2014

Multiple-Case Study and Exploratory Analysis of the Implementation of Value-Added Teacher Performance Assessment on Eighth Grade Student Achievement in Ohio, Pennsylvania, and Tennessee

James Carter
University of Central Florida

MULTIPLE-CASE STUDY AND EXPLORATORY ANALYSIS OF THE IMPLEMENTATION OF VALUE-ADDED TEACHER PERFORMANCE ASSESSMENT ON EIGHTH GRADE STUDENT ACHIEVEMENT IN OHIO, PENNSYLVANIA, AND TENNESSEE

by

JAMES DANIEL CARTER
B. A. University of North Florida, 2004
M. A. Jones International University, 2007

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

Fall Term
2014

Major Professor: Rosemarye Taylor

ABSTRACT

The goal of this research was to analyze the academic impact of the implementation of the Value Added Assessment Model. The researcher analyzed the Value Added Assessment Models in the three Value Added Assessment Model states that had implemented the Value Added Assessment Model for more than five years. Additionally, the research was done by analyzing the academic impact as measured by the eighth grade reading NAEP and the eighth grade mathematics NAEP. The researcher paired the three states that had implemented Value Added Assessment Model for more than five years, with three demographically matched states that had not implemented Value Added Assessment Model. The states were matched as follows: Ohio (Value Added Assessment Model implementing state) with Michigan (non Value Added Assessment Model state), Pennsylvania (Value Added Assessment Model implementing state) with Virginia (non Value Added Assessment Model state) and Tennessee (Value Added Assessment Model implementing state) with Georgia (non Value Added Assessment Model state). The mean composite scale score in NAEP from the following categories of students were compared and analyzed: 1) All students 2) White students 3) Black students 4) National School Lunch Program Eligible Students 5) National School Lunch Program Ineligible Students 6) Exceptional Education students.

The results of the study indicated that the impact of Value Added Assessment Model on academic impact as measured by the eighth grade reading NAEP and the eighth grade mathematics NAEP was negligible.

To my bride, Kate, who is the best earthly gift that God has given to me.

ACKNOWLEDGMENTS

As I sit here typing away on my keyboard I am reminded of all of the things that have changed since I first step foot in my first doctoral class. When I started my first class Kate and I had just one little beautiful curly haired girl, Payton (now 6). Now nearly five years later, we are blessed to have two more little girls Adison (4) and Avery (2) and a little boy, James, on his way in December.

I would like to thank Dr. Taylor for her patience as I persevered through this journey. Dr. Taylor has shown great understanding in allowing me to work on this paper over several years as I have experienced many life changes. In addition to Dr. Taylor, I would like to thank my entire committee (Dr. Doherty, Dr. Gordon and Dr. Hahs-Vaughn) as they have willingly served. Additionally, I would like to thank my new friend Dr. Tom Gollery for his sacrificial service in helping me get my dissertation fine-tuned.

Outside of the University of Central Florida I would like to thank all of the wonderful teachers that I have had the opportunity to learn from as a student and as a colleague. I want to offer a special thank you to my first writing teacher, my fourth grade teacher, Mrs. Wanda Carter or as I call her mom.

I would also like to thank my wife who has been incredibly supportive as I have spent many many hours working on this doctoral degree. She has been my number one fan and encourager. Finally, I would like to thank my father Dale and my brother David for always being proud of me and encouraging me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
THE PROBLEM AND ITS CLARIFYING COMPONENTS

Governor Ned McWherter, of Tennessee signed the Education Improvement Act in March 1992 marking the beginning of the Value Added Assessment Model (VAAM) movement (Bratton, Horn, & Wright, 1996). The Value Added Assessment Model "is a system that calculates the value teachers add to their students' achievement, based on changes in test scores from year to year and how the students perform compared with others in their grade" (Dillon, 2010, p. B 7). VAAM is a part of an ambitious education reform movement in the United States, initially developed by Sanders (1998) in the late 1980s and 1990s. Sander's background is in agriculture and he modified the process by which agricultural yields were evaluated for use in the evaluation of educators and educational institutions (Sanders & Horn, 1994). In 1996, Tennessee first used VAAM to provide teachers with VAAM scores to determine their effectiveness in increasing student achievement. The VAAM scores were based on predicted annual increases in student achievement and the actual increase in student achievement based on *The Comprehensive Test of Basic Skills (CTBS) 4th Edition.* The CTBS was used to indicate the difference between predicted academic achievement and actual academic achievement and reflects the teacher's impact on student academic achievement (Bratton, Horn, & Wright, 1996).

VAAM uses various factors to determine the value teachers, administrators, schools, school districts and state departments of education add to the education of the students. VAAM uses a statistical mixed-model to analyze the achievement of students and in turn determine the effectiveness of the classroom teacher. Using VAAM data, researchers have determined that the classroom teacher has the greatest impact on the success of students (Sanders, 1998).

As of the 2012-2013 school year, VAAM was being used statewide in Tennessee, Pennsylvania, and Ohio as well as in hundreds of individual school districts in 26 states

(Anderman, Anderman , Gimbert, & Yough, 2010). This movement created the need for research that would be helpful to school district and state level decision-makers as they determine to include or exclude VAAM from state's educator performance systems.

The current study will address the impact that VAAM implementation has had on student achievement in Ohio, Pennsylvania and Tennessee. The study compared eighth grade data for students in Ohio, Pennsylvania and Tennessee, who took the National Assessment in Educational Progress (NAEP) in the school-year 2012-2013 versus eighth grade data for students in three matched states; Michigan, Virginia and Georgia, who were not under the VAAM model and took the NAEP in SY 2012-2013.

Conceptual Framework

Studies and reports commissioned by government entities, beginning in the 1960s, have given rise to reform in public education (i.e. *Sputnik Crisis, Equality of Educational Opportunity, A Nation at Risk, No Child Left Behind, Race to the Top*). The studies called for more accountability for educators and educational institutions. One of the ways for accountability to increase would be to measure educator and educational institution performance based on student academic achievement through standardized testing (Chubb & Moe, 1990). Standardized testing would eventually provide the data necessary to measure educator and educational institution performance through VAAM.

The History of NAEP

To gain historical insights and understand some challenges of NAEP, it is necessary to examine the early forces that helped shape and direct NAEP. In 1963, the United States Commissioner of Education, Francis Keppel, first created a committee to look in the options for evaluating American education (Jones, 1996). Three years later in 1966 Ralph Tyler (the committee's chair) and Commissioner Keppel proposed the framework for NAEP (Jones, 1996). The idea that the federal government would assess state and local educational authorities (LEAs) was a source of contention regarding the role of the federal government in education. Many argued that the federal government was overstepping the bounds laid forth in the United States Constitution regarding federalism. Despite contention about the development of NAEP, it was first administered as a nationwide test in 1969 (Jones, 1996).

NAEP is now known as The Nation's Report Card, and has been assessing the state of education in the United States periodically since it was first given in 1969. NAEP is a standardized criterion referenced test, the analysis of NAEP performance provides stakeholders with the ability to know what testers have learned about a particular subject (Feuer, Holland, Green, Berenthal & Cadelle Hemphill, 1999). Assessments are given in mathematics, reading, science, history, geography, writing, and other fields (National Center for Education Statistics (NCES), 1999). NAEP is the only source for continuing statistically representative sample core subject area academic analyses for schools across the nation (NCES, 2001). NAEP allows for states and LEAs to be measured and compared to one another based on student academic achievement. The results of NAEP are made public through the National Center for Educational Statistics. The data that are published provide disaggregated academic performance data to policymakers and the public (U.S. Department of Education, 2002).

Prior to 1990, NAEP data were only provided for the nation as a whole and subgroups within the population. In 1988, the United States Congress passed legislation allowing for state participation in NAEP. These tests were to be administered, using separate representative samples within each state that agreed to participate. The trial state assessments were administered in 1990, 1992, and 1994. Starting with the 1996 assessment, NAEP administrations were no longer trial assessments. With the passage and signing into law Public Law 107-110, (No Child Left Behind Act of 2001 or NCLB), NAEP participation went from voluntary, to mandatory for states receiving Title I funding (National Center for Educational Statistics, 2010). Title I of the Elementary and Secondary Education Act (ESEA) provides financial assistance to local educational agencies (LEAs) with high numbers or high percentages of children from low-income families (United States Department of Education, 2012). The states are required to administer the NAEP every two years and test both reading and mathematics (National Center for Educational Statistics, 2010).

No Child Left Behind Act of 2001 (NCLB)

January 8, 2002, marked the signing of Public Law 107-110 in to law. Public Law 107-110, was developed by President George W. Bush's administration as No Child Left Behind Act of 2001. No Child Left Behind was largely agreed upon in Congress by parties on both sides of the aisle. The intent of President George W. Bush's signing into law the No Child Left Behind Act of 2001 (NCLB) was to "close the achievement gap with accountability, flexibility, and choice, so that no child is left behind" (p. 1). The passage of NCLB into law elevated educational reform through increased accountability in student achievement though standardized testing into a position of vanguard status in America (Shaul & Ganson, 2005).

NCLB required states to determine what annual progress each student subgroup would need to make in order to reach the proficiency goal. This progress became known as adequately yearly progress (AYP). States were allowed to come up with their own plans to meet AYP, but the federally mandated goal was for all students to meet academic proficiency in reading and mathematics by the end of the school year 2013-14.

NCLB awarded states a degree of autonomy, in the design as to how the goals would be met, but they were still guided by the federal government as to what those goals would accomplish. NCLB increased higher state education standards, which in turn offered students increased academic rigor. However, the only way for students to illustrate their academic proficiency was through standardized testing. According to NCLB Sec.1111 (b)(2)(A)(i), states were not allowed to use other measures of academic achievement in determining proficiency, because the other areas could have offset any deficiencies in academic proficiency (or lack thereof) in reading and mathematics (Linn, 2003).

Race to the Top (RTTT)

On February 17, 2009, President Barack Obama signed into law the American Recovery and Reinvestment Act of 2009 (AARA). The (AARA) provided $4.35 billion in funding for education in a grant program known as Race to the Top (RTTT). RTTT was designed to provide funding for states that fashioned conditions for academic excellence through innovation and reform. RTTT funds were allocated for states to implement agendas to academically prepare students to succeed in college, life, and the global workforce. RTTT also supported states financially that agreed to move towards a common national curriculum (United States Department of Education, 2009).

The adoption of RTTT further propelled Value Added Assessment Model (VAAM) in public education. As policymakers in education looked for ways to achieve academic excellence, much of the nation's attention turned to VAAM as a solution. The policymakers looked at VAAM as a way to identify what was working in education and more importantly to identify what was not working. As RTTT reached full implementation, VAAM would continue to be adopted by LEAs (United States Department of Education, 2009).

Problem Statement

At the time of this study there was no research to indicate the effectiveness of VAAM based on overall student performance as indicated by NAEP scores. This study focused on the relationship that VAAM has on the student achievement, in eighth grade students, in states that implemented VAAM as of the 2012-2013 school year. The researcher analyzed the difference of change in academic performance of the states with VAAM, compared to matched states that had not adopted VAAM for the 2012-2013 school year. A statistical analysis was performed to determine the relationship, if any, that VAAM had on student subgroups, such as White, Black, and economically disadvantaged (as indicated based on participation in the National School Lunch Program). The analysis was conducted based on the data that were synthesized, and observations that were made from researching state legislation pertaining to VAAM, and from information obtained from interviews conducted by the researcher with education experts in Ohio, Pennsylvania, and Tennessee.

Purpose of the Study

The purpose of this study was to collect and analyze data regarding the impact of VAAM on student achievement, in eighth grade students, in Ohio, Pennsylvania, and Tennessee whom had adopted VAAM matched against student achievement, in eighth grade students, in Michigan, Virginia and Georgia who had not adopted VAAM. The researcher (a) examined the relationship of VAAM implementation and student state level aggregated NAEP scores in the six states, (b) examined whether a relationship existed between VAAM implementation and student achievement of White students, (b) examined whether a relationship existed between VAAM implementation and student achievement of Black students, (c) examined whether a relationship existed between VAAM implementation and student achievement amongst economically disadvantaged students (as indicated based on participation in the National School Lunch Program), and (d) examined whether a relationship existed between VAAM implementation and student achievement amongst economically advantaged students (as indicated based on nonparticipation in the National School Lunch Program).

Research Questions

The questions that guided the research study were:

1. What are the Value Added Assessment Model designs in Ohio, Pennsylvania and Tennessee?
2. To what extent, if any, is there a difference in student achievement in eighth grade students, as measured by percent proficient in NAEP reading and mathematics scores among states with VAAM and matched states without VAAM?
3. To what extent, if any, is there a difference in student achievement in White eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
4. To what extent, if any, is there a difference in student achievement in Black eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
5. To what extent, if any, is there a difference in student achievement in National School Lunch Program eligible eighth grade students, as measured by the mean composite

scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

6. To what extent, if any, is there a difference in student achievement in National School Lunch Program ineligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
7. To what extent, if any, is there a difference in student achievement in Exceptional Education eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

Definition of Terms

Black students - Any student with origins in Africa, or Black racial groups of Caribbean Island nations (Burns, Wang, Henning, & National Center for Education Statistics, 2011).

National Assessment of Education Progress (NAEP) - a federally mandated assessment that is administered periodically to represent samples of students for the nation as a whole and for each state (Chudowsky, N., Chudowsky, M., & Center on Education Policy, 2010).

National Center for Education Statistics (NCES) - The primary federal organization for synthesizing data related to education in the United States and across the world. NCES is backed by a congressional mandate to synthesize the data in the field of education and use that data to complete statistics on the condition of education in the United States (Burns, Wang, Henning, & National Center for Education Statistics, 2011).

National School Lunch Eligible Students - "Any child at a participating school may purchase a meal through the National School Lunch Program. Children from families with incomes at or below 130% of the poverty level are eligible for free meals. Those with incomes between 130% and 185% of the poverty level are eligible for reduced-price meals" (p.2 USDA, 2012).

No Child Left Behind (NCLB) - A legislative Act from 2001. NCLB was designed to hold teachers, schools and administrators accountable to achieve the goal of all students being proficient in reading and mathematics by the 2013-14 school year (Taylor, Strecher, O'Day, Naftel, Le Floch & U.S. Department of Education, 2010).

Pay for Performance - A policy where educators are compensated based on the performance of students under their influence. The performance is measured in part by formal evaluations and/or student learning (Wells & Westat, 2011).

Race to the Top (RTTT) - The Obama administration's $4. 35 billion program, which was created for the purpose of sparking educational reforms (Manna & Ryan, 2013). The regulations for RTTT first appeared in the Federal Register in November 2009. The goals were identified as:

> (a) adopting internationally benchmarked standards and assessments that prepare students for success in college and the workplace; (b) building data systems that measure student success and inform teachers and principals in how they can improve their practices; (c) increasing teacher effectiveness and achieving equity in teacher distribution; and (d) turning around our lowest achieving schools. (U.S. Department of Education, 2009)

Socio-economic Status (SES) - A composite of five equally weighted, components: father's education, mother's education, family income, father's occupation, and household items. The term low SES refers to the lower quartile of the weighted SES composite (Burns, Wang, Henning, & National Center for Education Statistics, 2011).

Value Added Assessment Model (VAAM) - A statistical method for determining the impact of educators supporting academic progress for populations of students (Sanders, 1998).

Study Design

The following section illustrates the methodology used throughout the study. It comprises the population included in the study, sources of the data, data compilation methods, data analysis measures, organization of the study, instrumentation used for the study, and the significance of the study. The study is bicameral in nature; multiple-case study and quantitative descriptive statistical analysis.

Population

The population included eighth graders in Tennessee, Ohio, and Pennsylvania, who took the NAEP in SY 2012-2013 versus the test data of eighth graders in Michigan, Virginia and Georgia who took the NAEP in SY 2012-2013. The states were matched by analyzing geographic, demographic, and population statistics for the individual states. The states that had implemented any variation of VAAM were eliminated as potential matches for Ohio, Pennsylvania or Tennessee. The researcher matched Georgia with Tennessee, Michigan with Ohio, and Virginia with Pennsylvania.

Source of Data, Data Collection, and Analysis

For the quantitative portion of the study, the data were examined based on published state level aggregated student scores on the NAEP. Student NAEP 2013 data were collected during the spring of 2014. The data were collected from NAEP and NCES databases. The researcher obtained the data from the Internet, using http://nces.ed.gov/nationsreportcard/statecomparisons/ to obtain the information regarding NAEP scores. The variables that were examined are NAEP scores. Once the data were gathered from the NAEP and NCES databases the data were analyzed using SPSS software. The data from the NAEP scores were analyzed using descriptive statistics.

For the multiple-case study portion of the study, the researcher used a descriptive case study approach to analyze the states and compare legislation, policies, and procedures regarding K-12 education. The researcher conducted email interviews with state department of education designees in an attempt to get a clear understanding of the use of the Value Added Assessment Model design used in the three states studied. Email interviews were designed after determining evidence or documents that needed clarification by the researcher.

Limitations

The study was limited as follows:

1) Only analyzed data from three matched pairs of states.
2) The NAEP data used was aggregated state level data, not individual student level data.

Delimitations of the Study

1) The VAAM data were delimited to Tennessee, Ohio, and Pennsylvania.

2) The data are delimited to what is publicly published.

3) The comparison data are delimited to NAEP scores.

4) The matched states are Michigan, Virginia and Georgia.

Significance of the Study

Data were collected in this study to attempt to find a potential relationship between VAAM and student achievement. The findings would provide the decision makers in the educational community with case studies and analysis that could be suggestive to the impact of VAAM on student achievement.

Summary

Chapter 1 has delivered a summary of the research that was conducted by the researcher. Included were the conceptual framework, problem statement, purpose of the study, research questions, definition of terms, study design, and significance of the study. Within the conceptual framework the researcher included, a brief history of NAEP, a brief history of No Child Left Behind and a brief history of Race to the Top. Chapter 2 contains a review of pertinent literature related to Value-Added teacher performance and educator accountability. Chapter 3 contains an explanation and information describing the methodology used to conduct the research. Chapter 4 has description of the outcomes of the data analyses. Chapter 5 is a summary and description of the findings and conclusions obtained from the study.

CHAPTER 2
REVIEW OF THE LITERATURE

Introduction

This chapter is a review of relevant research pertaining to the impact of the Value Added Assessment Model (VAAM) approach to education. The review of literature begins with the history of the accountability of K-12 education and continues to the modern version of school accountability; with the inclusion of the Value Added Assessment Model. The researcher resolved to include the following review of literature through analysis of the most pertinent policies, processes and legislation pertaining to accountability in K-12 education.

History of Accountability in K-12 Education

In 1791, the authors of the United States Constitution's Bill of Rights included the 10^{th} Amendment. The 10^{th} Amendment of the Constitution states the following: "…the powers not delegated to the United States by the Constitution, nor prohibited by it to the states, are reserved to the states respectively, or to the people". Regarding education this Amendment applies Federalism, in favor of the states, to rule over the structure and function of their educational institutions within their respective boundaries. The role of the Federal government in education at the onset of our country's history was negligible. With the exception of the enforcement of the 14^{th} Amendment and how it impacts groups of individuals obtaining equal opportunities for education, the Federal government has taken a laissez-faire approach to handling education and left that responsibility to the individual states protecting their autonomy. Due to the Supremacy Clause, even though the Federal government attempted to play an insignificant role in education in our nation's early history, rulings and legislation made at the Federal level have greatly impacted education (Riley, 1977).

First in 1858, the United States Department of Education in the 1858 Newcastle report declared, “Our system of free schools was sustained directly by the people, without special care or direct aid from the government” (Riley, 1977, p. 3). During this time in our history, the expectations of schools by taxpayers were being met, and the problems that were identified were not expected to be fixed by the schools or the government. Instead the parents had the roles of being the teachers of their children in their own homes (Riley, 1977).

When this report was filed, our nation was quite different than it is today. At the time of the report, our nation was primarily agrarian and the family structure was the most effective instrument for augmenting behavior in society. As our nation embraced industrialization and urbanization, the role of the government in education began to morph. The morphology happened in response to our once homogenous society changing and our education system losing adeptness and efficiency. As a result, in the early 1900s, the United States Department of Education implemented reform that would result in a system which helped to insure that all citizens would have the rights to compulsory education (Riley, 1977).

Sputnik Crisis

In the late 1950s, the Soviet Union won the race to space with the launching of Sputnik Russian space satellite. With communism prevailing in a Cold War era, contest of intellect it became public and private opinion that our system of educating American youth was lacking. The scientists of United Socialist Soviet Republic had beaten the United States scientists to space. Consequently, America’s public education system became the focal point for a constant barrage of negative journalistic commentary regarding our ability to compete with other countries in the areas of mathematics and Science (Steeves, Bernhardt, Burns, & Lombard, 2009). The United States Office of Education’s life adjustment education humanistic curriculum was widely depicted by politicians and commentators as being too soft. A return to a more traditionalistic approach to education was heard by politicians, military leaders, and scientists. In reaction to the outcry for reform, Congress passed legislation that earmarked and dispensed considerable amounts of money for the advancement of public education. Following suit, President Dwight D. Eisenhower signed The National Defense Education Act into law, which increased the teaching of mathematics and science in public schools (Steeves, et al., 2009).

Sputnik set forth an unfortunate precedent by placing the blame of our societal shortfalls square on the shoulders of public education. Even worse, when the Soviet Union collapsed and the United States prevailed; public education received no credit for the victory (Bracey, 2007).

Elementary and Secondary Act of 1965

America's public schools were under attack again in the early 1960s as the disenchantment with schools continued in intensification. This would be the first time that the public's outrage would be rooted in the lack of results due to the influx of spending in the late 1950s. The public perception reflected the fact that the country had never truly recovered from the Sputnik Crisis. The results of the newly designed National Assessment of Educational Progress (NAEP) in the 1960s indicated that students still lacked competency in mathematics and science. Furthermore, the early 1960s brought with it social movements reflecting the injustices present in the apparent discrepancy concerning the qualities of education between different ethnic races. Furthermore, the Civil Rights Act of 1964 gave the United States Department of Education the authority to collect and analyze data pertaining to the racial makeup of schools (Johanningmeier, 2010).

