak, J. R., & Fuller, B. (2003, December). Penalizing diverse schools? Similar test scores, but different students, bring federal sanctions. *Policy Brief*, 3-4, University of California Berkeley: Retrieved from http://pace.berkeley.edu

Obama, Barack. (2009). *Remarks by the President to the Hispanic Chamber of Commerce on a Complete and Competitive American Education.* Retrieved from http://www.whitehouse.gov/the_press_office/Remarks-of-the-President-to-the-United-States-Hispanic-Chamber-of-Commerce/

Olson, L. (2002a). A 'proficient' score depends on geography. *Education Week, 21*(23), 1-15.

Olson, L. (2002b). Want to confirm state test scores? It's complex, but NAEP can do it. *Education Week, 21*(26), 1-11.

Olson, L. (2005). Defying predictions: State trends prove mixed on schools making NCLB targets. *Education Week, 25*(2), 1.

107th Congress. (2002, January 8). Public Law – 107 -110. Retrieved from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf

Passow, A. H. (1960). Educating the gifted in the U.S.A. *International Review of Education, 6,* 141–153.

Peterson, P.E. & Lastra-Anadon, C.X.O (2010). State standards rise in reading fall in math. *Education Next,10(4), p.12-16.*

Peterson, P. & West, M. (2003). *No child left behind? The politics and practice of school accountability*. Washington, D.C.: Brookings Press.

Plessy v. Ferguson, 163 U.S. 537, 539 (1896).

Prowker, A., & Camilli, G. (2007). Looking beyond the overall scores of NAEP assessments: Applications of generalized linear mixed modeling for exploring value-added item difficulty effects. *Journal of Educational Measurement, 44*(1), 69-87.

Rickover, H. G. (1959). *Education and freedom*. New York, NY: Dutton.

Sadker, M. P., & Sadker, D. M. (2005). *Teachers, schools and society*. Boston: McGraw Hill.

Taylor, R., & Gordon, W. R. (2011). National Assessment of Educational Progress and state assessments: What do the data really mean? *ERS Spectrum, 29*(1), 29-34.

U. S. Department of Education. (1983). *Introduction: A nation at risk.* Retrieved from http://www2.ed.gov/pubs/NatAtRisk/intro.html

U. S. Department of Education. (2004a). *Charting the course: States decide major provisions under No Child Left Behind.* Retrieved from http://www2.ed.gov/news/pressreleases/2004/01/01142004.html#elements

U. S. Department of Education. (2004b). *Four pillars of NCLB.* Retrieved from http://www.ed.gov/nclb/overview/intro/4pillars.html

U. S. Department of Education. (2007). *Assessment and item specifications for the NAEP 2009 mathematics assessment: National Assessment of Educational Progress*. Washington, DC: National Assessment Governing Board. Retrieved from http://purl.fdlp.gov/GPO/LPS117844

U. S. Department of Education. (2009). *Education department: American Recovery and Reinvestment Act of 2009*. Retrieved from http://www2.ed.gov/policy/gen/leg/recovery/implement.html

U. S. Department of Education. (2010). *Adequate yearly progress.* Retrieved from https://answers.ed.gov/app/answers/detail/a_id/6/kw/adequate%20yearly%20progress

Waltman, K. (1997, Summer). Using performance standards to link statewide achievement results to NAEP. *Journal of Education Measurement*, *34*(2) 101-121.