> No person in the United States shall; on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be otherwise subjected to discrimination under any program or activity receiving Federal financial assistance

(dol.gov/oasam/regs/statutes/titlevi.htm, 1964).

In 1964, Congress passed legislation to enact the Civil Rights Act of 1964, which made it illegal to discriminate on the basis of race in any federally funded institution (Orfield, 1969). In response to the mounting pressure to close the gap on the racial inequalities of America's schools, the Elementary and Secondary Act of 1965 (ESEA) was enacted. At the time that it was passed ESEA was considered the most groundbreaking legislation in history to limit the autonomy of the state and local governments in regards to public education, by the federal government (Murphy, 1971). The federal government enticed the state and local governments to forego some of their autonomy and comply with ESEA policies by offering over four billion dollars in aid to the states. The aid came in the form of Title I funds. Title I funds were intended to assist with underprivileged students in the educational progression through the grade levels (Office of Education, 1970). In addition to the Title I funds, funds were also set up to help with various other educational and social programs (Office of Education, 1970).

As states accepted these federal dollars to assist in educating the underprivileged and disadvantaged youth, expectations were placed on the states that these groups of students would improve academically. The students' academic growth was monitored in reading and mathematics and was compared to other students throughout their state who were not identified as being disadvantaged (Kirst & Jung, 1980). The effectiveness of ESEA was to be appraised on a quadrennial basis to ensure that the intended outcomes were observed (United States Congress, 1965a).

Equality of Educational Opportunity 1966

The Equality of Educational Opportunity was a manifestation of section 402 of the Civil Rights Act of 1964 which stated (United States Congress, 1965b):

> The Commissioner shall conduct a survey and make a report to the President and the Congress, within two years of the enactment of this title, concerning the lack of availability of equal educational opportunities for individuals by reason of race, color, religion, or national origin in public educational institutions at all levels in the United States, its territories and possessions, and the District of Columbia (p.iii).

The federal government commissioned The Equality of Educational Opportunity study in order to determine if inequities existed in the education of America's youth. If these inequities were found, to what extent were they impacting the academic performance of the students which were marginalized due to the discrepancies? The Equality of Educational Opportunity study would be a monumental undertaking by the federal government. The sample size of the study was 4,000 schools and 645,000 students. The participants in the study represented a cross section of the nation at the time. The respondents included six identified racial/ethnic classifications; Negros, Puerto Ricans, American Indians, Mexican Americans, Oriental Americans, and Whites (Coleman, Campbell, Hobson, McPartland , Mood & Weinfeld, 1966).

Four major questions provided the framework for the study. These were:

1) To what extent if any are schools segregated on the basis of race/ethnicity?
2) Do all schools offer equal opportunities for academic rigor to all racial and ethnic groups?
3) To what extent are students achieving academically. As measured by standard achievement test?
4) To what extent, if any, are there a relationship between academic achievement and the schools in which students attend?

Upon completion of the study, The Equality of Educational Opportunity's findings were controversial. In essence, three conclusions were drawn from the study. First, the per pupil funding had an insignificant impact on student achievement. Second, the discrepancy in capital resources for White and Black schools was less than previously thought. Lastly, the backgrounds of the students that attended the schools had the largest impact on student achievement (eg. the peers), even more so than the quality of the academic instruction provided by the teacher. The ramifications of the findings of the Equality of Educational Opportunity study led to the expedited desegregation of schools. The lawmakers believed that if the quality of the peer influence was increased, then the academic achievement of the underprivileged students would significantly increase (Coleman, et al., 1966).

The Equality of Educational Opportunity study concluded that teachers and schools only account for 10% of the variance in academic achievement of students. This was controversial and contested at the time of the publication of the findings (Chubb & Moe, 1990). Educational researchers in the 21st century contended that schools and teachers account for closer to 30% of the variance in academic achievement (Marzano, 2003).

A Nation at Risk: The Imperative for Educational Reform

In 1969, public opinion began to be swayed by the Gallup organization's polling on the basis of public education. The results of the surveys further magnified the sentiment that American citizens wanted more for their public schools. In 1978, Gallup published that 41% of Americans responded that schools had declined (Tyack & Cuban, 1995).

Even after The Sputnik Crisis, ESEA, and The Equality of Educational Opportunity study, American sentiment was that schools were not educating students as well as they had educated the previous generations. In response, educational reform was once again at the forefront of political campaigns and legislative directives. By 1981, President Ronald Reagan directed his Secretary of Education, T. H. Bell, to create the National Commission on Excellence in Education (NCES). The NCES was instructed to analyze public schools, while paying specific attention to junior/high school students and colleges. The commission wanted to focus on the next generation of the workforce (National Commission on Excellence in Education, 1983).

The *Nation at Risk* report proclaimed, "The educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people" (p. 5). Upon publication, *A Nation at Risk*, identified that our nation's school system was in decline. The report identified that the educational system was no better than it had been prior to the Sputnik Crisis (National Commission on Excellence in Education, 1983). This was a major blow to the American psyche, considering the billions and billions of dollars poured into education during the preceding decades, with nothing to show for it. *A Nation at Risk*, further identified that the social programs that were in place in schools were detracting from the schools' ability to educate students. Schools resources were being stretched thin, for personal and political gains. The focus of schools should be solely to educate, and not to be a place for a multitude of other activities (National Commission on Excellence in Education, 1983).

At the time of the report, America was in a war of ideologies with the communists. President Reagan could allow for the perception that Americans were falling behind prevail. In response, to the *A Nation at Risk* findings, President Reagan called for reform in public education from kindergarten to post-secondary. President Reagan attributed our problems in education to the economy and to the enforcement of civil rights legislation. President Reagan argued that the legislation provided too many obstacles, which took the focus off of the role of schools. President Reagan's opinions were not accepted by all and were debated by many educators. President Regan's adversaries argued that the NAEP scores were not indicative of the decline that was reported in *A Nation at Risk* and that the findings of the study were being sensationalized (Mondale & Patton, 2001).

The NCES presented *A Nation at Risk* with several recommendations for the improvement of public schools. The suggestions for areas of improvement were set for both short term and long term improvement. The specific areas for improvement were comprehensive and included: covering curriculum, expectations, students' seat time, quality of instruction, leadership, and funding. The commission concluded that all students, college bound or not, deserved an adequate education (National Commission on Excellence in Education, 1983).

The first area recommended by the Commission for improvement was in the area of curriculum. The determination by NCES was that the course expectations were not high enough. In response, the commission instituted the 5 New Basics. The objective was to create a workforce that was second to none in the world. High school graduates were recommended to have the following as graduation requirements: six semesters of high school math, science, and social studies; eighth semesters of high school English; one semester of computer classes and four semesters of foreign language (for college bound students). For each subject, the commission had suggestions for what students should master as a result of passing the course (National Commission on Excellence in Education, 1983).

The second area recommended by the Commission for improvement was in the capacity of expectations and standards. The commission implied that grade inflation was causing infidelity in the education of students across America. The implication was made when the commission addressed the purpose of grade it was to provide students and other entities with evidence of the students' preparedness for future education in the subject matter. The commission further implied that the lack of integrity in the public school grading procedures by suggesting that standardized tests be administered and that the scores should be used as a basis for academic readiness. The commission suggested that formative assessments and summative assessments be given periodically to serve as in-process measures to check for students' academic progress and achievement levels (National Commission on Excellence in Education, 1983).

The third area recommended by the Commission for improvement was in the area of student seat time. The Commission found that American students were in school 35 days less per year than other industrialized nations, 185 days compared to 220 days. The commission further suggested that instructional time should focus on academics and not other superfluous subjects. The commission also found that class time was not being used effectively and that teachers needed training and support in the areas of classroom management and organization (National Commission on Excellence in Education, 1983).

The fourth area recommended by the Commission for improvement was in the area of teacher preparation and teacher incentive models. This area addressed the need for colleges and universities to better prepare teachers graduating from their colleges of education for the art of teaching. The Commission suggested that teacher preparation programs should be evaluated based on the readiness rate of their graduates. The incentives for educators in the field of education should be merit based not years of service based. This would allow for expert teachers to be rewarded for their expertise and novice teachers to be paid based on of their skills, not their longevity. The commission suggested that having an incentive based pay scale would serve two purposes. First, it would allow for recent college graduates to make more money in a shorter period of time. This would increase the quality of students that would choose education as a field of study. Second, it would help schools to retain the expert teachers and not have them leave the field. The final suggestion in the area of teacher quality was that the commission suggested extending the teacher calendar to an eleven month calendar to give adequate time for professional growth (National Commission on Excellence in Education, 1983).

The concluding area recommended by the Commission for improvement was in the area of leadership. The Commission suggested that administrators should not simply serve as managers, but rather should serve in the capacity of a leader. The administrators should serve to assist with the implementation of reforms in education. Having leaders work to change to meet the needs of the students as identified by research would make schools more efficient. The Commission also suggested that the Federal government should take a larger role in providing financing for schools to assist in making sure that schools were preparing students for the future. The specific areas where The Commission identified for the Federal government to have responsibility over was in the areas of: student rights, collection and analyzing of data, research related to curriculum, professional development of educators in areas of critical need and providing monetary assistance to impoverished students (National Commission on Excellence in Education, 1983).

Both *The Equality of Educational Opportunity* and *A Nation at Risk* were critical analyses of public education in the United States. These studies both served to impact and redirect the direction of education in America. *The Equality of Educational Opportunity* identified that schools and teachers only had a marginal impact on academic achievement in comparison to family and peer influence (Coleman, et al., 1966). *A Nation at Risk* perpetuated the fact that schools were functioning below expectations and were not making the most efficient use of the opportunities awarded to them and was identified as a liability to the security of our nation. These reports served as the impetus for the modern American education system (Johanningmeier, 2010). These results were presented a generation ago and still have not come to full fruition, and they are still in the national education debate. Both of these studies perpetuated the standardized testing and school accountability movement.

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

As the 21st century approached, the Clinton Administration implemented standards based curriculum that would be directed from the federal level, but based at the state and local level. The primary focus of the efforts of the Clinton administration was to focus on establishing national standards for education. These standards would serve as a guideline to provide students with rigorous academic programs and provide remediation for students that were failing to meet expectations. This would allow for educators to increase their expectations of students as a whole (Congress, 1994).

The Improving America's Schools Act of 1994 (IASA) is considered the reauthorization of ESEA. IASA of 1994 continued to maintain the influence of the federal government upon public education. The federal government was able to influence policy changes at the state and local levels by offering grants and other sources of monies for compliance. As state Department of Education's and Local Education Agencies (LEAs) came into compliance, they were offered funds with the purpose of aligning federal standards and curriculum. With the federal government offering substantial funding for compliance, LEAs were able to be held accountable to federal guidelines regarding student achievement (Congress, 1994).

IASA would ultimately end up paving the way for the standards-based education. In

Title I of IASA, it is required that states establish rigorous academic standards. The rigorous academic standards are created not only from an instructional perspective, but also by from a state standardized testing perspective. As federally guided standards based education gained momentum and political traction, standardized testing became a reality in school districts around the nation (Schwartz & Robinson, 2000).

The Goals 2000: Educate America Act (P.L. 103-227) was signed into law on March 31, 1994 by President Bill Clinton (Hise, 1994). The Act provides resources to states and communities to ensure that all students reach their full potential. The passage of this act further exemplified the federal government's aspiration to be involved with local level education. The goals that were set forth were goals that could, and should, be met as a nation, thus illustrating the fact that the federal government saw it as a duty to manage local level education through federal mandates. Goals 2000 started off in Title I of the legislation proclaiming that: by the year 2000, the nation's graduation rate would increase to 90%, all children would enter the school system ready to learn, all students in grades 4, 8, and 12 would demonstrate a level of academic mastery prior to proceeding to the next grade. Also, the United States will be number 1 in academic achievement in the world in mathematics and science. In addition, every school will be absent from drugs and violence, and every adult in America will be able to read and write. Parents will work with schools to assist in, and support the education process, and that all students will strive to be productive citizens. Lastly, every teacher will work to develop themselves professionally to embrace the skills of the 21st century school (Congress, 1994).

The Goals 2000 legislation had a polarizing effect on the country. Conservatives believed that it paved the way for the expansion of the federal government into local governments' territory. Liberals believed that the standardized testing would unfairly impact the already impoverished areas of the nation (Schwartz & Robinson, 2000).

Never the less, Goals 2000 helped to provide guidance for the development of standards based education and standardized testing. For decades, the federal government identified in study after study that American students were not achieving at a level in which the federal government deemed acceptable. As the federal government gained more and more influence over the LEAs, it began to impose its will more and more. The federal government set forth with a plan of making sure that students were aware of grade level expectations. The direction of National Education Policies of the 21st century currently follows the federal government's lead to promote higher standards and more academically rigorous state curriculum (Schwartz & Robinson, 2000).

LEAs were primarily influenced by grant funding, given from the federal level. Grants are awarded after a plan is submitted mapping the direction that LEAs will proceed. One LEA is then compared to another LEA to decide which one will receive the funding. As LEAs receive the funding they then were required to follow through with the direction of the grant proposal. After the inception of Goals 2000, over 1.4 billion dollars, in grant money was dispersed over a four year span. Over 90% of the grants that were awarded at this time required local level educational reform (Congress, 1994).

The federal governments believed that increased academic rigor would result in higher academic achievement. As the federal government pushed for increased academic standards to education, educators communicated reservations. Educators believed that a one size fit all business approach to education would not obtain the intended results (Congress, 1994).

As Goals 2000 was implemented, it became a concern as to the fidelity in which the standards were being implemented. The learning curve for teachers to understand the new standards and the willingness of the teachers to adjust their teaching practices to accommodate the new standards was unknown (Riley, 1996).

No Child Left Behind Act of 2001 (NCLB)

January 8, 2002, marked the signing of Public Law 107-110 into law. Public Law 107-110, was developed by President George W. Bush's administration and was known as No Child Left Behind. No Child Left Behind was largely agreed upon in Congress by parties on both sides of the aisle. The intent of President George W. Bush's signing into law the No Child Left Behind Act of 2001 (NCLB) was to "close the achievement gap with accountability, flexibility, and choice, so that no child is left behind" (p. 1). The No Child Left Behind Act is the reauthorization of The Elementary and Secondary Education Act of 1965 (ESEA).

The passage of NCLB into law elevated educational reform through increased accountability in student achievement though standardized testing into a position of vanguard status in America (Shaul & Ganson, 2005). As the 107th Congress and President George W. Bush reauthorized ESEA, they authorized some wide reaching expansions of ESEA. NCLB would work to:

1. Require annual standardized testing in reading and mathematics.
2. Increase the level of accountability in LEAs and state DOEs.
3. Provide opportunities for students to leave under performing schools.
4. Increase the level of teacher quality, mandating that all teachers would be highly qualified.
5. Adjust the funding structure of schools LEAs and state DOEs.

NCLB followed suit with the direction of the original intentions of President Lyndon B.

Johnson's Elementary and Secondary Education Act (ESEA), which was a part of his administration's War on Poverty. President Johnson's War on Poverty allocated large quantities of federal dollars to racial and minority groups, in an effort to increase the academic achievement of those groups of students. NCLB took the initiative set forth by the Johnson Administration and made it a mandate for all students to be academically proficient in reading and mathematics by the year 2014 (Shaul & Ganson, 2005).

NCLB required states to determine what annual progress each student group would need to make in order to reach the proficiency goal. This progress became known as Adequate Yearly Progress (AYP). States were allowed to develop their own plans to meet AYP, but the federally mandated goal was for all students to meet academic proficiency in reading and mathematics by school year 2013-14 (Linn, 2003).

NCLB awarded states a degree of autonomy, in the design as to how the goals would be met, but they were still guided by the federal government as to what those goals would accomplish. NCLB increased higher state education standards, which in turn offered students increased academic rigor. However, the only way for students to illustrate their academic proficiency was through standardized testing. According to NCLB Sec. 1111 (b)(2)(A)(i), states were not allowed to use other measures of academic achievement in determining proficiency, because the other areas could have offset any deficiencies in academic proficiency (or lack thereof) in reading and mathematics (Linn, 2003).

NCLB further aligned with President Johnson's War on Poverty by mandating that academic progress be tracked on the following student groups: economically disadvantaged, ethnic and racial minorities, limited English proficient students and, students with disabilities. The tracking of these students was accomplished through the creation of subgroups. The subgroups were created and state DOEs and local LEAs were required to track and report the academic progress of each of the student subgroups. In essence, NCLB admitted that specific subgroups had a history of underperforming academically in relation to White middle class students. NCLB disseminated the notion that through gathering and disaggregation of data on these student subgroups, that academics would increase amongst the students within those groups. LEAs and schools that failed to meet AYP on these student subgroups would be potentially subject to a litany of mandates, corrective actions, and increased monitoring. NCLB made it so that student subgroups could no longer be ignored within the majority, but rather that they would have to be monitored as a reflection of the LEA or school as a whole.

NCLB guidelines were laid out in the ten proceeding "titles" of NCLB:

Title I – Improving the Academic Achievement of the Disadvantaged

Title II – Preparing, Training, and Recruiting High Quality Teachers and Principals

Title III – Language Instruction for LEP and Immigrant Students

Title IV – 21st Century Schools

Title V – Promoting Informed Parental Choice

Title VI - Flexibility and Accountability

Title VII - Indian, Native Hawaiian, and Alaska Native Education

Title VIII - Impact Aid Program

Title IX - General Provisions

Title X -Repeals, Re-designations, and Amendments to Other Statutes (107th Congress, 2002)

For the purposes for this study, the researcher focused on what Title I and Title VI do. Title I affords for the focus on disadvantaged students and Title VI allows states the autonomy in determining how they will account and report academic achievement.

NCLB Title 1:
Improving the Academic Achievement of the Disadvantaged

Title I of NCLB focused on the disadvantaged students of America. Title I provides funding to the LEAs in order to provide for the academic needs of the disadvantaged students in their jurisdiction (Bejoian & Reid, 2005). The purpose of Title I was stated in Section 1001:

> Ensure that all children have a fair, equal, and significant opportunity to obtain a high quality education and reach, at minimum, proficiency on challenging state academic standards and state academic assessments (No Child Left Behind Act of 2001).

As an imperative part of Title I, LEAs were to use disaggregated data as a means to address the learning of the disadvantaged students. The disaggregated data would provide instructors with a plethora of data that would support data driven decision making. The adjustment of instruction-based data, would afford disadvantaged students with better educational opportunities, thus attempting to close the disparities in academic achievement between disadvantaged students and White middle class students (Shaul & Ganson, 2005).

Fundamental to Title I of NCLB was increased funding for schools with high percentages of disadvantaged students. These funds could go to increasing the teacher-to-student ratio, extended day programs, instructional development, technology, and additional programs to increase student achievement. The expectations of the billions of dollars in additional funding to LEAs, was that LEAs would help be more responsible for closing the achievement gap between the subgroups and the general student population (Azzam, 2004).

The acceptance of NCLB mandates by state DOEs and LEAs were not met by a large amount of resistance, due to the large sums of money that were tied to compliance. States were able to comply with NCLB's specification of meeting performance standards rather easily, since the federal government gave only minimal guidance as to setting these standards. States therefore, set their own performance indicators and the disparity between states' cuts scores made it so that one state's AYP could not be accurately compared to another state. The reliability of testing with fidelity from state to state was not maintained (Schafer, Liu & Wang, 2007). The absence of reliable state to state comparison data made it sensible to consider national assessments for reliable data.

The National Assessment of Educational Progress (NAEP)

As states received federal funds through NCLB, they agreed to have fourth and eighth grade students partake in NAEP testing. The NAEP is a test that students would be assessed on a semiannual basis and would assess reading and mathematics proficiency. The Department of Education communicated that NAEP testing would empower parents by giving them information as to the state of education in our union (United States Department of Education, 2004).

NAEP was produced in 1969 by the United States Department of Education (DOE). The DOE created NAEP testing as a way to study the academic performance of students (in mathematics, reading, science, and writing). NAEP tests are created and scored by The National Center for Educational Statistics (NCES). NCES is under the control of the DOE and is responsible for making the data received from NAEP testing be available to the public (Jones, 1996).

Race to the Top (RTTT)

On February 17, 2009 President Barrack Obama signed into law the American Recovery and Reinvestment Act of 2009. The American Recovery and Reinvestment Act, would provide

$4.35 billion in funding for education in a grant program known as Race to the Top (RTTT). RTTT was designed to provide funding for states that were fashioning conditions for academic excellence through innovation and reform. RTTT funds were allocated for states to implement agendas to academically prepare students to succeed in college, life, and the global workforce. RTTT also supported states financially that agreed to move towards a common national curriculum (United States Department of Education, 2009).

The United States Department of Education publicized the rulers for states to receive Race to the Top funding. One of the main requirements for states to receive the Race to the Top funding is that the states must authorize legislation which delineates an effective educator as one "whose students achieve acceptable rates (eg. at least one grade level in an academic year) of student growth." (United States Department of Education, 2009). The teacher evaluations must be based in significant part on student growth (Lomax & Kuenzi, 2012).

The adoption of Race to the Top further propelled Value Added Assessment Model (VAAM) in public education. As policymakers in education looked for ways to achieve academic excellence much of the nation's attention turned to VAAM as a solution. The policymakers looked at VAAM as a way to identify what was working in education and more importantly to identify what was not working. As RTTT reached full implementation VAAM will continued to be adopted into LEAs (United States Department of Education, 2009).

Common Core Initiative

The National Governors Association Center for Best Practices (NGA Center) in conjunction with, the Council of Chief State School Officers (CCSSO) facilitated the Common Core Standards Initiative. These groups worked alongside classroom educators, school based administrators, district administrators, and other experts in the field of education to provide guidance as to what was necessary for students to learn in order for them to be adequately prepared for college and career at the culmination of their high school experience. In an attempt to provide the students with adequate education, the consortium used individual case studies of successful classrooms, schools, and districts to provide evidence of what methods and content yielded the largest effect on assisting students reach college and career readiness. The consortium operated under the understanding that consistent standards and expectations for students were attainable and achievable for students regardless of where they live (CCSSI, 2012).

The Common Core State Standards (CCSS) were officially released on June 2, 2010. The entities that revised and released the CCSS were Achieve and The Common Core Standards Initiative. These entities used the following guidelines when developing the CCSS: use effective models for education that are currently in practice in the United States as well as abroad, and provide all stakeholders with a comprehensive understanding as to what students are accountable for learning at each grade level. The CCSS are intended to be "building on the strength of current state standards, the CCSS are designed to be focused, coherent, clear, and rigorous; internationally benchmarked; anchored in college and career readiness; and evidence and research based." (parcconline.org/implementation, 2012b).

As a result of RTTT and CCSS two primary consortia, Partnership for Assessment of Readiness for College and Careers (PARCC) and Smarter Balanced emerged that are responsible for academic assessments and accountability data for CCSS. The two consortia are funded by the United States Department of Education. "PARCC is a consortium of 23 states plus the U.S. Virgin Islands working together to develop a common set of K-12 assessments in English and math anchored in what it takes to be ready for college and careers" (parcconline.org/about-parcc, 2012a). Smarter Balanced is made up of 24 states "These states share a commitment to developing a next-generation assessment system aligned to the Common Core State Standards that provide educators with meaningful feedback and actionable data" (Smarter Balanced, 2012). Both PARCC and Smarter Balanced have received grants from the United States Department of Education, Race To the Top Assessment Grants ($186 million and $175 million grant respectively) (parcconline.org/about-parcc, 2012a).

Value Added Assessment Model (VAAM)

As standardized testing and national educational reform provided a wide array of data, new opportunities for analyzing the data for improved educational practice developed. With the enormous amount of data available from state standardized test and NEAP tests, a new method for analyzing the data developed. In VAAMs, standardized tests are the most common measure of student achievement as their use simplifies the statistical modeling process. Concerns have been raised with using standardized test scores as the primary index to measure teacher effectiveness. However, the availability of data allowed for VAAM to be developed as a way to analyze the academic success of students. In VAAM, standardized tests are the most common measure of student achievement as their use simplifies the statistical modeling process. Concerns have been raised with using standardized test scores as the primary index to measure teacher effectiveness (Konstantopoulos, 2014).

VAAM was a part of an ambitious educational reform movement in the United States. VAAM first came to the education community in 1992 when the State of Tennessee adopted the program. The program was developed in the late 1980s and early 1990s by William L. Sanders Ph.D. Sanders had a background in agriculture and modified the process by which agricultural yields were evaluated for use in the evaluation of educators and educational institutions. Sanders was working as a statistician at the University of Tennessee when he devolved the concept of VAAM (Hershberg, Simon, & Lea-Kruger, 2004).

Since 1992, VAAM has been used as an educational accountability instrument. VAAM first showed up in legislation in included as part of Tennessee's educational reform bill (Tennessee HB752) (1992). Tennessee HB752 outlined value-added as a "statistical system for educational outcome assessment" (p. 6). In 1996, Tennessee first used VAAM to provide teachers with their VAAM scores to determine their effectiveness as educators based on student achievement (Bratton, Horn, & Wright, 1996).

VAAM uses various factors to determine the value teachers, administrators, schools, districts, and states add to the education of the students that they influence. Though controversial, VAAM is considered by many scholars to be the best method for evaluating educational effectiveness of stakeholders (Kupermintz, 2003). VAAM uses a statistical mixed-model to analyze the achievement of students and in turn determine the effectiveness of the classroom teacher. Using VAAM data, researchers have determined that the classroom teacher has the greatest impact on the success of students (Sanders, 1998).

The VAAM statistical model makes an attempt to control for variances outside of the control of schools and educators and only measure what is under the control of the educators. VAAM attempts to measure the impact that an educator, educational institution or educational process has had on student achievement. Through the analysis of VAAM results, one can identify what institutions, teachers, or methods have had the most significant impact on student achievement. In completion of this quantitative analysis of a highly qualitative process one can identify what modifications to the education process facilitate the most improvement in academic achievement. In summation, VAAM System uses students' standardized test scores to determine if the academic environment that the students were in contributed to higher than or lower than anticipated test scores (McCaffrey, Lockwood, Koretz & Hamilton, 2003).

(Center for Greater Philadelphia, n.d.)

Figure 1. Map of States with VAAM (VAAM states are shaded dark)

Tennessee was the first state to adopt VAAM as means for analyzing the academic impact of educators on student performance. Between Tennessee's adoption of VAAM in 1992 and 2006, 17 other states have used VAAM to one degree or another. The states in Figure 1 shaded darker have adopted a variation of VAAM (Center for Greater Philadelphia). The only states that have implemented VAAM statewide for more than five years (as of SY 2012-2013) are: Ohio, Pennsylvania, and Tennessee (Center for Greater Philadelphia, n.d.).

Summary

As the United States progressed into the modern era in education, reform for public education was the intonation of office-bearers, the private sector, parentages, journalist, political commentators, and a litany of other groups. With all of these groups impacting the role of education in our society, our schools have become much more than institutions of learning. They have become institutions to address every area of societal deficiency. The fact that educational institutions today now focus on a wide range of areas, schools are always a political target for one group or another. Even so, our schools remain a necessary part and are irreplaceable to our republic (Kober & Rentner, 2011).

CHAPTER 3
METHODOLOGY

Introduction

The intent of this study is to provide information for policymakers and stakeholders in the field of education, by analyzing the impact of the implementation of the Value Added Assessment Model on student academic performance as indicated on the 2013 NAEP. The multiple-case study and exploratory analysis methodology were selected for this study because these methodologies best met the needs of the population analyzed. The researcher analyzed the Value Added Assessment Models in Ohio, Pennsylvania, and Tennessee through a multiple-case study analysis. Additionally, the researcher analyzed the NAEP scores of the six selected states utilizing quantitative descriptive statistics.

Case study analysis was one of the methodologies selected for this study because of the qualitative analysis necessary to obtain a depiction of the Value Added Assessment Models in the six selected states. As is quoted in Robert K. Yin's 2009 book titled, *Case Study Research: Design and Methods*, from Wilbur Schramm's 1971 working paper, commissioned by the Department of Education "The essence of a case study, the central tendency among all types of case study, is that it tries to illuminate a decision or set of decisions: why they were taken, how they were implemented, and with what result" (p. 17).

Utilizing quantitative descriptive statistics the researcher further analyzed the NAEP scores of the six selected states; the difference in NAEP scores was taken in account, amongst other factors in providing a description of teacher impact in the selected states. The three selected states that implemented the Value Added Assessment Model for more than five years as of SY 2012-2013 (Pennsylvania, Ohio, and Tennessee) were compared with selected matched states, which had not implemented Value Added Assessment Model as of SY 2012-2013 (Michigan, Virginia and Georgia) (United States Department of Education, 2014).

Each year the data are reported through the NCES as a part of the Nation's Report Card. The data reported are accessible by the general public on the internet. The data are reported without student names, so the students' anonymity are maintained. However, the data are still connected to the demographic data of the students (United States Department of Education, 2011).

Statement of the Problem

At the time of this study there was no research to indicate the effectiveness of VAAM based on overall student performance as indicated by NAEP scores. This study will help to determine the relationship that VAAM has on the student achievement, in eighth grade students, in states and school districts that implemented VAAM as of the 2012-2013 school year. This research will analyze the difference of change in academic performance of the states with VAAM, compared to matched states that had not adopted VAAM for the 2012-2013 school year. An analysis of the relationship, if any, that VAAM has had on student subgroups, such as economically disadvantaged, Black, and English Language Learners will be conducted based on the data synthesized and observations made.

Population

The intended population of this study was comprised of the states which had implemented VAAM for more than five years, statewide. As of January 2013, out of the 50 states in the United States only three states had implemented VAAM statewide for more than five years. The states which had implemented VAAM for more than five years were Pennsylvania, Ohio, and Tennessee.

The researcher analyzed the 2010 United States Census data for the 33 states that had not implemented VAAM as of 2012-2013 school year and matched them with Pennsylvania, Ohio, and Tennessee. The states were matched by finding states that were similar geographically (ie. a northeastern state would not be matched with a southern state). Demographically, a very diverse state would not be matched with a homogenous state (ie. Tennessee 16.7% Black would not be matched with Vermont 1% Black). Economically, based on average household income (ie. Tennessee, average household income $51,083, would not be matched with Maryland, average household income $83,137. The individual state comparisons can be found in Appendix A, B and C. The matched states are as follows: Ohio and Michigan, Pennsylvania and Virginia, and Tennessee and Georgia. The charts that follow illustrate the demographics of the matched states.

The students tested were the students who participated the state's aggregate on the 2013 NAEP of NAEP. The data that were analyzed were aggregated at the state level. The student performance on the NAEP was the focus of this study.

Table 1. Racial Demographic by Percent of Population

	One Race White alone	One race Black or African American Alone	One race American Indian, Alaska Native Alone	One race Asian Alone	One race Native Hawaiian and other Pacific Islander Alone	One Race Some other Race	Two or more races
Tennessee [a]	77.6	16.7	0.3	1.4	1.7	-	1.7
Georgia [a]	59.7	30.5	0.3	3.2	2.1	-	2.1
Pennsylvania [b]	81.9	10.8	0.2	2.7	-	2.4	1.9
Virginia [b]	68.6	19.4	0.4	5.5	0.1	3.2	2.9
Ohio [c]	82.7	12.2	0.2	1.7	-	1.1	2.1
Michigan [c]	78.9	14.2	0.6	2.4	-	1.5	2.3

(Census, 2010) *Notes. Bracketed letters indicate matched states.*

Table 2. Total Population and National School Lunch Participation

	Total Population 2010 (in thousands)	Total National School Lunch Eligible 2010 (in thousands)	Proportion of the total population National School Lunch Eligible 2010	National School Lunch participation percentage eighth grade 2013
United States	308,746	31,752	9.72%	-
Tennessee [a]	6,346	699	9.08%	53%
Georgia [a]	9,688	1,303	7.44%	55%
Pennsylvania [b]	12,702	1,159	10.96%	40%
Virginia [b]	8,001	757	10.57%	32%
Ohio [c]	11,537	1,136	10.15%	43%
Michigan [c]	9,884	920	10.74%	42%

(Census, 2010, USDA, 2013, NAEP, 2013) *Notes. Bracketed letters indicate matched states.*

Table 3. Median Family Income and Education Level by state as a Percentage of the Adult Population 2009

	Median Family Income 2010	High school graduate or more	Bachelor's degree or more	Advanced degree or more
United States	$61,627	85.3	27.9	10.3
Tennessee [a]	$51,083	83.1	23.0	7.9
Georgia [a]	$55,209	83.9	27.5	9.9
Pennsylvania [b]	$61,890	86.6	34.0	14.1
Virginia [b]	$72,476	87.9	26.4	10.2
Ohio [c]	$56,518	87.6	24.1	8.8
Michigan [c]	$56,101	87.9	24.6	9.4

(Census, 2010) *Notes. Bracketed letters indicate matched states.*

Research Questions

The study was guided by the following research questions:

1. What are the Value Added Assessment Model designs in Ohio, Pennsylvania and Tennessee?
2. To what extent, if any, is there a difference in student achievement in eighth grade students, as measured by percent proficient in NAEP reading and mathematics scores among states with VAAM and matched states without VAAM?
3. To what extent, if any, is there a difference in student achievement in White eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
4. To what extent, if any, is there a difference in student achievement in Black eighth grade students, as measured by the mean composite scale score in NAEP reading and

mathematics scores among states with VAAM and matched states that have not implemented VAAM?

5. To what extent, if any, is there a difference in student achievement in National School Lunch Program eligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
6. To what extent, if any, is there a difference in student achievement in National School Lunch Program ineligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?
7. To what extent, if any, is there a difference in student achievement in Exceptional Education eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

Instrumentation

The researcher analyzed the Value Added Assessment Models in Ohio, Pennsylvania and Tennessee through case studies. The researcher used both quantitative and qualitative analysis. The researcher analyzed the eighth grade NAEP data from SY 2012-2013 in Georgia, Michigan, Ohio, Pennsylvania, Tennessee, and Virginia. The researcher also, analyzed the Value Added Assessment Model legislation and polices in Ohio, Pennsylvania and Tennessee. The researcher conducted research through analysis of published state law and policies on the individual state department of education websites. The researcher further investigated the policies regarding the use of Value Added Assessment Model through email correspondence interviews with the commissioners of education (or their designees) in Georgia, Pennsylvania, and Tennessee.

Reliability

NAEP is considered to be of the standard in measuring educational achievement for American K-12 education (Jones & Olkin, 2004). NAEP is the largest data source for the quality of education in the United States. The NAEP is administered to over 3 million students on a semiannual basis. Since the NAEP assesses such a large number of students, NAEP has taken widespread measures to ensure reliability. NAEP requires extensive quality controls to insure the highest level of accuracy to score the over 3 million test that are graded annually (United States Department of Education, 2011).

The Education Testing Service (ETS) will be responsible for the instrumentation, examination, and reporting of the 2013 NAEP data. NAEP assessments give students a variety of questioning formats including multiple-choice and open ended questions (United States Department of Education, 2013). The multiple-choice portions permit the students to choose the best answer among the five options permitted. Because NAEP findings have an impact on the public's understanding of student academic achievement, precautions are taken to ensure the reliability of these findings. In its current legislation, as in previous legislative mandates, Congress has called for an ongoing evaluation of the assessment as a whole. In response to these legislative mandates, the National Center for Education Statistics (NCES) has established various panels of technical experts to study NAEP, and panels are formed periodically by NCES or external organizations, such as the National Academy of Sciences, to conduct evaluations. The Buros Center for Testing, in collaboration with the University of Massachusetts/Center for

Educational Assessment and the University of Georgia, recently conducted an external evaluation of NAEP (United States Department of Education, 2013).

NAEP

The researcher collected data from the 2013 NAEP on the website of the National Center for Education Statistic. The National Center for Education Statistics is a division of the United States Department of Education. The data that were collected were from the 2012-2013 school year. The report is a part of The Nation's Report Card.

Legislation and Policy Analysis

The researcher also, analyzed the Value Added Assessment Model, legislation and polices in Ohio, Pennsylvania, and Tennessee. The researcher conducted research through analysis of published state law and policies on the individual state department of education websites.

Interviews

The researcher elected to obtain information through utilizing the principles of basic interpretive study. Basic interpretive study utilizes data collected through a variety of methods including the use of email interviews, in order to acquire perspective of the experience of another entity, in the case of the study it would be the interviewee (Dillman, Smyth & Christian, 2009).

The email interviews were used as one component in an attempt to construct an understanding of the Value Added Assessment Models used in Pennsylvania, Ohio and Tennessee. The interviews were created with semi-structured questions that enabled the interviewer to obtain the information in a consistent manner. In having written questions it allowed the researcher to attempt to have an element of consistency from one interviewee to the next (Dillman, Smyth & Christian, 2009).

In creating the questions for the interviews, the researcher first referenced the 2009 book *Internet, Mail, and Mixed Mode Surveys*, by Dillman, Smyth and Christian (2009). Particular attention was given to chapters 4, 5, and 6, as a guide in an attempt to write questions that would evoke the appropriate response from the interviewee, meanwhile minimizing interviewer bias. Additionally, the researcher had the questions reviewed by University of Central Florida College of Education professors.

Instrument Reliability and Validity Background

The NAEP was created in 1969 by the United States federal government. NAEP was created as a result of a contract with the Education Commission on States. The purpose of NAEP is to assess achievement of students across multiple subject areas. After the creation of NAEP, policy makers would have data to influence their policy decision making (Resnick, 1980).

Value Added Assessment Model (VAAM) is a part of an ambitious education reform movement in the United States. VAAM first came to the education community in 1992 when the State of Tennessee adopted the program. The program was developed in the late 1980s and early 1990s by William L. Sanders. Sanders has a background in agriculture and modified the process by which agricultural yields were evaluated for use in the evaluation of educators and educational institutions. In 1996, Tennessee first used VAAM to provide teachers with their VAAM scores to determine their effectiveness as educators based on student achievement (Bratton, Horn, & Wright, 1996).

VAAM used various factors to determine the value teachers, administrators, schools, districts, and states add to the education of their students. Though controversial, VAAM was considered by many scholars to be the best method for evaluating educational effectiveness of stakeholders (Hershberg, Simon, & Lea-Kruger, 2004). VAAM uses a statistical mixed-model to analyze the achievement of students and in turn determine the effectiveness of the classroom teacher. Using VAAM data, researchers have determined that the classroom teacher has the greatest impact on the success of students (Sanders, 1998).

Qualitative Data Collection and Analysis

In answering Research Question 1 the researcher performed the qualitative data collection and analysis by following two primary steps: semi-structured email interviews, and the examination of state legislation and polices regarding the Value Added Assessment Models in, Pennsylvania, Ohio and Tennessee. The first step required the researcher to gain a perspective on the Value Added Assessment Models in the states studied. The researcher did this by identifying legislation and policies pertinent to educational accountability and then quoting and summarizing the information found. The research was recorded and sorted within a word document. The second step was semi-structured email interviews; during the interviews the researcher emailed state officers in charge of public relations and inquiry related to Value Added Assessment Model.

Quantitative Data Collection and Analysis

In March 2014, eighth grade reading and mathematics performance data were collected from the NCES 2013 State Snapshot Report. The following 2013 data were collected and disaggregated within an SPSS worksheet:(a) name of state, (b) state adoption status of Value Added Assessment Model, (c) the state average NAEP reading scale score of eighth grade students, (d) the state average NAEP reading scale score of eighth grade White students, (e) the state average NAEP reading scale score of eighth grade Black students, (f) the state average NAEP reading scale score of eighth grade National School Lunch Program eligible students, (g) the state average NAEP reading scale score of eighth grade National School Lunch Program ineligible students, (h) the state average NAEP math scale score of eighth grade students, (i) the state average NAEP math scale score of eighth grade White students, (j) the state average NAEP math scale score of eighth grade Black students, (k) the state average NAEP math scale score of eighth grade National School Lunch Program eligible students and (j) the state average NAEP math scale score of eighth grade National School Lunch Program ineligible students. This information was analyzed using simple descriptive statistics and a t-test of independent means, if the difference was statistically significant then a Cohen's d was run to determine the effect size, this was done to answer Research Questions 2, 3, 4, 5, 6 and 7.

Summary

Upon analyzing the data from NAEP, the interviews, legislation, and policy analysis, the researcher was able to compare states that implemented Value Added Assessment Model for more than five years and states that did not implement the Value Added Assessment Model. In answering Research Question 1 the researcher was able to identify the Value Added Assessment Model used in each of the states and the legislation and policies that drove the states' to adopt the Value Added Assessment Model. Furthermore, the researcher was able to analyze the data obtained from researching questions 2, 3, 4, 5, 6, and 7 by imputing the data in SPSS and running and analyzing the basic descriptive statistics and a t-test of independent means that were derived from the data. The data were then disaggregated and analyzed. The points for disaggregation answered Research Questions 2, 3, 4, 5, 6, and 7. The data were then disaggregated by states that had implemented the Value Added Assessment Model and states that had not implemented the Value Added Assessment Model.

CHAPTER 4
ANALYSIS OF THE DATA

Introduction

This study was chosen to analyze states that had implemented a Value Added Assessment Model (VAAM) for education accountability in K-12 education for more than three years. The researcher analyzed the quantitative data of 2013 National Assessment for Education Progress (NAEP) reading and mathematics scores by state disaggregated for eighth grade students, as published in the Nation's Report card by the United States Department of Education. The states that had implemented VAM for more than three years were Ohio, Pennsylvania, and Tennessee. The VAM implementing states were matched with three states that had not implemented VAM as of SY2012-2013 and were matched on demographic variables: Michigan with Ohio, Virginia with Pennsylvania, and Georgia with Tennessee. The states were matched by analyzing geographic, demographic, and population statistics for the individual states. The states that had implemented any variation of VAAM were eliminated as potential matches for Ohio, Pennsylvania or Tennessee. The researcher further analyzed the states by performing multiple case studies on each of the three VAM states. The researcher analyzed published data, state legislation enacted in each state, and conducted email interviews with state designees in Ohio, Pennsylvania and Tennessee.

Research Question 1) What are the Value Added Assessment Model designs in Ohio, Pennsylvania and Tennessee?

SAS EVAAS

SAS EVAAS in the contracted organization in that is contracted by Ohio, Pennsylvania and Tennessee. Ohio, Pennsylvania and Tennessee all use the same Value Added Assessment Model. The model used did not account for poverty as an individual factor to be considered. The model used was the SAS EVAAS. SAS EVAAS collected and analyzed the data from each state. There were two different models used by SAS EVAAS, both models were used in Ohio, Pennsylvania and Tennessee. The data sets drive which model is used at which time. The two analyses that were used were the EVAAMS Multivariate Response Model and the EVAAMS Univariate Response Model. These models were used by the three states whose results were analyzed (SAS Factsheet, 2014). *Details of the models can be found in Appendix L.*

The EVAAMS Multivariate Response Model is a linear mixed model. The EVAAMS Multivariate Response Model uses gains to determine a value-added rate. The EVAAMS Multivariate Response Model requires two conditions to give an appropriate evaluation of the learning gains. The first factor that is considered is that the data must be scaled. The data must have two comparable means, to determine if an appropriate level of gain is achieved. The second requirement for the EVAAMS Multivariate Response Model is that there must be a before and after point. This works best with end of year tests, which permits a clear before and after introduction of one year's worth of instruction. The EVAAMS Multivariate Response Model does not work as well with high school courses that are generally one year of a particular subject and not a continuation over multiple years. In these instances SAS EVAAS utilizes the EVAAMS Univariate Response Model (SAS Factsheet, 2014).

The EVAAMS Univariate Response Model is an analysis of covariance (ANCOVA) model. The EVAAMS Univariate Response Model is similar to the EVAAMS Multivariate Response Model, but does not project future scores, instead uses the previous year's data to predict current year scores. The EVAAMS Univariate Response Model works when students have multiple teachers and only have the subject for one semester or one year. Utilizing both the EVAAMS Multivariate Response Model and the EVAAMS Univariate Response Model insures that SAS EVAAS is able to properly analyze all students, teachers, schools and districts using the Value Added Assessment Model (SAS Factsheet, 2014).

Ohio

VAM was initiated in Ohio in 2005 with the passage of House Bill 107 by the 126th General Assembly of the state of Ohio's House of Representatives. The Bill was authored and sponsored by Representatives Setzer, Webster, Seitz, Kearns, Distel, C. Evans, Chandler, Combs, Domenick, D. Evans, Flowers and Hagan. The legislation reads "Within 180 days after the Department of Education implements the "value-added progress dimension," the curricula of the program, including methods of interpreting data, are aligned with that value-added progress dimension" (126th General Assembly H.B. 107). Furthermore, the bill states that the Ohio Department of Education is to develop (or have developed for them) Value Added Assessment Model and to begin implementing it not earlier than July 1, 2005, and not later than July 1, 2007. With this passage of legislation the state of Ohio adopted the Value Added Assessment Model as the method for measuring student growth over time (126th General Assembly H.B. 107).

Pennsylvania

Pennsylvania began implementing VAM in 2002 when the Pennsylvania League of Urban Schools and the Pennsylvania Department of Education (PDE) sponsored a program to provide value-added reports to school districts in the state. The program, now referred to as the Pennsylvania Value-Added Assessment System, started with a pilot project that included 32 districts in the spring of 2002 sponsored by the Pennsylvania League of Urban Schools. In September 2002, the State Board of Education approved plans to fund PVAAMS and developed a plan for introducing and implementing to the remaining school districts in the state. Full implementation of the Pennsylvania Value Added Assessment Model began in September 2006 (Hamilton & McCaffrey, 2007).

Tennessee

Tennessee Implemented VAM in 1992 with the passage of the Education Improvement Act which stated that the Tennessee Value Added Assessment Model System (TVAAMS) would take effect July 1, 1992. Tennessee Governor Ned Mc Wherter, signed the Education Improvement Act in March 1992 marking the beginning of the Value Added Assessment Model (VAAM) movement. In 1996, Tennessee first used VAAM to provide teachers with their VAAM scores to determine their effectiveness in increasing student achievement. The TVAAMS used an algorithm based on past individual student achievement to determine the appropriate rate of growth for each individual student for the current year. This method was used instead of a simple competence rate. Each individual student received his/her own level of student growth expected based on passed student performance. The VAAM scores were based on predicted annual increase in student achievement and the actual increase in student achievement based on *The Comprehensive Test of Basic Skills (CTBS) 4th Edition.* The CTBS was used to indicate the difference between predicted academic achievement and actual academic achievement and indicated the teacher effect on student academic achievement. The data were used to analyze the effectiveness of the current year's added value to their academic abilities (Bratton, Horn, & Wright, 1996).

Research Question 2) To what extent, if any, is there a difference in student achievement in eighth grade students, as measured by percent proficient in NAEP reading and mathematics scores among states with VAAM and matched states without VAAM?

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of all NAEP tested eighth grade students' performance comparisons in Value Added Assessment Model states versus non Value Added Assessment Model states. The matched states, Ohio and Michigan will be discussed first.

Ohio and Michigan NAEP All Eighth Grade Tested Students' Performance

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Ohio had 79% of Ohio's tested population at or above proficient. In comparison, Ohio's matched state Michigan had 77% of Michigan's tested population at or above proficient on the 2013 eighth grade reading NAEP. Since Ohio mandated VAAM statewide in 2007, Ohio went from 78% of the tested population proficient on the eighth grade reading NAEP in 2003 to 79% of the tested population proficient on the eighth grade reading NAEP in 2013. Over the same span in time the matched state, Michigan, went from 75% of the tested population proficient on the eighth grade reading NAEP in 2003 to 77% of the tested population proficient on the eighth grade reading NAEP in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were -3%. Changes in Michigan's NAEP reading scores were 0% for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 4.71%; $p<.002$. The mean difference between Ohio and Michigan from 2002-2013 was 5% (4.71) in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=2.33$.

Table 4. Percent proficient on the eighth grade reading NAEP all tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	82%	77%	5%
2003	78%	75%	3%
2005	78%	72%	6%
2007	79%	72%	7%
2009	80%	72%	8%
2011	79%	77%	2%
2013	79%	77%	2%
Change	-3%	0%	-3%

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Ohio had 79% of Ohio's tested population at or above proficient. In comparison, Ohio's matched state Michigan had 70% of Michigan's tested population at or above proficient on the

2013 eighth grade mathematics NAEP. Since Ohio mandated VAAM statewide in 2007, Ohio went from 76% of the tested population proficient on the eighth grade mathematics NAEP in 2003 to 79% of the tested population proficient on the eighth grade mathematics NAEP in 2013. Over the same span in time the matched state, Michigan, went from 75% of the tested population proficient on the eighth grade mathematics NAEP in 2003 to 77% of the tested population proficient on the eighth grade mathematics NAEP in 2013.

Changes in the NAEP mathematics scores for Ohio from 2002-2013 were 5%. Changes in Michigan's NAEP mathematics scores were 2% for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 7.5%; $p<.000$. The mean difference between Ohio and Michigan from 2002-2013 was 7.5% in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=4.06$.

Table 5. Percent proficient on the eighth grade mathematics NAEP all tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2003	74%	68%	6%
2005	74%	68%	6%
2007	76%	66%	10%
2009	76%	68%	8%
2011	79%	71%	8%
2013	79%	70%	9%
Change	5%	2%	3%

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Ohio had a mean NAEP score of 269 for Ohio's tested population. In comparison, Ohio's matched state, Michigan, had a mean NAEP reading score of 266 for Michigan's tested population. Since Ohio mandated VAAM statewide in 2007, Ohio's rank went from 16th out of 50 states on the

2003 NAEP (the test prior to VAAM implementation) to 18th out of 50 states in 2013. The matched state, Michigan's, rank went from 27th out of 50 states in 2007 to 32nd out of 50 states in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were 1. Changes in Michigan's NAEP reading scores were 1 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 4.71; $p<.001$. The mean difference between Ohio and Michigan from 2002-2013 was 4.71 in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=2.74$.

Table 6. Mean composite scale score on the eighth grade reading NAEP all tested students in Ohio and *Michigan 2002-2013*

Year	Ohio	Michigan	Difference
2002	268	265	3
2003	267	264	3
2005	267	261	6
2007	268	260	8
2009	269	262	7
2011	268	265	3
2013	269	266	3
Change	1	1	0

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Ohio had a mean NAEP score of 290 for Ohio's tested population. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 280 for Michigan's tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 17th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 10th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 34th out of 50 states in 2007 to 37th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Ohio from 2002-2013 were 8. Changes in Michigan's NAEP mathematics scores were 4 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 7.83; p<.000. The mean difference between Ohio and Michigan from 2002-2013 was 7.83 in favor of Ohio. That difference in turn was statistically significant with an effect size of d=3.13.

Table 7. Mean composite scale score on the eighth grade mathematics NAEP all tested students in Ohio and Michigan 2003-2013

Year	Ohio	Michigan	Difference
2003	282	276	6
2005	283	277	6
2007	285	277	8
2009	286	278	8
2011	289	280	9
2013	290	280	10
Change	8	4	3

Pennsylvania and Virginia NAEP all eighth grade tested students which were tested on the NAEP

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Pennsylvania had 81% of Pennsylvania's tested population at or above proficient. In comparison, Pennsylvania's matched state, Virginia, had 78% of Virginia's tested population at or above proficient on the 2013 eighth grade reading NAEP. Since Pennsylvania mandated VAAM statewide in 2006, Pennsylvania went from 76% of the tested population proficient on the eighth grade reading NAEP in 2003 to 81% of the tested population proficient on the eighth grade reading NAEP in 2013. Over the same span in time the matched state, Virginia, went from

79% of the tested population proficient on the eighth grade reading NAEP in 2003 to 78% of the tested population proficient on the eighth grade reading NAEP in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 4%. *Changes* in Virginia's NAEP reading *scores were down* 2% *for the same period. A t-test of independent means on the two groups,* Pennsylvania and Virginia *was run with a mean difference* = -.29%; p=.741. The mean difference between Pennsylvania and Virginia from 2002-*2013 was* nearly 0% (-.29%) in favor of Virginia. That difference *in turn was not statistically significant*.

Table 8. Percent proficient on the eighth grade reading NAEP all tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	77%	80%	-3%
2003	76%	79%	-3%
2005	77%	78%	-1%
2007	79%	79%	0%
2009	81%	78%	3%
2011	77%	78%	-1%
2013	81%	78%	3%
Change	4%	-2%	6%

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Pennsylvania had 78% of Pennsylvania's tested population at or above proficient. In comparison, Pennsylvania's matched state, Virginia, had 77% of Virginia's tested population at or above proficient on the 2013 eighth grade mathematics NAEP. Since Pennsylvania mandated VAAM statewide in 2006, Pennsylvania went from 69% of the tested population proficient on the eighth grade mathematics NAEP in 2003 to 78% of the tested population proficient on the eighth grade mathematics NAEP in 2013. Over the same span in time the matched state, Virginia, went from 72% of the tested population proficient on the eighth grade mathematics NAEP in 2003 to 77% of the tested population proficient on the eighth grade mathematics NAEP in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 9%. *Changes* in *Virginia's* NAEP mathematics scores were 5% *for the same period. A t-test of independent means on the two groups*, Pennsylvania and *Virginia was run with a mean difference* = -1.17%; p=.516. The mean difference between the Pennsylvania and *Virginia* from 200*3-2013 was* approximately -1% (-1.17%) in favor of Virginia. That difference *in turn was not statistically significant*.

Table 9. Percent proficient on the eighth grade mathematics NAEP all tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2003	69%	72%	-3%
2005	72%	75%	3%
2007	77%	77%	0%
2009	78%	76%	2%
2011	74%	78%	-4%
2013	78%	77%	1%
Change	9%	5%	4%
mean Difference			-1.17%

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Pennsylvania had a mean NAEP score of 272 for Pennsylvania's tested population. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 268 for Virginia's tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 29th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 7th out of 50 states in 2013 with Pennsylvania's matched state, Virginia, went from 8th out of 50 states in 2003 to 33rd out of 50 states in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 7. Changes in Virginia's NAEP reading scores were down 1 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -.286; p=.810. The mean difference between the Pennsylvania and Virginia from 2002-2013 was -.286 in favor of Virginia. That difference in turn was not statistically significant.

Table 10. Mean composite scale score on the eighth grade reading NAEP all tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	265	269	-4
2003	264	268	-4
2005	267	268	-2
2007	268	267	1
2009	271	266	5
2011	268	267	-1
2013	272	268	4
Change	7	-1	8

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 290 for Pennsylvania's tested population. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 288 for Virginia's tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 28th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 7nd out of 50 states in 2013. Pennsylvania's matched state, Virginia, went from 16th out of 50 states in 2003 to 15th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 11. Changes in Virginia's NAEP mathematics scores were 6 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -1.17; p=.580. The mean difference between the Pennsylvania and Virginia from 2003-2013 was -1.17 in favor of Virginia. That difference in turn was not statistically significant.

Table 11. Mean composite scale score on the eighth grade mathematics NAEP all tested students in Pennsylvania and Virginia 2003-2013

Year	Pennsylvania	Virginia	Difference
2003	279	282	-3
2005	281	284	-3
2007	286	288	-2
2009	288	286	2
2011	286	289	-3
2013	290	288	2
Change	11	6	5

Tennessee and Georgia NAEP all eighth grade tested students which were tested on the NAEP

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Tennessee had 77% of Tennessee's tested population at or above proficient. In comparison, Tennessee's matched state, Georgia, had 75% of Georgia's tested population at or above proficient. Since Tennessee mandated VAAM statewide in 1992, Tennessee's went from 71% of the tested population proficient on the eighth grade reading NAEP in 1998 to 77% of the tested population proficient on the eighth grade reading NAEP in 2013. Over the same span in time the matched state, Georgia, went from 68% of the tested population proficient on the eighth grade reading NAEP in 1998 to 75% of the tested population proficient on the eighth grade reading NAEP in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 6%. Changes in Georgia's NAEP reading scores were 7% for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = 1%; p=.462. The mean difference between Tennessee and Georgia from 1998-2013 was 1% in favor of Tennessee. That difference in turn was not statistically significant.

Table 12. Percent proficient on the eighth grade reading NAEP all tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	71%	68%	2%
2002	71%	70%	1%
2003	69%	69%	0%
2005	71%	67%	4%
2007	71%	70%	0%
2009	73%	72%	1%
2011	70%	74%	-4%
2013	77%	75%	2%
Change	6%	7%	0%

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Tennessee had 69% of Tennessee's tested population at or above proficient. In comparison, Tennessee's matched state, Georgia, had 68% of Georgia's tested population at or above proficient on the 2013 eighth grade mathematics NAEP. Since Tennessee mandated VAAM statewide in 1992, Tennessee's went from 47% of the tested population proficient on the eighth grade mathematics NAEP in 1992 to 69% of the tested population proficient on the eighth grade mathematics NAEP in 2013. Over the same span in time the matched state, Georgia, went from

48% of the tested population proficient on the eighth grade mathematics NAEP in 1992 to 68% of the tested population proficient on the eighth grade mathematics NAEP in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 22%. Changes in Georgia's NAEP mathematics scores were 20% for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = .778%; $p<.826$. The mean difference between Tennessee and Georgia from 1998-2013 was about 1% (.778%) in favor of Tennessee. That difference in turn was not statistically significant.

Table 13. Percent proficient on the eighth grade mathematics NAEP all tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1992	47%	48%	-1%
1996	53%	51%	2%
2000	52%	54%	-2%
2003	59%	59%	0%
2005	61%	62%	-1%
2007	64%	64%	0%
2009	65%	67%	-2%
2011	64%	68%	-4%
2013	69%	68%	1%
Change	22%	20%	0%

The state aggregate on the 2013 NAEP of eighth grade reading tested students in Tennessee had a mean NAEP score of 265 for Tennessee's tested population. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 265 for Georgia's tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 24th out of 37 states on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 34th out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 26th out of 37 states in 1998 to 19th out of 50 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 7. Changes in Georgia's NAEP reading scores were 8 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = 4.71; p<.002. The mean difference between Tennessee and Georgia from 1998-2013 was .374 in favor of Tennessee. That difference in turn was not statistically significant.

Table 14. Mean composite scale score on the eighth grade reading NAEP all tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	258	257	1
2002	260	258	2
2003	258	258	0
2005	259	257	2
2007	259	259	0
2009	261	260	1
2011	259	262	-3
2013	265	265	0
Change	7	8	1

The state aggregate on the 2013 NAEP of eighth grade mathematics tested students in Tennessee had a mean NAEP score of 278 for Tennessee's tested population. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 279 for Georgia's tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 32nd out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the1996 NAEP (the first state aggregated published state ranking NAEP results available) to 43rd out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 34th out of 40 states in 1996 to 40th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 19. Changes in Georgia's NAEP mathematics scores were 20 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -1.6; p=.646. The mean difference between Tennessee and Georgia from 1998-2013 was -1.6 in favor of Georgia. That difference in turn was not statistically significant.

Table 15. *Mean composite scale score on the eighth grade mathematics NAEP all tested students in Tennessee and Georgia 1992-2013*

Year	Tennessee	Georgia	Difference
1992	259	259	0
1996	263	262	1
2000	262	265	-3
2003	268	270	-2
2005	271	272	-1
2007	274	275	-1
2009	275	278	-3
2011	274	278	-4
2013	278	279	-1
Change	19	20	-1

Combined Comparison of VAAM and non VAAM States for Research Question 2

On the reading NAEP, the state aggregate of all eighth grade tested students had a mean difference of 2.0 in favor of the Value Added Assessment Model states of Ohio, Pennsylvania, and Tennessee. However, this mean difference was not statistically significant (t=1.47; p=.150). Conversely, on the mathematics NAEP of all eighth grade tested students, the mean difference was 6.11 again favoring the Value Added Assessment Model states, this was statistically significant with an effect size of 1.34 ($t=2.11$; $p<.042$, $d=1.34$).

Research Question 3) To what extent, if any, is there a difference in student achievement in White eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of NAEP tested White eighth grade students' performance comparisons in Value Added Assessment Model states versus non Value Added Assessment Model states. The matched states, Ohio and Michigan will be discussed first.

Ohio and Michigan White eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of White eighth grade reading tested students in Ohio had a mean NAEP score of 273 for Ohio's White tested population. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 271 for Michigan's White tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 20th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 29th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 12th out of 50 states in 2003 to 41st out of 50 states in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were 0. Changes in Michigan's NAEP reading scores were 1 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 3.57; p<.001. The mean difference between Ohio and Michigan from 2002-2013 was 3.57 in favor of Ohio. That difference in turn was statistically significant with an effect size of d=2.42.

Table 16. Mean composite scale score on the eighth grade reading NAEP White tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	273	270	3
2003	271	272	-1
2005	272	268	4
2007	274	267	7
2009	273	268	5
2011	274	269	5
2013	273	271	2
Change	0	1	-1
Mean Difference			3.57

The state aggregate on the 2013 NAEP of White eighth grade mathematics tested students in Ohio had a mean NAEP score of 294 for Ohio's White tested population. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 285 for Michigan's White tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 21st out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 18th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 24th out of 50 states in 2003 to 41st out of 50 states in 2013.

Changes in the NAEP mathematics scores for Ohio from 2003-2013 were 7. Changes in Michigan's NAEP mathematics scores were 1 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 5.33; $p<.006$. The mean difference between Ohio and Michigan from 2003-2013 was 5.33 in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=2.45$.

Table 17. Mean composite scale score on the eighth grade mathematics NAEP White tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2003	287	286	1
2005	289	285	4
2007	291	285	6
2009	291	286	5
2011	295	286	9
2013	294	287	7
Change	7	1	6

Pennsylvania and Virginia all White eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of White eighth grade reading tested students in Pennsylvania had a mean NAEP score of 279. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 275 for Virginia's White tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 31st out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 8th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 31st out of 50 states in 2003 to 15th out of 50 states in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 8. Changes in Virginia's NAEP reading scores were 0 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -.571; p=.699. The mean difference between Pennsylvania and Virginia from 2002-2013 was -.571 in favor of Virginia. That difference in turn was not statistically significant.

Table 18. Mean composite scale score on the eighth grade reading NAEP White tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	271	275	-4
2003	268	275	-7
2005	273	275	-2
2007	272	273	-1
2009	276	272	4
2011	275	273	3
2013	279	275	4
Change	8	0	8

The state aggregate on the 2013 NAEP of White eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 297. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 296 for Virginia's White tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 29th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 6th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 9th out of 50 states in 2003 to 23rd out of 50 states in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 12. Changes in Virginia's NAEP mathematics scores were 6 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -2.67; p=.246. The mean difference between Pennsylvania and Virginia from 2003-2013 was 2.67 in favor of Virginia. That difference in turn was not statistically significant.

Table 19. Mean composite scale score on the eighth grade mathematics NAEP White tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2003	285	290	-5
2005	287	293	-6
2007	293	296	-3
2009	294	294	0
2011	294	297	-3
2013	297	296	1
Change	12	6	6

Tennessee and Georgia NAEP all White eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of White eighth grade reading tested students in Tennessee had a mean NAEP score of 270. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 274 for Georgia's White tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 23rd out of 36 states on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 43rd out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 18th out of 36 states in 1998 to 19th out of 50 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 6. Changes in Georgia's NAEP reading scores were 6 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -3.63; p<.005. The mean difference between Tennessee and Georgia from 1998-2013 was -3.63 in favor of Georgia. That difference in turn was statistically significant with an effect of d=2.69.

Table 20. Mean composite scale score on the eighth grade reading NAEP White tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	264	268	-3
2002	265	268	-2
2003	265	268	-4
2005	265	268	-3
2007	267	271	-4
2009	267	268	-1
2011	265	272	-7
2013	270	274	-4
Change	6	6	0

The state aggregate on the 2013 NAEP of White eighth grade mathematics tested students in Tennessee had a mean NAEP score of 284. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 292 for Georgia's White tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 35th out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the1996 NAEP (the first state aggregated published state ranking NAEP results available) to 5th out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 29th out of 40 states in 1996 to 28th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 18. Changes in Georgia's NAEP mathematics scores were 22 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -7.71; p<.047. The mean difference between Tennessee and Georgia from 1998-2013 was -7.71 in favor of Georgia. That difference in turn was statistically significant with an effect size of d=1.02.

Table 21. Mean composite scale score on the eighth grade mathematics NAEP White tested students in Tennessee and Georgia 1992-2013

Year	Tennessee	Georgia	Difference
1992	266	270	-4
1996	270	276	-6
2000	269	279	-10
2003	277	284	-7
2005	278	284	-6
2007	282	288	-4
2009	282	289	-7
2011	281	291	-10
2013	284	292	-8
Change	18	22	-4

Combined Comparison of VAAM and non VAAM States for Research Question 3

Overall, NAEP tested White eighth grade students in both Value Added Assessment Model adopting states and non Value Added Assessment Model adopting states performed at the same level on reading NAEP (mean Difference=0.00). Interestingly, students enrolled in non Value Added Assessment Model states outperformed, on average, their peers from Value Added Assessment Model states by 7.11 in mathematics. The difference, however, was not statistically significant (t=1.26; p=.216).

Research Question 4) To what extent, if any, is there a difference in student achievement in Black eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of Black student performance in Value Added Assessment Model states versus Non VAM Model States.

Ohio and Michigan Black eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of Black eighth grade reading tested students in Ohio had a mean NAEP score of 247. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 246 for Michigan's Black tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 6^{th} out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the 2003 NAEP (the test prior to VAAM implementation) to 27^{th} out of 43 states in 2013 with Ohio's matched state, Michigan, going from 29^{th} out of 40 states in 2003 to 33^{rd} out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were 1. Changes in Michigan's NAEP reading scores were 4 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 5.7; $p<.005$. The mean difference between Ohio and Michigan from 2002-2013 was 5.7 in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=2.01$.

Table 22. Mean composite scale score on the eighth grade reading NAEP black tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	246	242	4
2003	249	242	7
2005	243	239	4
2007	246	236	10
2009	247	238	9
2011	247	244	3
2013	247	246	1
Change	1	4	-3

The state aggregate on the 2013 NAEP of Black eighth grade mathematics tested students in Ohio had a mean NAEP score of 267. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 251 for Michigan's Black tested population. Since Ohio mandated VAAM statewide in 2007, Ohio went from 10th out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the2003 NAEP (the test prior to VAAM implementation) to 11th out of 43 states in 2013 with Ohio's matched state, Michigan, going from 35th out of 40 states in 2003 to 41st out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013.

Changes in the NAEP mathematics scores for Ohio from 2003-2013 were 10. Changes in Michigan's NAEP mathematics scores were 6 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 12.8; p<.000. The mean difference between Ohio and Michigan from 2003-2013 was 12.8 in favor of Ohio. That difference in turn was statistically significant with an effect size of d=3.49.

Table 23. Mean composite scale score on the eighth grade mathematics NAEP Black tested students in Ohio and Michigan 2003-2013

Year	Ohio	Michigan	Difference
2003	257	245	12
2005	255	247	8
2007	258	244	14
2009	260	246	14
2011	263	250	13
2013	267	251	16
Change	10	6	4

Pennsylvania and Virginia all Black eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of Black eighth grade reading tested students in Pennsylvania had a mean NAEP score of 250. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 249 for Virginia's Black tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 25th out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the2003 NAEP (the test prior to VAAM implementation) to 13th out of 43 states in 2013 Pennsylvania's matched state, Virginia, went from 4th out of 40 states in 2003 to 40th out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 14. Changes in Virginia's NAEP reading scores were down 3 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -6.57; $p<.016$. The mean difference between Pennsylvania and Virginia from 2002-

2013 was -6.57 in favor of Virginia. That difference in turn was statistically significant with an effect size of d=1.73.

Table 24. Mean composite scale score on the eighth grade reading NAEP Black tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	236	252	-16
2003	243	250	-7
2005	239	251	-12
2007	248	252	-4
2009	249	250	-1
2011	244	251	-7
2013	250	249	1
Change	14	-3	17

The state aggregate on the 2013 NAEP of Black eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 262. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 267 for Virginia's Black tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 31st out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the 2003 NAEP (the test prior to VAAM implementation) to 21st out of 43 states in 2013 Pennsylvania's matched state, Virginia, went from 4th out of 40 states in 2003 to 10th out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 15. Changes in Virginia's NAEP mathematics scores were 5 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -10.50; p<.003. The mean difference between Pennsylvania and Virginia from

2003-2013 was -10.5 in favor of Virginia. That difference in turn was statistically significant with an effect size of d=2.31.

Table 25. Mean composite scale score on the eighth grade mathematics NAEP Black tested students in Pennsylvania and Virginia 2003-2013

Year	Pennsylvania	Virginia	Difference
2003	247	262	-15
2005	250	263	-13
2007	257	268	-11
2009	260	268	-8
2011	257	268	-11
2013	262	267	-5
Change	15	5	10

Tennessee and Georgia NAEP all Black eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of Black eighth grade reading tested students in Tennessee had a mean NAEP score of 251. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 252 for Georgia's Black tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 25th out of 30 states (only 30 states reported disaggregated scores for the subgroup) on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 18th out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013 with Tennessee's matched state, Georgia, going from 18th out of 30 states in 1998 to 14th out of 43 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 16. Changes in Georgia's NAEP reading scores were 11 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -5.25; p<.032. The mean difference between Tennessee and Georgia from 1998-2013 was -5.25 in favor of Georgia. That difference in turn was statistically significant with an effect size of d=1.20.

Table 26. Mean composite scale score on the eighth grade reading NAEP Black tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	235	241	-6
2002	240	246	-6
2003	239	244	-5
2005	240	241	-1
2007	240	246	-6
2009	243	249	-6
2011	240	251	-11
2013	251	252	-1
Change	16	11	5

The state aggregate on the 2013 NAEP of Black eighth grade mathematics tested students in Tennessee had a mean NAEP score of 257. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 262 for Georgia's Black tested population. Since Tennessee mandated VAAM statewide in 1992, Tennessee went from 28th out of 30 states on the 1996 NAEP (the first state aggregated published state ranking NAEP results available) to 34th out of 43 states in 2013 with Tennessee's matched state, Georgia, going from 23rd out of 30 states in 1996 to 19th out of 43 states (only 43 states reported disaggregated scores for the subgroup) in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 22. Changes in Georgia's NAEP mathematics scores were 21 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference =

-7.67; p=.104. The mean difference between Tennessee and Georgia from 1998-2013 was -7.67 in favor of Georgia. That difference in turn was not statistically significant.

Table 27. Mean composite scale score on the eighth grade mathematics NAEP Black tested students in Tennessee and Georgia 1992-2013

Year	Tennessee	Georgia	Difference
1992	234	241	-7
1996	234	240	-6
2000	235	244	-9
2003	242	250	-8
2005	246	255	-9
2007	254	261	-7
2009	254	262	-8
2011	252	262	-10
2013	257	262	-5
Change	22	21	2

Combined Comparison of VAAM and non VAAM States for Research Question 4

On the eighth grade reading NAEP, Black students, on average, performed at a rate 1.44 points greater in non Value Added Assessment Model states than their counterparts enrolled in states adopting a Value Added Assessment Model. The difference, however, was not statistically significant (t=.934; p=.357). On the eighth Grade mathematics NAEP, Black students enrolled in Value Added Assessment Model states outperformed their counterparts in non Value Added Assessment Model states by an average of 3.72 points. The difference was not found to be statistically significant (t=.598; p=.554).

Research Question 5) To what extent, if any, is there a difference in student achievement in National School Lunch Program eligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of National School Lunch Program eligible eighth grade NAEP tested students enrolled in Value Added Assessment Model states versus their counterparts in non Value Added Assessment Model states. The matched states, Ohio and Michigan will be discussed first.

Ohio and Michigan National School Lunch Program eligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade reading tested students in Ohio had a mean NAEP score of 254. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 254 for Michigan's National School Lunch Program eligible tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 20th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 28th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 31st out of 50 states in 2003 to 31st out of 50 states in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were -3. Changes in Michigan's NAEP reading scores were down 3 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 3.71; p=.106. The mean difference between Ohio and Michigan from 2002-2013 was 3.71 in favor of Ohio. That difference in turn was not statistically significant.

Table 28. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Eligible tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	257	257	0
2003	251	247	4
2005	251	246	5
2007	251	244	7
2009	255	247	8
2011	255	253	2
2013	254	254	0
Change	-3	-3	0

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade mathematics tested students in Ohio had a mean NAEP score of 274. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 265 for Michigan's National School Lunch Program eligible tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 18th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 17th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 34th out of 50 states in 2003 to 44th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Ohio from 2002-2013 were 11. Changes in Michigan's NAEP mathematics scores were 8 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 8.0; p<.008. The mean difference between Ohio and Michigan from 2002-2013 was 8.0 in favor of Ohio. That difference in turn was statistically significant with an effect size of d=1.92.

Table 29. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Eligible tested students in Ohio and Michigan 2003-2013

Year	Ohio	Michigan	Difference
2003	263	257	6
2005	265	258	7
2007	268	259	9
2009	269	260	9
2011	274	266	8
2013	274	265	9
Change	11	8	3

Pennsylvania and Virginia all National School Lunch Program eligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade reading tested students in Pennsylvania had a mean NAEP score of 258. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 251 for Virginia's National School Lunch Program eligible tested population. Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 30th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 13th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 16th out of 50 states in 2003 to 40th out of 50 states in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 12. Changes in Virginia's NAEP reading scores were down 5 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -1.29; p=.491. The mean difference between Pennsylvania and Virginia from 2002-2013 was -1.29 in favor of Pennsylvania. That difference in turn was not statistically significant.

Table 30. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Eligible tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	246	256	-10
2003	247	252	-5
2005	247	253	-6
2007	253	252	1
2009	253	251	2
2011	252	250	2
2013	258	251	7
Change	12	-5	17

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 273. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 270 for Virginia's National School Lunch Program eligible tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 34th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 18th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 24th out of 50 states in 2003 to 28th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 16. Changes in Virginia's NAEP mathematics scores were 9 for the time period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -.833; p=.768. The mean difference between Pennsylvania and Virginia from 2003-2013 was -.833 in favor of Virginia. That difference in turn was not statistically significant.

Table 31. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Eligible tested students in Pennsylvania and Virginia 2003-2013

Year	Pennsylvania	Virginia	Difference
2003	257	261	-5
2005	262	263	-1
2007	267	268	-1
2009	268	268	0
2011	268	270	-2
2013	273	270	3
Change	16	9	8

Tennessee and Georgia NAEP all National School Lunch Program eligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade reading tested students in Tennessee had a mean NAEP score of 256. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 255 for Georgia's National School Lunch Program eligible tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 25th out of 36 states (only 36 states reported disaggregated scores for the subgroup) on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 22nd out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 30th out of 36 states in 1998 to 24th out of 50 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 16. Changes in Georgia's NAEP reading scores were 15 during the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = .75; p=.067. The mean difference between Tennessee and Georgia from 1998-2013 was .75 in favor of Tennessee. That difference in turn was not statistically significant.

Table 32. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Eligible tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	240	240	0
2002	246	245	1
2003	245	243	2
2005	246	243	3
2007	247	247	0
2009	250	249	1
2011	250	253	-3
2013	256	255	1
Change	16	15	1

The state aggregate on the 2013 NAEP of National School Lunch Program eligible eighth grade mathematics tested students in Tennessee had a mean NAEP score of 265. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 267 for Georgia's National School Lunch Program eligible tested population. Since Tennessee mandated VAAM statewide in 1992, Tennessee went from 34th out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the1996 NAEP (the first state aggregated published state ranking NAEP results available) to 46th out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 37th out of 40 states in 1996 to 37th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 16. Changes in Georgia's NAEP mathematics scores were 15 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -2.13; p=.638. The mean difference between Tennessee and Georgia from 1998-2013 was -2.13 in favor of Georgia. That difference in turn was not statistically significant.

Table 33. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Eligible tested students in Tennessee and Georgia 1996-2013

Year	Tennessee	Georgia	Difference
1996	246	242	4
2000	242	246	-4
2003	244	248	-4
2005	250	253	-3
2007	256	257	-1
2009	262	262	0
2011	261	265	-4
2013	262	267	-5
Change	16	15	1

Combined Comparison of VAAM and non VAAM States for Research Question 5

National School Lunch Program eligible eighth grade NAEP tested students which were enrolled in Value Added Assessment Model states outperformed their counterparts in non Value Added Assessment Model states by an average of 1.72 in the area of reading. The difference, however, was not found to be statistically significant (t=1.37; p=.180). In the area of mathematics, Value Added Assessment Model states manifested an average performance advantage of 1.33 points over their counterparts in non Value Added Assessment Model states. The difference was not found to be statistically significant (t- .567; p=.574).

Research Question 6) To what extent, if any, is there a difference in student achievement in National School Lunch Program ineligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of National School Lunch Program ineligible eighth grade NAEP tested students by enrollment in Value Added Assessment Model states versus non Value Added Assessment Model states. The matched states, Ohio and Michigan will be discussed first.

Ohio and Michigan National School Lunch Program ineligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade reading tested students in Ohio had a mean NAEP score of 280. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 276 for Michigan's National School Lunch Program ineligible tested population. Since Ohio mandated VAAM statewide in 2007, Ohio went from 19th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 11th out of 50 states in 2013 with Ohio's matched state, Michigan, went from 21st out of 50 states in 2003 to 31st out of 50 states in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were 7. Changes in Michigan's NAEP reading scores were 6 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 4.43; $p<.015$. The mean difference between Ohio and Michigan from 2002-2013 was 4.43 in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=1.52$.

Table 34. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Ineligible tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	273	270	3
2003	273	272	0
2005	274	267	7
2007	275	268	7
2009	276	271	5
2011	278	274	4
2013	280	276	4
Change	7	6	1

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade mathematics tested students in Ohio had a mean NAEP score of 301. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 293 for Michigan's National School Lunch Program ineligible tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 17th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 9th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 31st out of 50 states in 2003 to 39th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Ohio from 2002-2013 were 12. Changes in Michigan's NAEP mathematics scores were 8 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 6.33; p<.026. The mean difference between Ohio and Michigan from 2002-2013 was 6.33 in favor of Ohio. That difference in turn was statistically significant with an effect size of d=1.50.

Table 35. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Ineligible tested students in Ohio and Michigan 2003-2013

Year	Ohio	Michigan	Difference
2003	289	285	4
2005	290	285	5
2007	293	285	8
2009	294	289	5
2011	299	291	8
2013	301	293	8
Change	12	8	4

Pennsylvania and Virginia all National School Lunch Program ineligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade reading tested students in Pennsylvania had a mean NAEP score of 282. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 276 for Virginia's National School Lunch Program ineligible tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 27th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 7th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 9th out of 50 states in 2003 to 26th out of 50 states in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 8. Changes in Virginia's NAEP reading scores were 12 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = 2.57; p=.112. The mean difference between Pennsylvania and Virginia from 2002-2013 was 2.57 in favor of Pennsylvania. That difference in turn was not statistically significant.

Table 36. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Ineligible tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	274	274	0
2003	271	274	-3
2005	276	273	3
2007	275	272	3
2009	279	272	7
2011	278	276	3
2013	282	276	5
Change	8	2	5

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 301. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 298 for Virginia's National School Lunch Program ineligible tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 21st out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 7th out of 50 states in 2013. Pennsylvania's matched state, Virginia, went from 16th out of 50 states in 2003 to 19th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2003-2013 were 13. Changes in Virginia's NAEP mathematics scores were 9 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = .333; p=.900. The mean difference between Pennsylvania and Virginia from 2003-2013 was .333 in favor of Pennsylvania. That difference in turn was not statistically significant.

Table 37. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Ineligible tested students in Pennsylvania and Virginia 2003-2013

Year	Pennsylvania	Virginia	Difference
2003	288	289	-1
2005	289	292	-3
2007	294	295	-1
2009	298	294	4
2011	298	298	0
2013	301	298	3
Change	13	9	4

Tennessee and Georgia NAEP all National School Lunch Program ineligible eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade reading tested students in Tennessee had a mean NAEP score of 276. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 278 for Georgia's National School Lunch Program ineligible tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 25th out of 36 states on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 33rd out of 50 states in 2013. Tennessee's matched state, Georgia, moved from 24th out of 36 states in 1998 to 19th out of 50 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 13. Changes in Georgia's NAEP reading scores were 10 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -1.88; p=.293. The mean difference between Tennessee and Georgia from 1998-2013 was -1.88 in favor of Georgia. That difference in turn was not statistically significant.

Table 38. Mean composite scale score on the eighth grade reading NAEP National School Lunch Program Ineligible tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	267	268	-1
2002	268	267	1
2003	265	269	-4
2005	268	269	-1
2007	269	270	-1
2009	269	272	-3
2011	270	274	-4
2013	276	278	-2
Change	13	10	3

The state aggregate on the 2013 NAEP of National School Lunch Program ineligible eighth grade mathematics tested students in Tennessee had a mean NAEP score of 292. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 296 for Georgia's National School Lunch Program ineligible tested population. Since Tennessee mandated VAAM statewide in 1992, Tennessee went from 24th out of 40 states (only 40 states reported disaggregated scores for the subgroup) on the 1996 NAEP (the first state aggregated published state ranking NAEP results available) to 42nd out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 31st out of 40 states in 1996 to 25th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 21. Changes in Georgia's NAEP mathematics scores were 23 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -4.13; p=.279. The mean difference between Tennessee and Georgia from 1998-2013 was -4.13 in favor of Georgia. That difference in turn was not statistically significant.

Table 39. Mean composite scale score on the eighth grade mathematics NAEP National School Lunch Program Eligible tested students in Tennessee and Georgia 1996-2013

Year	Tennessee	Georgia	Difference
1996	271	273	-2
2000	273	278	-5
2003	279	284	-5
2005	282	285	-3
2007	284	287	-3
2009	285	290	-5
2011	287	293	-6
2013	292	296	-4
Change	21	23	-2

Combined Comparison of VAAM and non VAAM States for Research Question 6

National School Lunch Program ineligible eighth grade NAEP tested students did not manifest statistically significant performance in either reading or mathematics NAEP performance within comparisons of Value Added Assessment Model states and non Value Added Assessment Model states (reading: t=1.32; p=.197; mathematics: t=.411; p=.684).

<u>Research Question 7) To what extent, if any, is there a difference in student achievement in Exceptional Education eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?</u>

The researcher's objective was to gain an understanding to what extent, if any, that there was in student achievement in the states that had implemented the Value Added Assessment Model as compared to the states that had not implemented the Value Added Assessment Model. The research conducted the comparison by performing an Independent *t* Test on the data. The data were arranged in two categories: Value Added Assessment Model implementing states and non Value Added Assessment Model implementing states. The Independent *t* Test that was performed was on the matched pairs comparing the Value Added Assessment Model implementing states with the matched pair on the non Value Added Assessment Model states. If the difference was found to be statistically significant <.05, then the researcher performed a Cohen's *d* to determine the magnitude of the effect. Upon completion of the matched state comparisons, then the researcher ran a summative Independent *t* Test on the data set of all of the matched pair states, comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model states.

The following depicts the pairwise comparisons of ESE student performance in Value Added Assessment Model states versus non Value Added Assessment Model states. The matched states, Ohio and Michigan will be discussed first.

Ohio and Michigan Exceptional education eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of exceptional education eighth grade reading tested students in Ohio had a mean NAEP score of 229. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 235 for Michigan's exceptional education tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 30th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 28th out of 50 states in 2013 with Ohio's matched state, Michigan, moved from 23rd out of 50 states in 2003 to 15th out of 50 states in 2013.

Changes in the NAEP reading scores for Ohio from 2002-2013 were 4. Changes in Michigan's NAEP reading scores were 1 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 2.28; p=.412. The mean difference between Ohio and Michigan from 2002-2013 was 2.28 in favor of Ohio. That difference in turn was not statistically significant.

Table 40. Mean composite scale score on the eighth grade reading NAEP Exceptional Education tested students in Ohio and Michigan 2002-2013

Year	Ohio	Michigan	Difference
2002	225	234	-9
2003	225	228	-3
2005	231	230	1
2007	235	224	11
2009	238	222	16
2011	236	230	6
2013	229	235	-6
Change	4	1	3

The state aggregate on the 2013 NAEP of exceptional education eighth grade mathematics tested students in Ohio had a mean NAEP score of 252. In comparison, Ohio's matched state, Michigan, had a mean NAEP score of 243 for Michigan's exceptional education tested population. Since Ohio mandated VAAM statewide in 2007 Ohio went from 24th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 11th out of 50 states in 2013 with Ohio's matched state, Michigan, going from 35th out of 50 states in 2003 to 34th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Ohio from 2002-2013 were 4. Changes in Michigan's NAEP mathematics scores were 1 for the same period. A t-test of independent means on the two groups, Ohio and Michigan, was run with a mean difference = 10.67; $p<.001$. The mean difference between Ohio and Michigan from 2002-2013 was 10.67 in favor of Ohio. That difference in turn was statistically significant with an effect size of $d=2.69$.

Table 41. Mean composite scale score on the eighth grade mathematics NAEP Exceptional Education tested students in Ohio and Michigan 2003-2013

Year	Ohio	Michigan	Difference
2003	245	240	4
2005	251	243	8
2007	250	238	12
2009	255	239	16
2011	258	246	12
2013	255	244	11
Change	4	1	3

Pennsylvania and Virginia all exceptional education eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of exceptional education eighth grade reading tested students in Pennsylvania had a mean NAEP score of 240. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 234 for Virginia's exceptional education tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 26th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 8th out of 50 states in 2013 Pennsylvania's matched state, Virginia, dropped from 8th out of 50 states in 2003 to 21st out of 50 states in 2013.

Changes in the NAEP reading scores for Pennsylvania from 2002-2013 were 12. Changes in Virginia's NAEP reading scores were down 5 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = 1.29; p=.4.91. The mean difference between Pennsylvania and Virginia from 2002-2013 was 1.29 in favor of Pennsylvania. That difference in turn was not statistically significant.

Table 42. Mean composite scale score on the eighth grade reading NAEP Exceptional Education tested students in Pennsylvania and Virginia 2002-2013

Year	Pennsylvania	Virginia	Difference
2002	246	256	-10
2003	247	252	-5
2005	247	253	-6
2007	253	252	-1
2009	253	251	2
2011	252	250	2
2013	258	251	7
Change	12	-5	17

The state aggregate on the 2013 NAEP of exceptional education eighth grade mathematics tested students in Pennsylvania had a mean NAEP score of 258. In comparison, Pennsylvania's matched state, Virginia, had a mean NAEP score of 251 for Virginia's exceptional education tested population. Since Pennsylvania mandated VAAM statewide in 2006 Pennsylvania went from 26th out of 50 states on the 2003 NAEP (the test prior to VAAM implementation) to 5th out of 50 states in 2013 Pennsylvania's matched state, Virginia, went from 3rd out of 50 states in 2003 to 15th out of 50 states in 2013.

Changes in the NAEP mathematics scores for Pennsylvania from 2002-2013 were 15. Changes in Virginia's NAEP mathematics scores were down 3 for the same period. A t-test of independent means on the two groups, Pennsylvania and Virginia was run with a mean difference = -4.17; p=.145. The mean difference between Pennsylvania and Virginia from 2003-2013 was -4.17 in favor of Virginia. That difference in turn was not statistically significant.

Table 43. Mean composite scale score on the eighth grade mathematics NAEP Exceptional Education tested students in Pennsylvania and Virginia 2003-2013

Year	Pennsylvania	Virginia	Difference
2003	244	255	-11
2005	245	256	-11
2007	254	260	-6
2009	254	253	1
2011	252	257	-4
2013	259	252	7
Change	15	-3	18

Tennessee and Georgia NAEP all exceptional education eighth grade students who were tested on the NAEP

The state aggregate on the 2013 NAEP of exceptional education eighth grade reading tested students in Tennessee had a mean NAEP score of 226. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 230 for Georgia's exceptional education tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 29th out of 31 states (only 31 states reported disaggregated scores for the subgroup) on the 1998 NAEP (the first state aggregated published state ranking NAEP results available) to 38th out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 17th out of 31 states in 1998 to 26th out of 50 states in 2013.

Changes in the NAEP reading scores for Tennessee from 1998-2013 were 15. Changes in Georgia's NAEP reading scores were down 3 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -1.5; p=.691. The mean difference between Tennessee and Georgia from 1998-2013 was -1.54 in favor of Georgia. That difference in turn was not statistically significant.

Table 44. Mean composite scale score on the eighth grade reading NAEP Exceptional Education tested students in Tennessee and Georgia 1998-2013

Year	Tennessee	Georgia	Difference
1998	211	225	-14
2002	223	216	7
2003	235	212	23
2005	216	226	-10
2007	228	231	-3
2009	223	224	-1
2011	224	234	-10
2013	226	230	-4
Change	15	5	18

The state aggregate on the 2013 NAEP of exceptional education eighth grade mathematics tested students in Tennessee had a mean NAEP score of 235. In comparison, Tennessee's matched state, Georgia, had a mean NAEP score of 242 for Georgia's exceptional education tested population. Since Tennessee mandated VAAM statewide in 1992 Tennessee went from 35th out of 37 states (only 37 states reported disaggregated scores for the subgroup) on the 2000 NAEP (the first state aggregated published state ranking NAEP results available) to 45th out of 50 states in 2013 with Tennessee's matched state, Georgia, going from 19th out of 37 states in 2000 to 33rd out of 50 states in 2013.

Changes in the NAEP mathematics scores for Tennessee from 1992-2013 were 15. Changes in Georgia's NAEP mathematics scores were down 3 for the same period. A t-test of independent means on the two groups, Tennessee and Georgia, was run with a mean difference = -4.43; p = .313. The mean difference between Tennessee and Georgia from 1998-2013 was -4.43 in favor of Georgia. That difference in turn was not statistically significant.

Table 45. Mean composite scale score on the eighth grade mathematics NAEP Exceptional Education tested students in Tennessee and Georgia 2000-2013

Year	Tennessee	Georgia	Difference
2000	216	232	-16
2003	242	234	8
2005	237	241	-4
2007	246	246	0
2009	239	245	-6
2011	239	244	-5
2013	236	244	-8
Change	20	12	8

Combined Comparison of VAAM and non VAAM States for Research Question 7

Eighth grade reading NAEP exceptional education tested students in Value Added Assessment Model states manifested a slight performance edge (mean difference=.555) over their peers enrolled in states not adopting Value Added Assessment Model. The difference was not statistically significant (t=.255; p=.800). In mathematics, an average performance difference of 1.33 in favor of exceptional education enrolled in Value Added Assessment Model adopting states existed. The difference, however, was not found to be statistically significant (t=.544; p=.590).

Summary

The data were presented and analyzed in Chapter 4 in an effort to answer the seven research questions which guided this study. The data were presented in both narrative and in tables to provide the clearest picture possible of the researcher's findings when comparing the Value Added Assessment Model implementing states with the non Value Added Assessment Model implementing states.

CHAPTER 5
CONCLUSIONS AND RECOMENDATIONS

Introduction

Chapter 5 is a summary of the entirety of the research and a discussion of the findings, inferences, and outcomes of the research, along with conclusions. Chapter 5 also contains discussion of the limitations of the study. Plausible implications of the findings of the study are also presented. Implications are provided to give a framework of understanding for the continued use of the Value Added Assessment Model as a summative measure of educational success. Additionally, directions for future possible research are recommended.

Inferences and outcomes are discussed to identify the usefulness of this study for further academic discussion regarding accountability as it relates to the Value Added Assessment Model in education. Directions for future research are identified to give direction for further dialogue and research in the area of the Value Added Assessment Model and student academic success. Limitations are listed in order for the reader to know the confines in which the conclusions of this research should be used. Concluding statements are provided to summarize the practical application of this study and the contribution that this study may provide to the current understanding of the Value Added Assessment Model and measure of educational success for stakeholders. Final discourse is provided for the reader with an understanding of the potential impact of use Value Added Assessment Model legislation on student achievement.

Purpose of the Study

The Value Added Assessment Model (VAAM) quantifies the impact of the academic environment on students' standardized test scores to determine the students' academic environment contributed to higher than or lower than anticipated test scores (McCaffrey, Lockwood, Koretz & Hamilton, 2003). The purpose of this study was to collect and analyze data regarding the impact of VAAM on student achievement, in eighth grade students, in Ohio, Pennsylvania, and Tennessee which had adopted VAAM. They were matched with state aggregate student achievement of eighth grade students, in Georgia, Michigan and Virginia. These states had not adopted VAAM. The six states were matched, Ohio and Michigan, Pennsylvania and Virginia, and Tennessee and Georgia. The researcher (a) examined the mean difference of VAAM implementation and student NAEP scores in the six states, (b) examined whether a mean difference existed between VAAM implementation and student achievement of White students, (b) examined whether a mean difference existed between VAAM implementation and student achievement of Black students, (c) examined whether a mean difference existed between VAAM implementation and student achievement amongst economically disadvantaged students (as indicated based on participation in the National School Lunch Program), and (d) examined whether a mean difference existed between VAAM implementation and student achievement amongst economically advantaged students (as indicated based on nonparticipation in the National School Lunch Program) and (e) examined whether a mean difference existed between VAAM implementation and student achievement of Exceptional Student Education students.

Methodology

The researcher analyzed data from states that implemented the Value Added Assessment Model for more than five years with matched states that had not implemented the Value Added Assessment Model. The researcher analyzed policy and legislation surrounding the Value Added Assessment Model implementing states' use of the Value Added Assessment Model in educator accountability. The states have the authority to govern over their own educational policies as is outlined in the 10th Amendment to the United States Construction (U.S. Constitution). The researcher also analyzed the eighth grade reading and mathematics NAEP data from the three Value Added Assessment Model states with their matched non-Value Added Assessment Model states for the years preceding the full implementation of the Value Added Assessment Model up to 2013. The researcher conducted email interviews with state designees from Ohio, Pennsylvania and Tennessee. The NAEP is considered to be the highest standard in measuring student academic performance (Jones & Olkin, 2004).

Population

The population included state aggregated data of eighth graders in Tennessee, Ohio, and Pennsylvania, who took the NAEP in school year 2012-2013 versus the test data of eighth graders in Michigan, Virginia and Georgia who took the NAEP in school year 2012-2013. The states were matched by analyzing geographic, demographic, and population statistics for the individual states. The states that had implemented any variation of VAAM were eliminated as potential matches for Ohio, Pennsylvania, and Tennessee. The researcher matched Georgia with Tennessee, Michigan with Ohio, and Virginia with Tennessee.

Summary and Discussion of Findings

The researcher's intent in performing this study was to analyze if there was a relationship between states' usage of the Value Added Assessment Model and eighth grade student academic performance as indicated on the reading and mathematics NAEP.

Research Question 1

What are the Value Added Assessment Model designs in Ohio, Pennsylvania and Tennessee?

It was determined through the analysis of the legislation, policies and through email interviews that Ohio, Pennsylvania and Tennessee all use the same Value Added Assessment Model, the model used is the SAS EVAAS.

Research Question 2

To what extent, if any, is there a difference in student achievement in eighth grade students, as measured by mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states without VAAM?

The findings from Research Question 2 indicated there was not a statistically significant difference in student performance on the NAEP reading assessment of all eighth grade tested students. However, the results of the NAEP mathematics assessment of all eighth grade tested students indicated a statistically significant difference in favor of states that have implemented the Value Added Assessment Model. The findings of Research Question 2 indicate that implementing the Value Added Assessment Model does have a positive impact on student mathematics performance on the NAEP mathematics assessment in the category of all eighth grade tested students.

Research Question 3

To what extent, if any, is there a difference in student achievement in White eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The findings from Research Question 3 indicated there was not a significant difference in student performance on the NAEP reading and mathematics assessment of White eighth grade tested students. The findings of Research Question 3 indicate that implementing the Value Added Assessment Model does not have a statistically significant academic impact on student performance as indicated by the mean composite scale scores of White eighth grade tested students on the NAEP reading and mathematics assessment. The accountability of using the

Value Added Assessment Model, with a result of increased quality of education is not evident in the results of the White eighth grade NAEP tested students evaluated in this study. The researcher maintains that the pressure on educators and the increased degradation of teacher autonomy and creativity would adversely impact student achievement at an inverse rate similar to any potential gains that would come from the implementation of the Value Added Assessment Model. With the White student population being the largest subgroup tested, it is very important to look at the gains that are made or not made in this category.

Research Question 4

To what extent, if any, is there a difference in student achievement in Black eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The findings from Research Question 4 indicated there was not a significant difference in student performance on the NAEP reading and mathematics assessment of Black eighth grade tested students. The findings of Research Question 4 indicated that implementing the Value Added Assessment Model does not have a statistically significant academic impact on student performance as indicated by the mean composite scale scores of Black eighth grade tested students on the NAEP reading and mathematics assessment. The implementation of the Value Added Assessment Model did not have a worthwhile impact on the Black students tested on the NAEP, with the Black subgroup traditionally scoring below the All and White subgroup, the researcher underscores the potential area for gains in this subgroup. If the implementation of the Value Added Assessment Model is incapable of making gains in the subgroups needing gains then it is presumed that the time, energy and finances be better used in a different manner.

Research Question 5

To what extent, if any, is there a difference in student achievement in National School Lunch Program eligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The findings from Research Question 5 indicated there was not a significant difference in student performance on the NAEP reading and mathematics assessment of National School Lunch Program eligible eighth grade tested students. The findings of Research Question 5 indicate that implementing the Value Added Assessment Model did not have a statistically significant academic impact on student performance as indicated by the mean composite scale scores of National School Lunch Program eligible eighth grade tested students on the NAEP reading and mathematics assessment, similar to the lack of influence on the reading and mathematics achievement of White and Black eighth grade students. This was similar to the Black subgroup potential for gains are available in the National School Lunch Program eligible subgroup and the gains did not happen.

Research Question 6

To what extent, if any, is there a difference in student achievement in National School Lunch Program ineligible eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The findings from Research Question 6 indicated there was not a significant difference in student performance on the NAEP reading and mathematics assessment of National School Lunch Program ineligible eighth grade tested students. The findings of Research Question 6 indicated that implementing the Value Added Assessment Model did not have a statistically significant academic impact on student performance as indicated by the mean composite scale scores of National School Lunch Program ineligible eighth grade tested students on the NAEP reading and mathematics assessment. Although many believe that VAAM models favor students in poverty, the VAAM did not have significant impact on achievement of those who participated in the National School Lunch Program nor on those who did not.

Research Question 7

To what extent, if any, is there a difference in student achievement in exceptional education eighth grade students, as measured by the mean composite scale score in NAEP reading and mathematics scores among states with VAAM and matched states that have not implemented VAAM?

The findings from Research Question 7 indicated a not statistically significant difference in student performance on the NAEP reading and mathematics assessment of exceptional education eighth grade tested students. The findings of Research Question 7 indicate that implementing the Value Added Assessment Model does not have a statistically significant academic impact on student performance as indicated by the mean composite scale scores of exceptional education eighth grade tested students on the NAEP reading and mathematics assessment. Similar to the Black and the National School Lunch Program eligible subgroups potential for gains are available in the exceptional education subgroup and the gains did not happen.

Implications for Practice

Through the comparison of eighth grade students' NAEP performance the lack of difference in Value Added Assessment Model states and non Value Added Assessment Model was revealed. Out of the 12 different comparisons only one area was statistically significant overall: the performance of all eighth grade tested students on the NAEP mathematics Assessment. The other 11 categories analyzed were found not to have a statistically significant difference. This information could be used to prompt Local Education Authorities (LEAs) and state and national policy makers to further investigate the funding level required and the resulting academic impact of the implementation of the Value Added Assessment Model. As indicated by this study VAAM implementation did not have a statistically positive impact on eight graders who took the reading and mathematics NAEP. If the goal of the implication of VAAM is increased academic achievement, VAAM implication should be further investigated prior to widespread usage and investment of resources.

The researcher agrees with Goldhaber & Hansen, 2010, in the fact that he Value Added Assessment Model does provide quantitative data that allows for politicians, the media and LEAs to use to measure effectiveness of education at various levels. The researcher maintains that this usage is the only justifiable usage of the data. The researcher believes that there are many factors that go into the microcosm of a school and a student that cannot be captured in an algorithm. The goal of increasing academic quality and performance is not met as a justifiable means for the widespread implementation of the Value Added Assessment Model. The researcher concurs with the research performed by Koretz, 2005, which found that teacher quality is the largest school based predictor of student achievement. As was identified and argued by Gordon, Kane & Staiger, 2006, increasing teacher quality is not done by scrutinizing teachers, but lowering teacher turnover, developing current teachers and recruiting talented teachers.

Recommendations for Future Research

1. Examining teacher effectiveness rates in states that have implemented the Value Added Assessment Model.

2. Examining the impact of the implantation of the Value Added Assessment Model on student academic performance as indicated on fourth grade students as indicated NAEP performance on reading and mathematics.

3. Examining the impact of the implementation of the Value Added Assessment Model on student academic performance as indicated by high school graduation rates.

4. Examining the impact of the implantation of the Value Added Assessment Model on student academic performance as indicated by composite scores of students taking the American College Test (ACT) or Scholastic Assessment Test (SAT) in high school.

5. Examine the impact of VAAM on mathematics on fourth grade and tenth grade NAEP scores, with analysis for correlation with mathematical standards and student achievement.

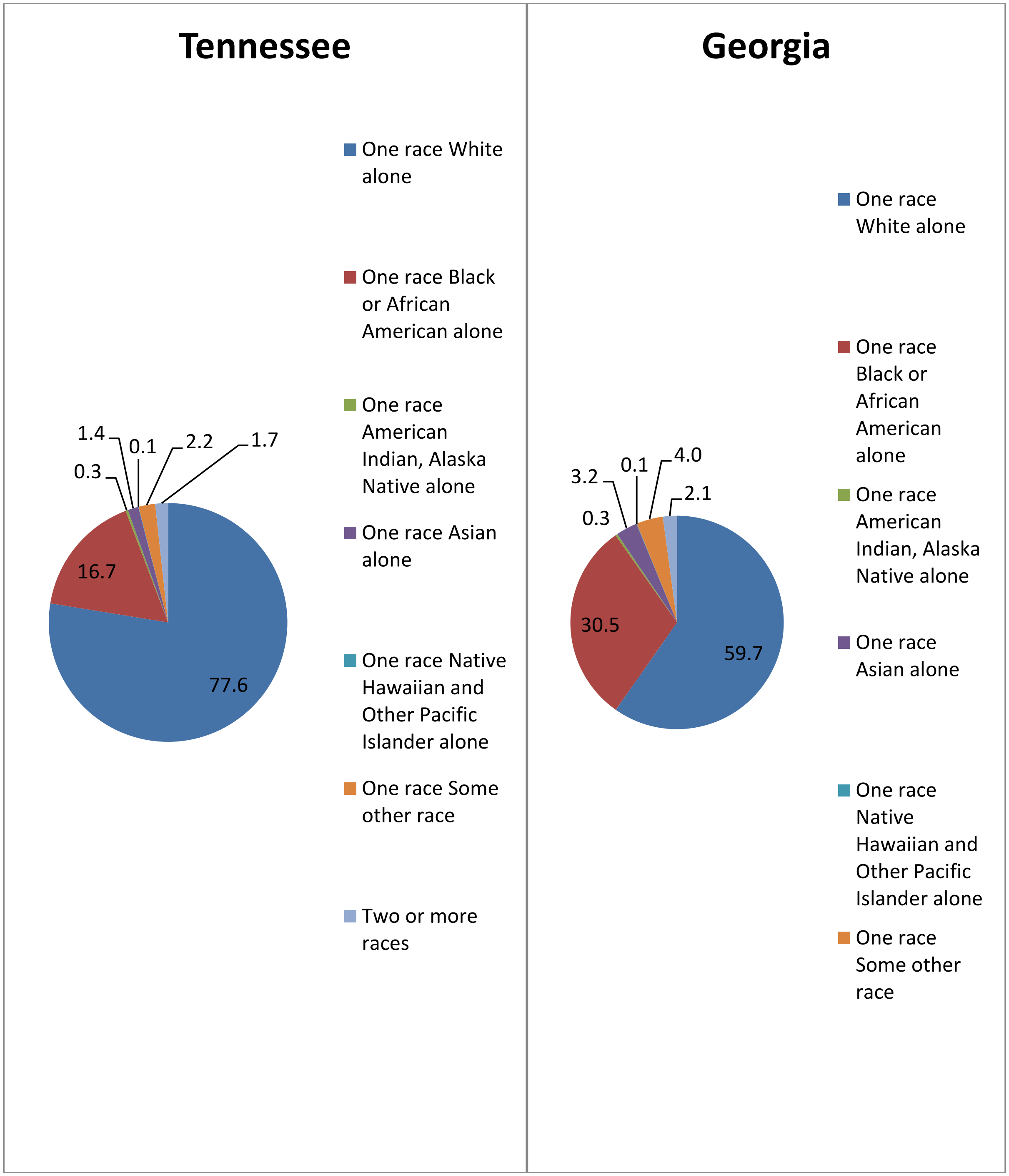
Summary

As our nation continues to call on evermore accountability in education more and more quantitative indicators of the educational effectiveness in states, districts, schools and classrooms will be required. Unfortunately, the researcher was able to find very little quantitative evidence to support the ever expanding use of the Value Added Assessment Model to improve student academic performance. The researcher maintains that the most impactful teachers possess qualities that are not measurable in mathematical formulas or through educational initiatives, but rather come from the heart of the teacher to want to make a difference in the lives of the students in their classroom.

The researcher expresses concern for the long term impact on education of increased restrictions on teachers and lower job satisfaction as a result of the increased accountability. Teachers have always been the biggest advocates for attracting the next generation of teachers into the field. With the current environment surrounding education many excellent teachers are now suggesting for their pupils to go into a field other than education. Had this been the case when I was coming of age I would have never gone into education. I became and educator because the teachers in my life advocated for the profession. The long term trend for attracting quality educators is alarming and is worth consideration.

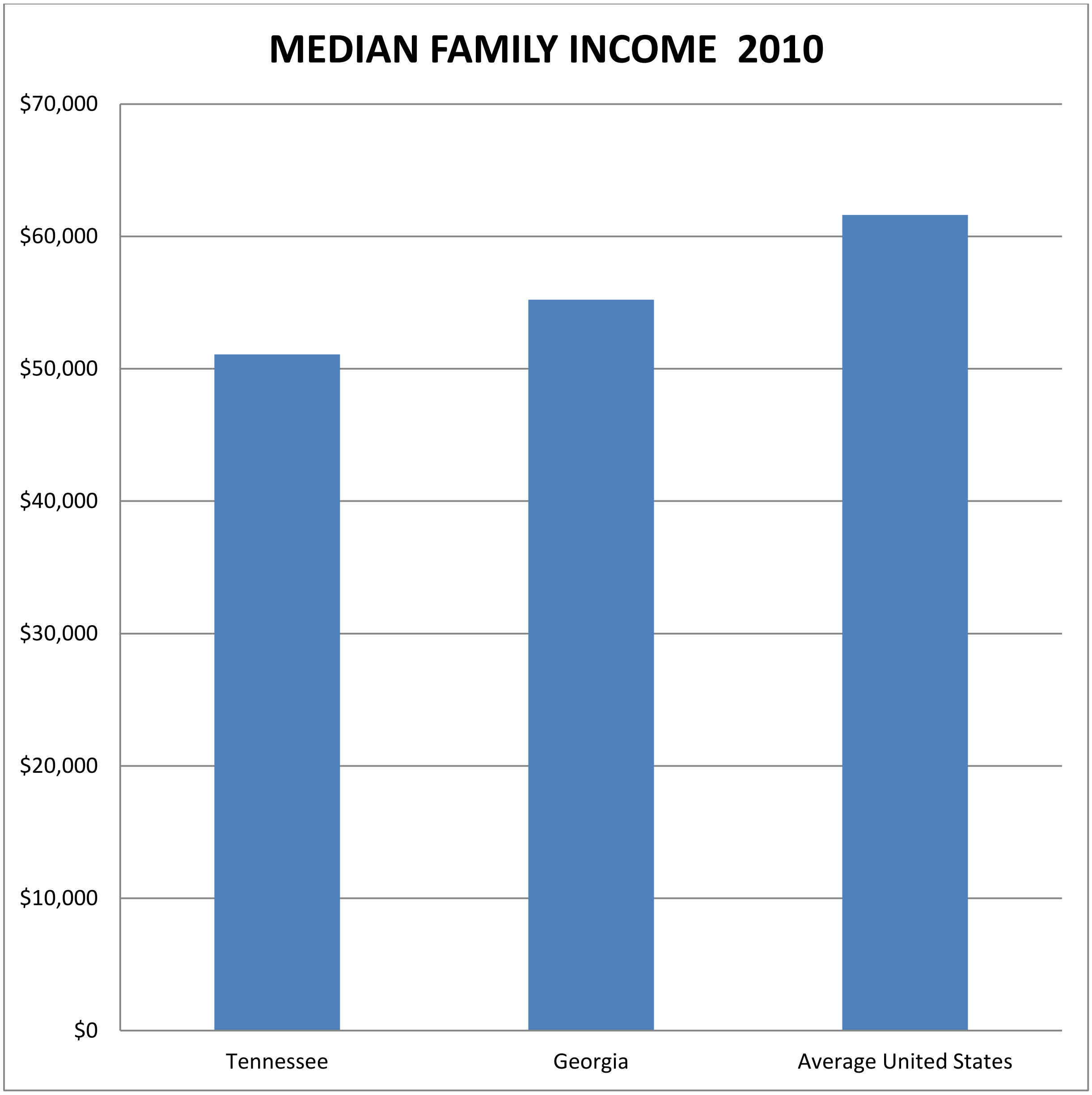
.

APPENDIX A: STATE DATA TENNESSEE AND GEORGIA



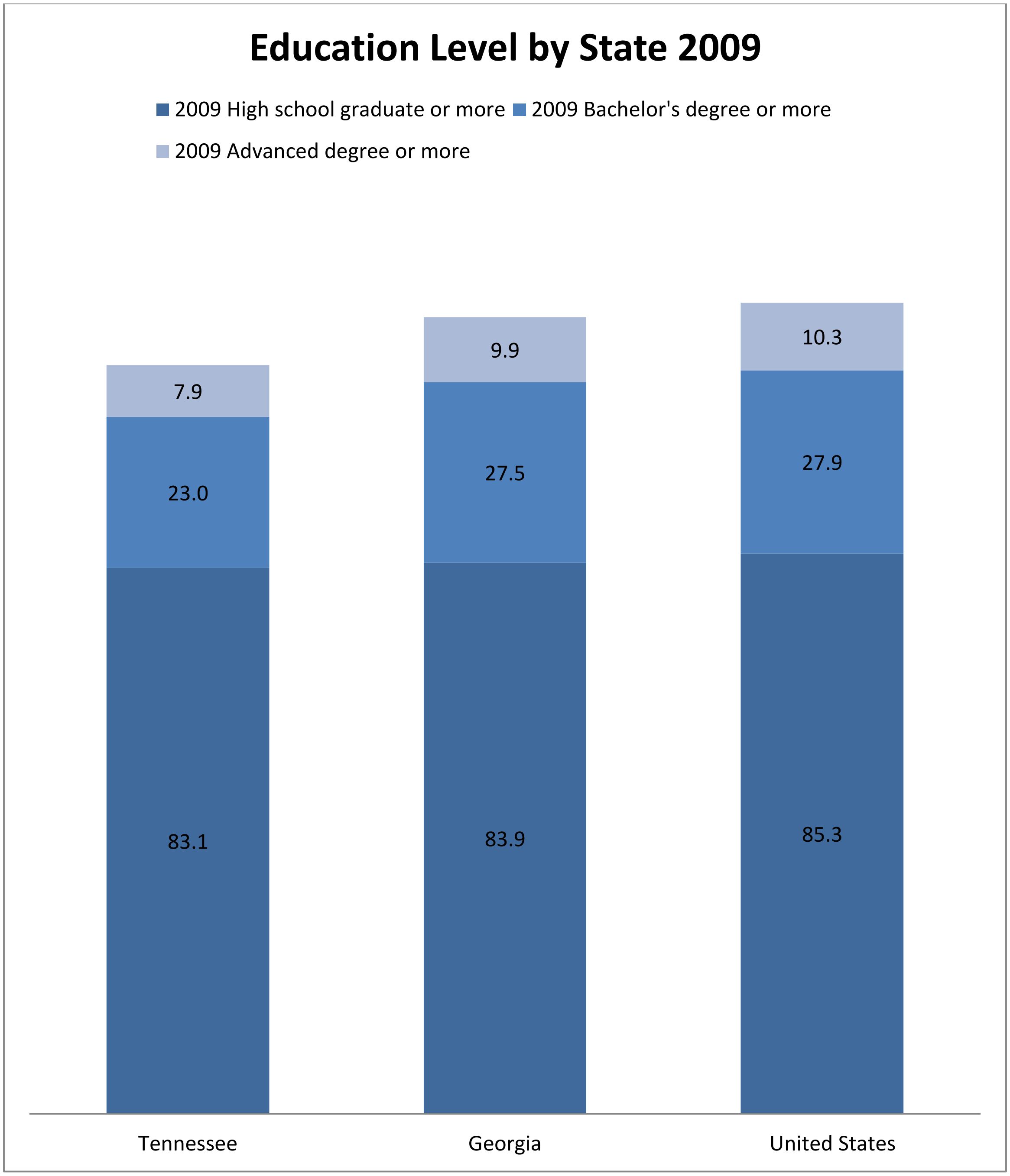
(U.S. Census, 2010)

Figure 2. Demographic Chart: Tennessee and Georgia



(U.S. Census, 2010)

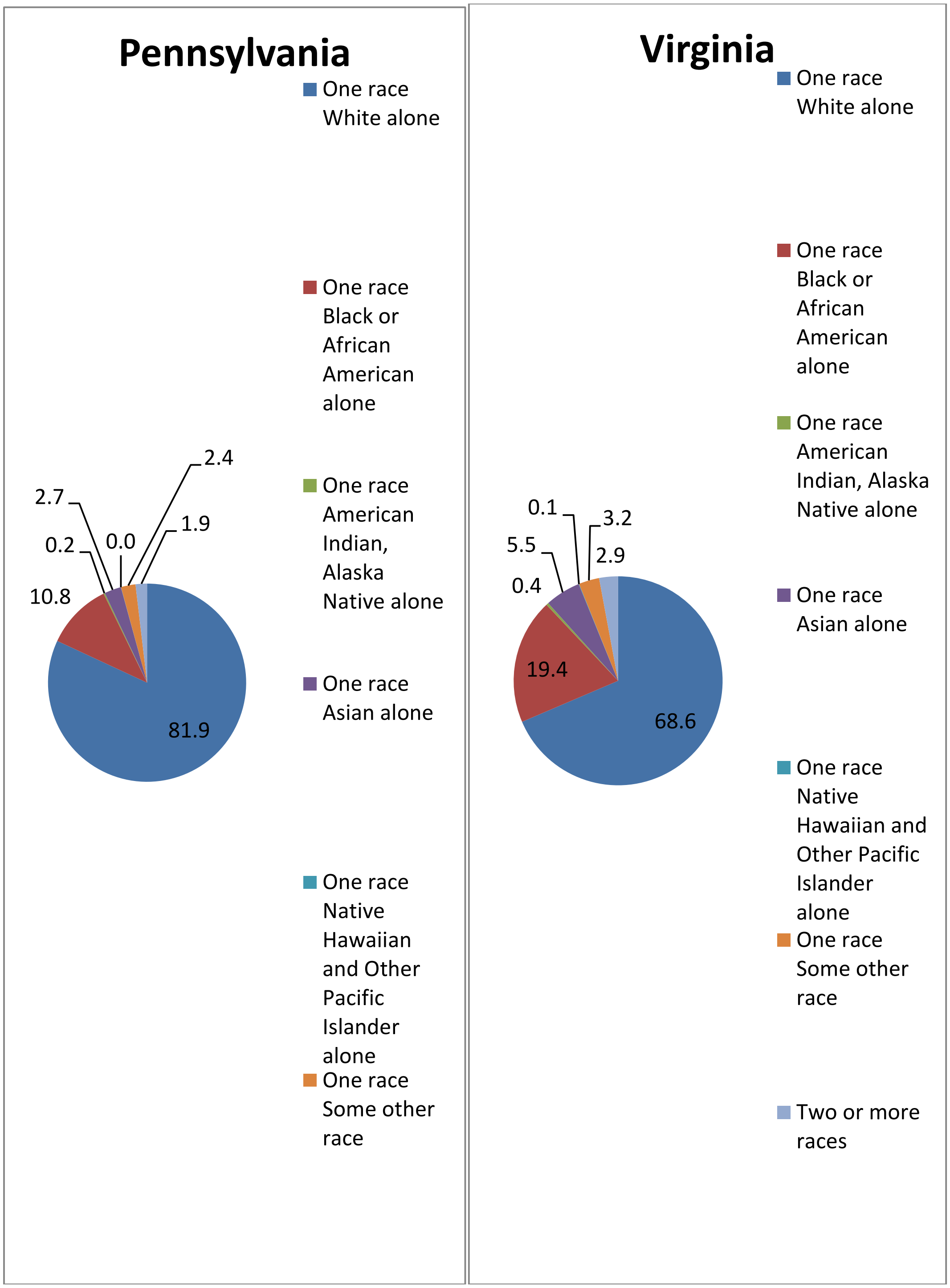
Figure 3. Median Family Income: Tennessee and Georgia



(U.S. Census, 2010)

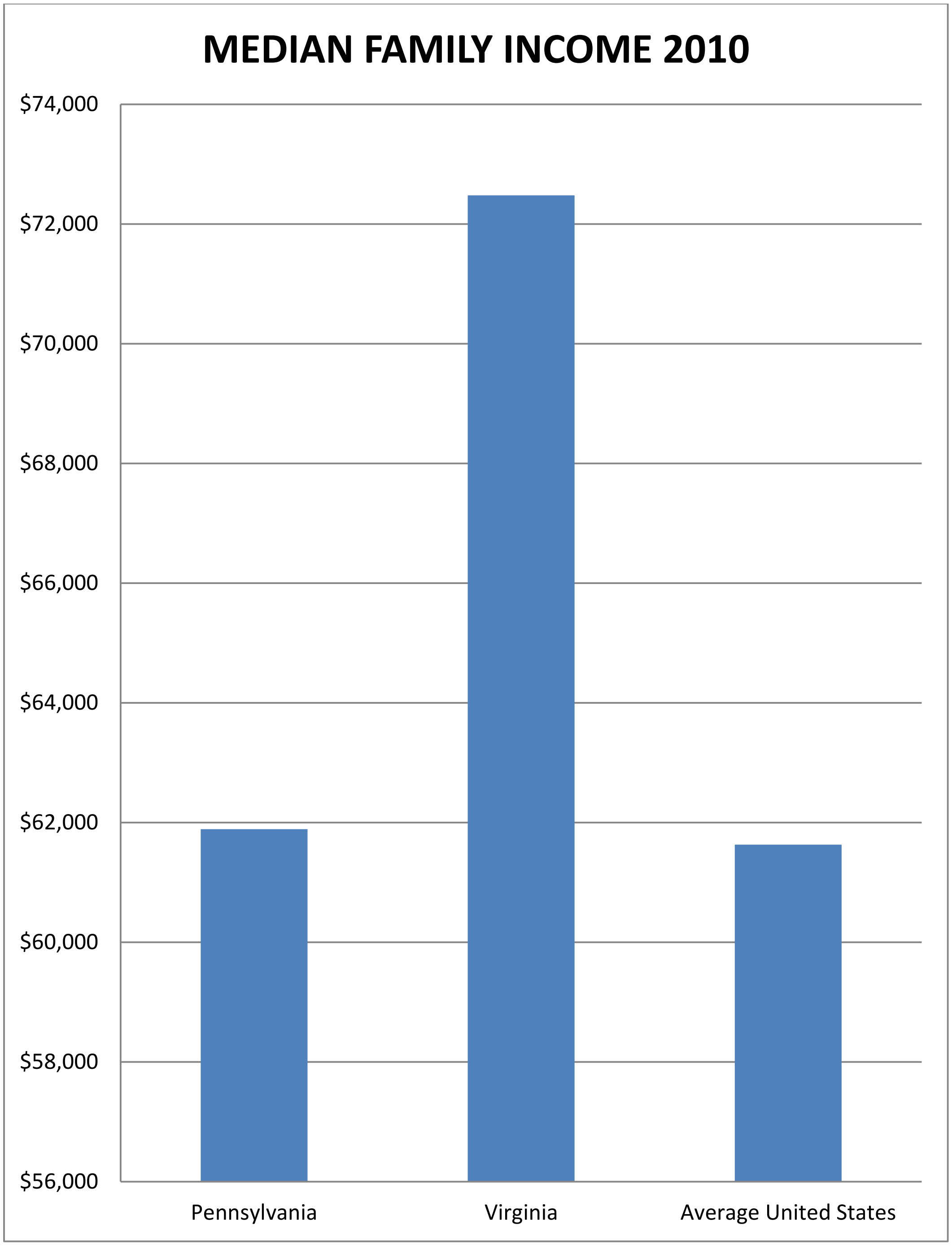
Figure 4. Education Level by State: Tennessee and Georgia

APPENDIX B: STATE DATA PENNSYLVANIA AND VIRGINIA



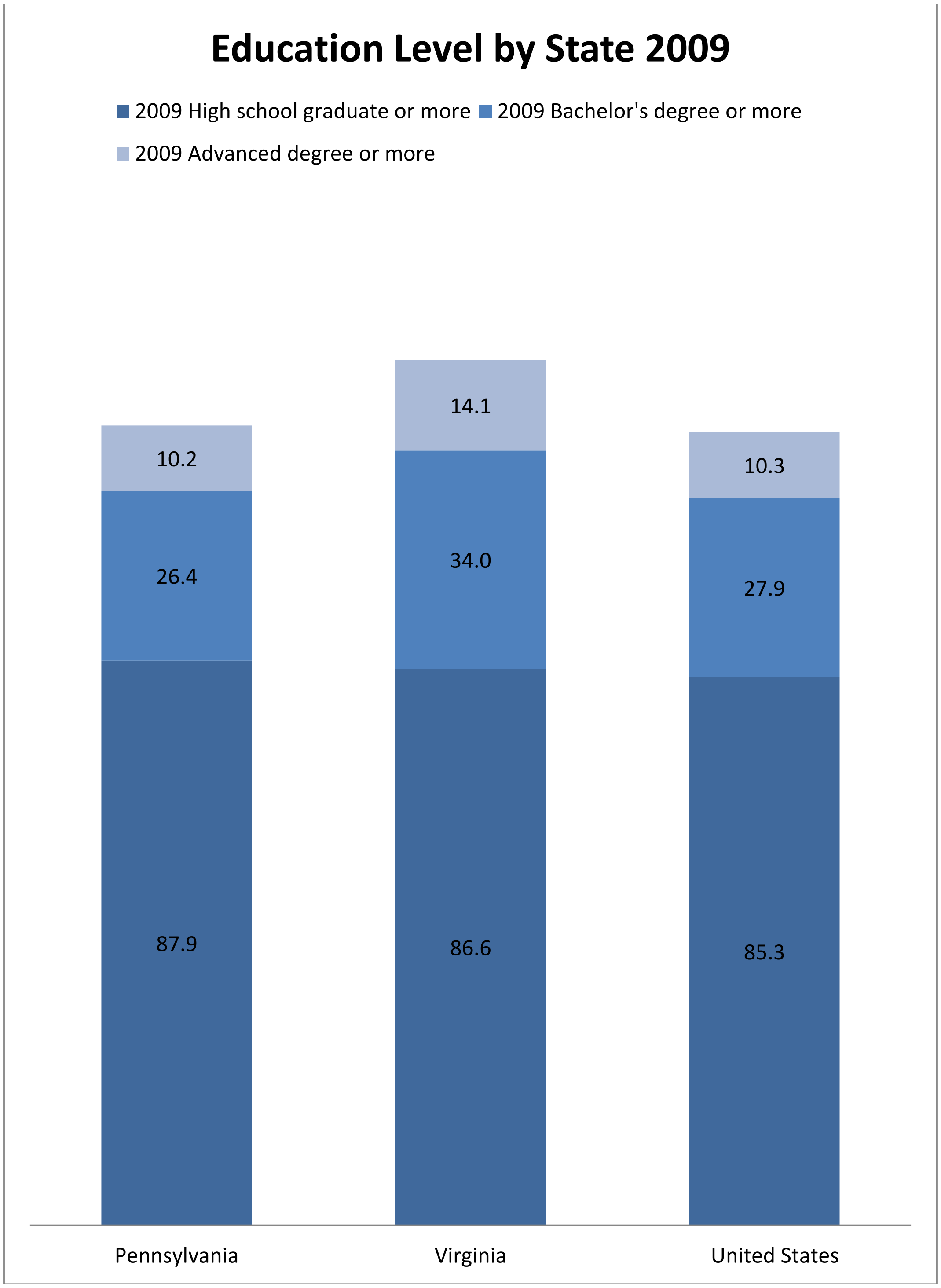
(U.S. Census, 2010)

Figure 5. Demographic Chart: Pennsylvania and Virginia



(U.S. Census, 2010)

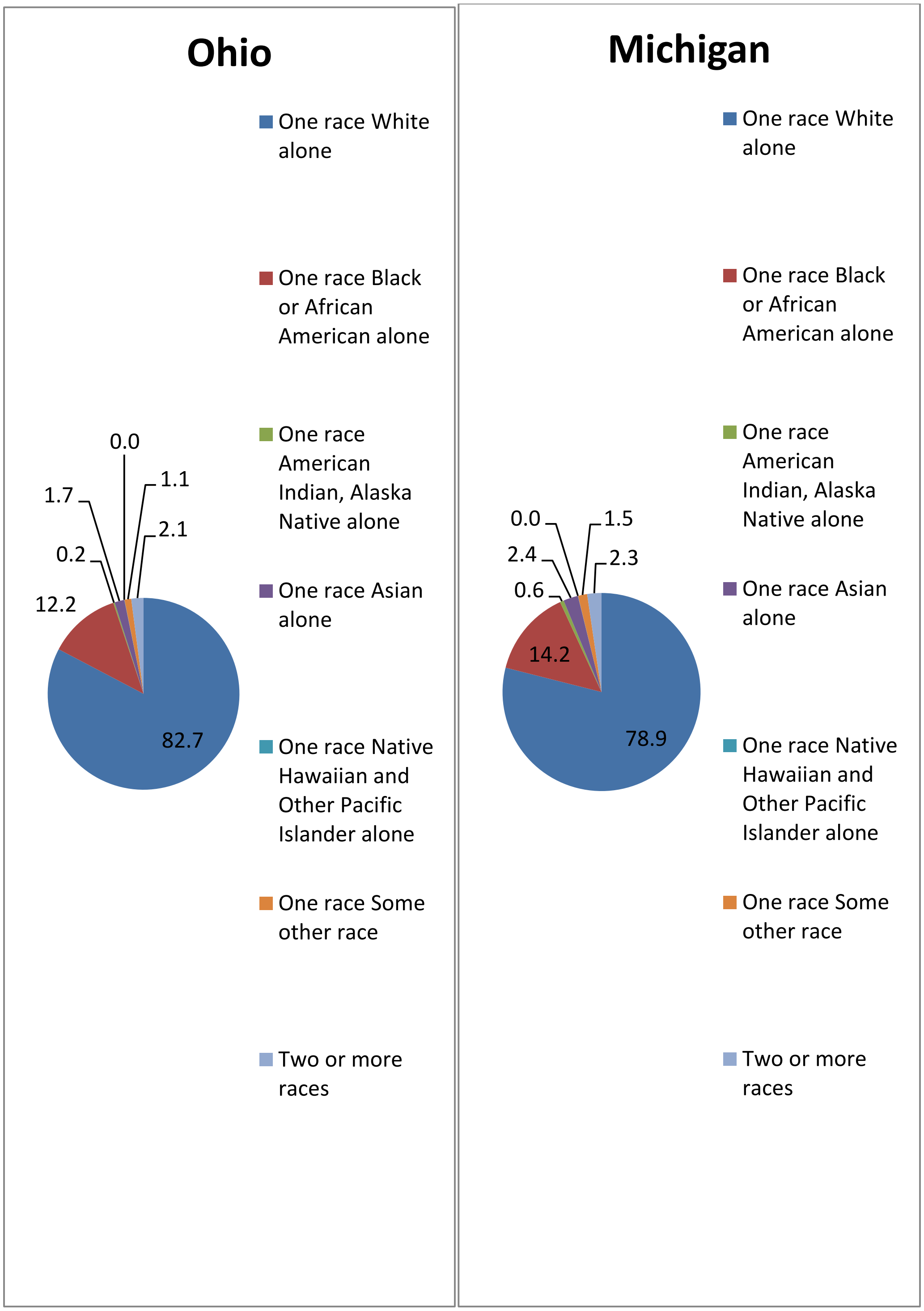
Figure 6. Median Family Income: Pennsylvania and Virginia



(U.S. Census, 2010)

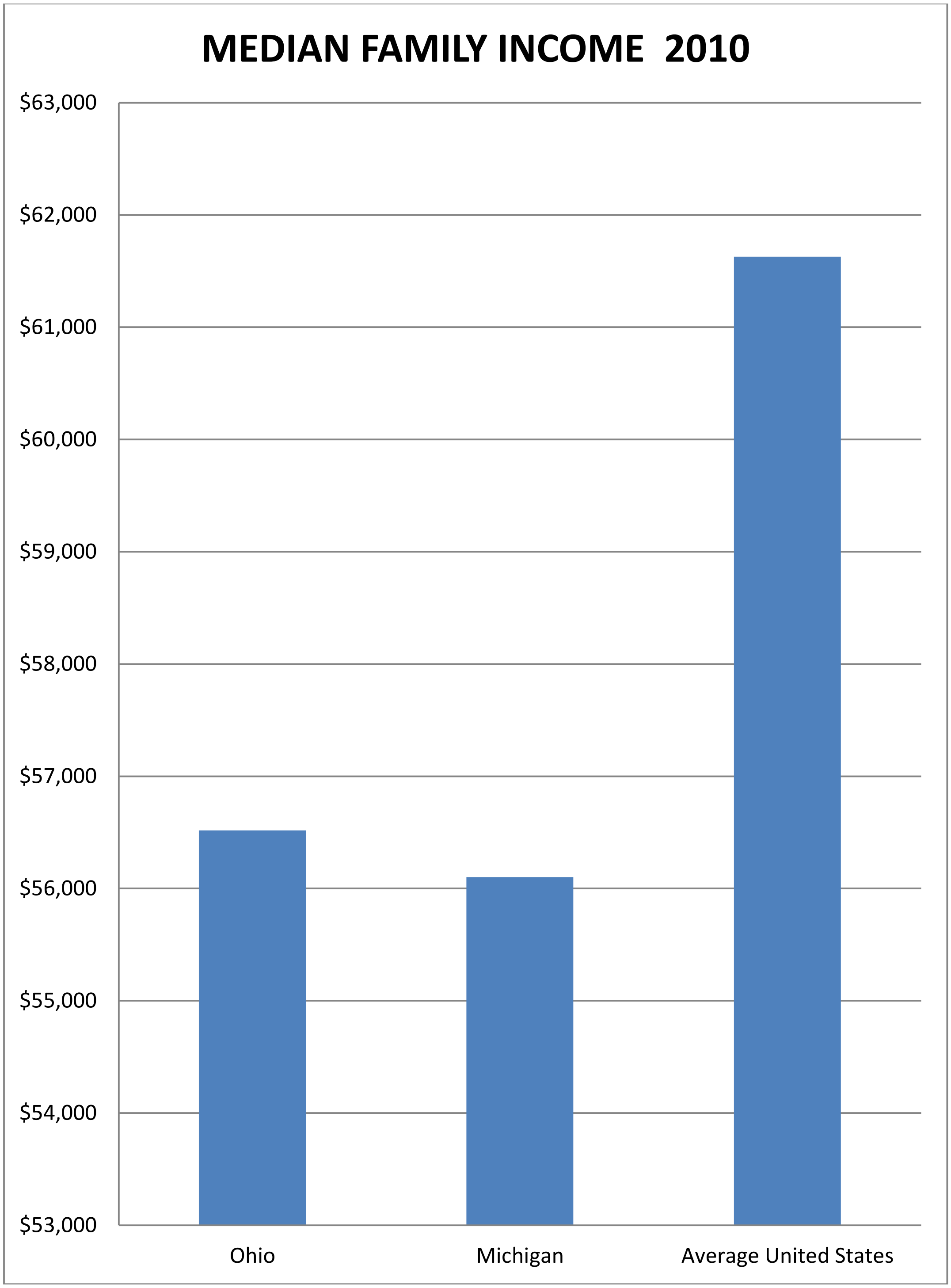
Figure 7. Education Level by State: Pennsylvania and Virginia

APPENDIX C: STATE DATA OHIO AND MICHIGAN



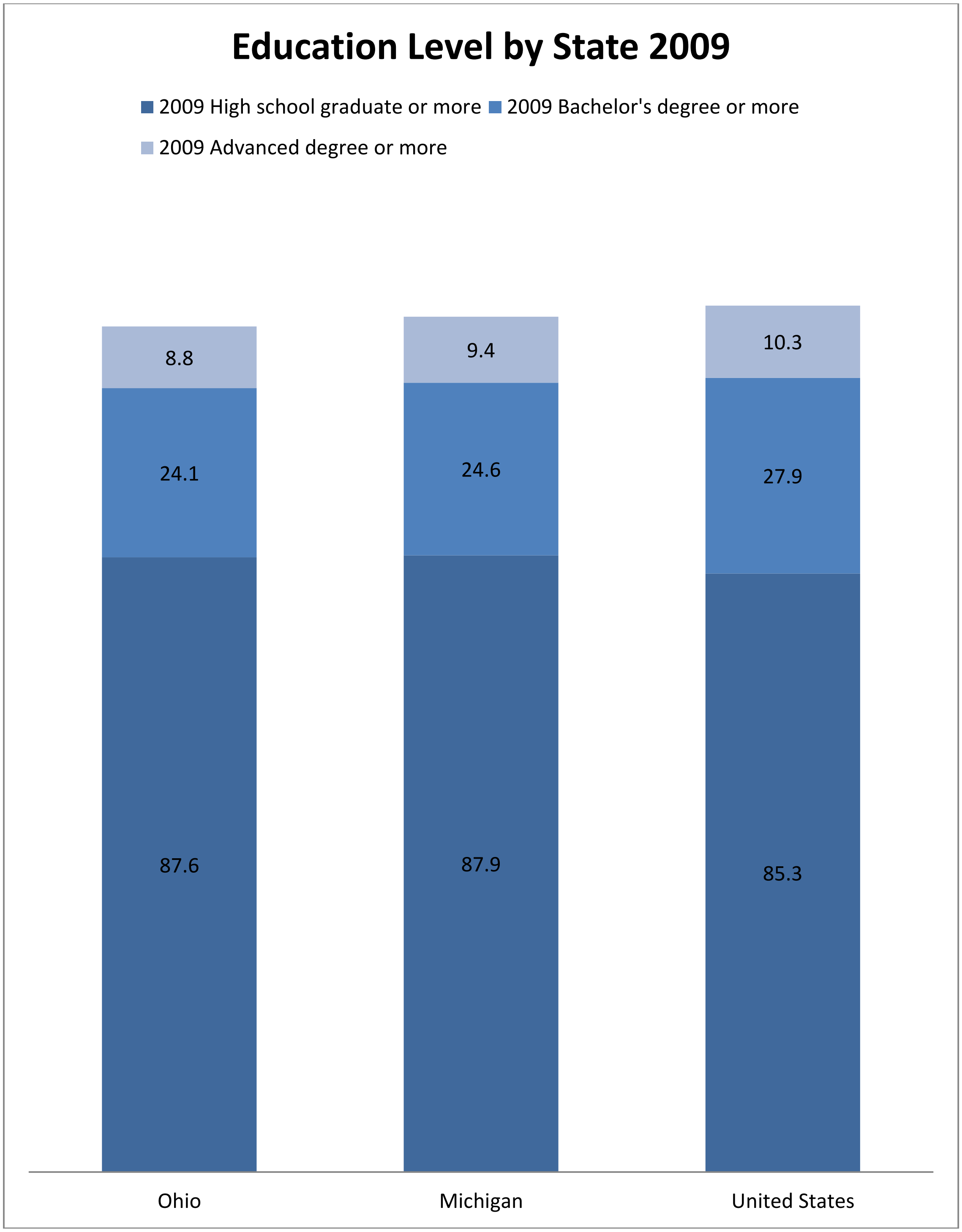
(U.S. Census, 2010)

Figure 8. Demographic Chart: Ohio and Michigan



(U.S. Census, 2010)

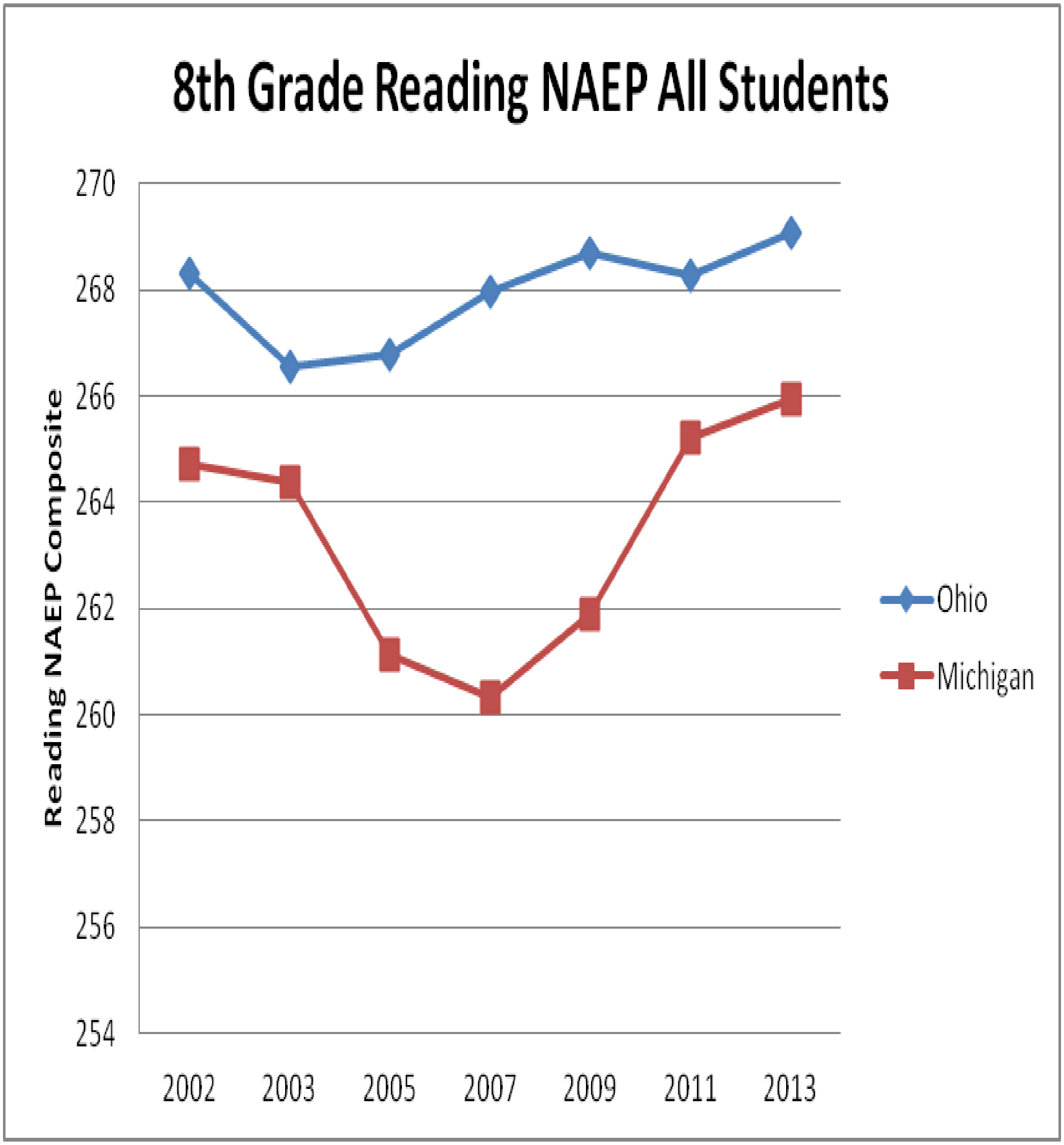
Figure 9. Median Family Income: Ohio and Michigan



(U.S. Census, 2010)

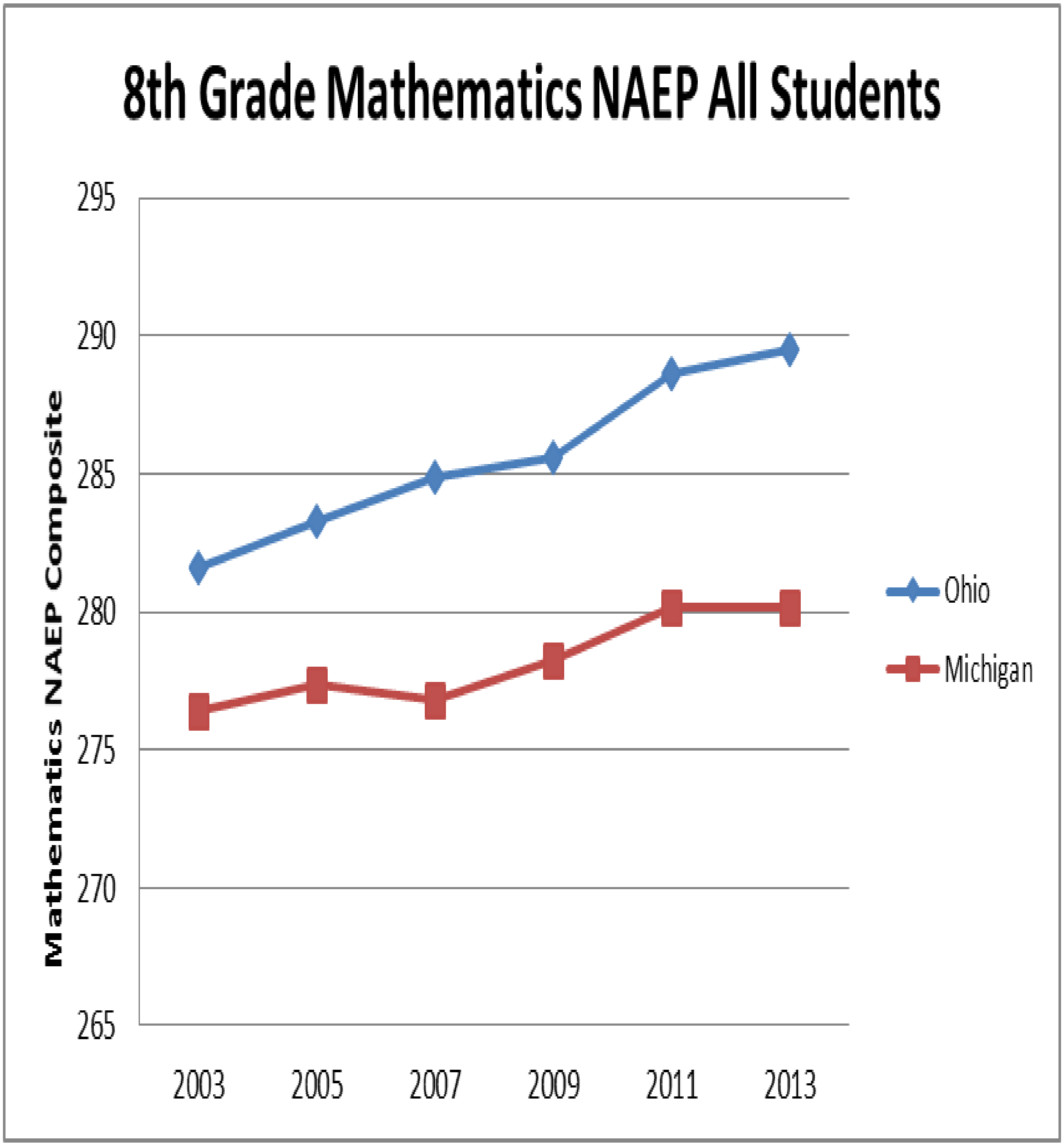
Figure 10. Education Level by State: Ohio and Michigan

APPENDIX D: RESEARCH QUESTION #2



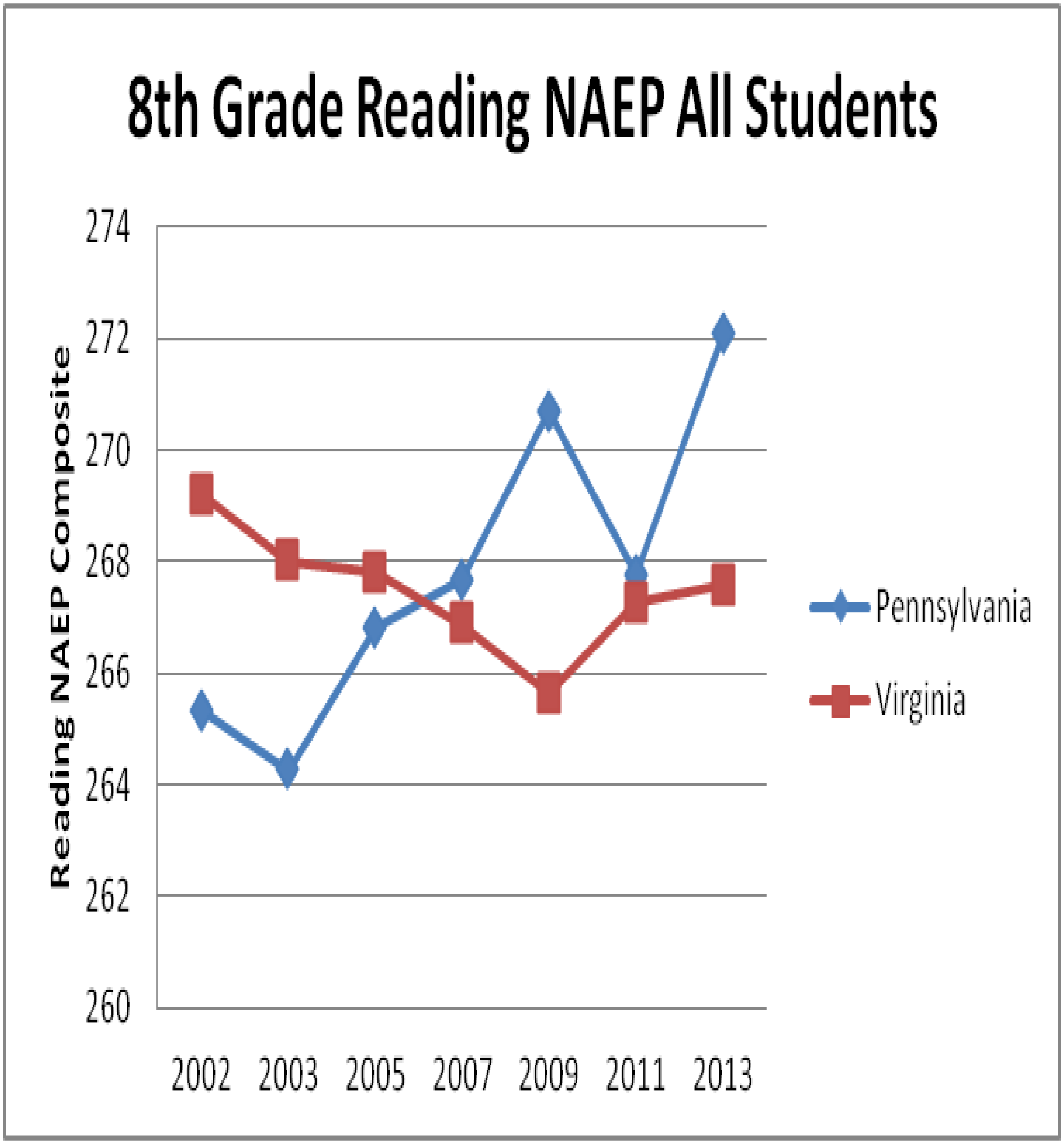
(NAEP, 2014)

Figure 11. *Reading NAEP composite All tested students in Ohio and Michigan 2002-2013*



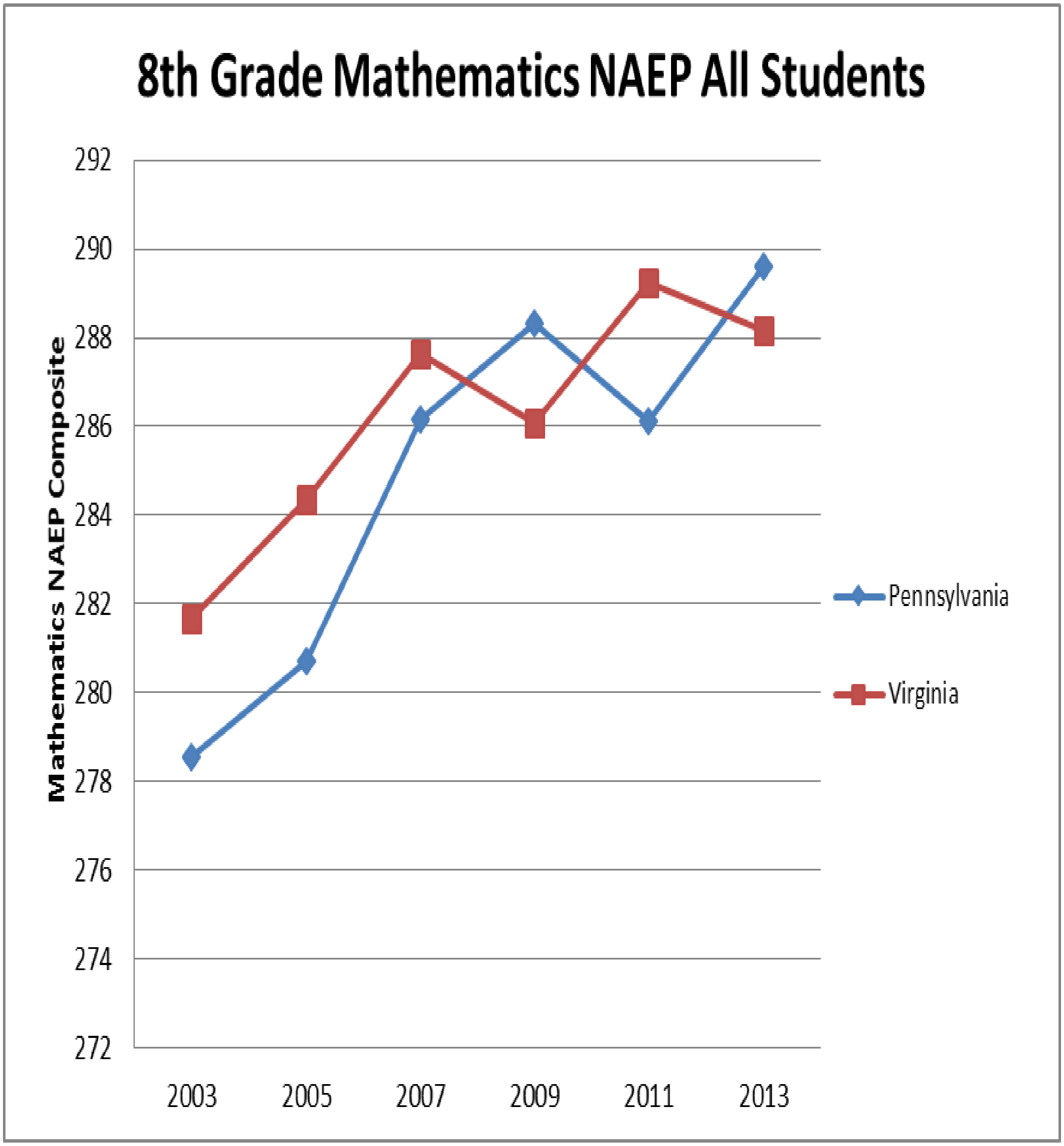
(NAEP, 2014)

Figure 12. Mathematics NAEP composite All tested students in Ohio and Michigan 2003-2013



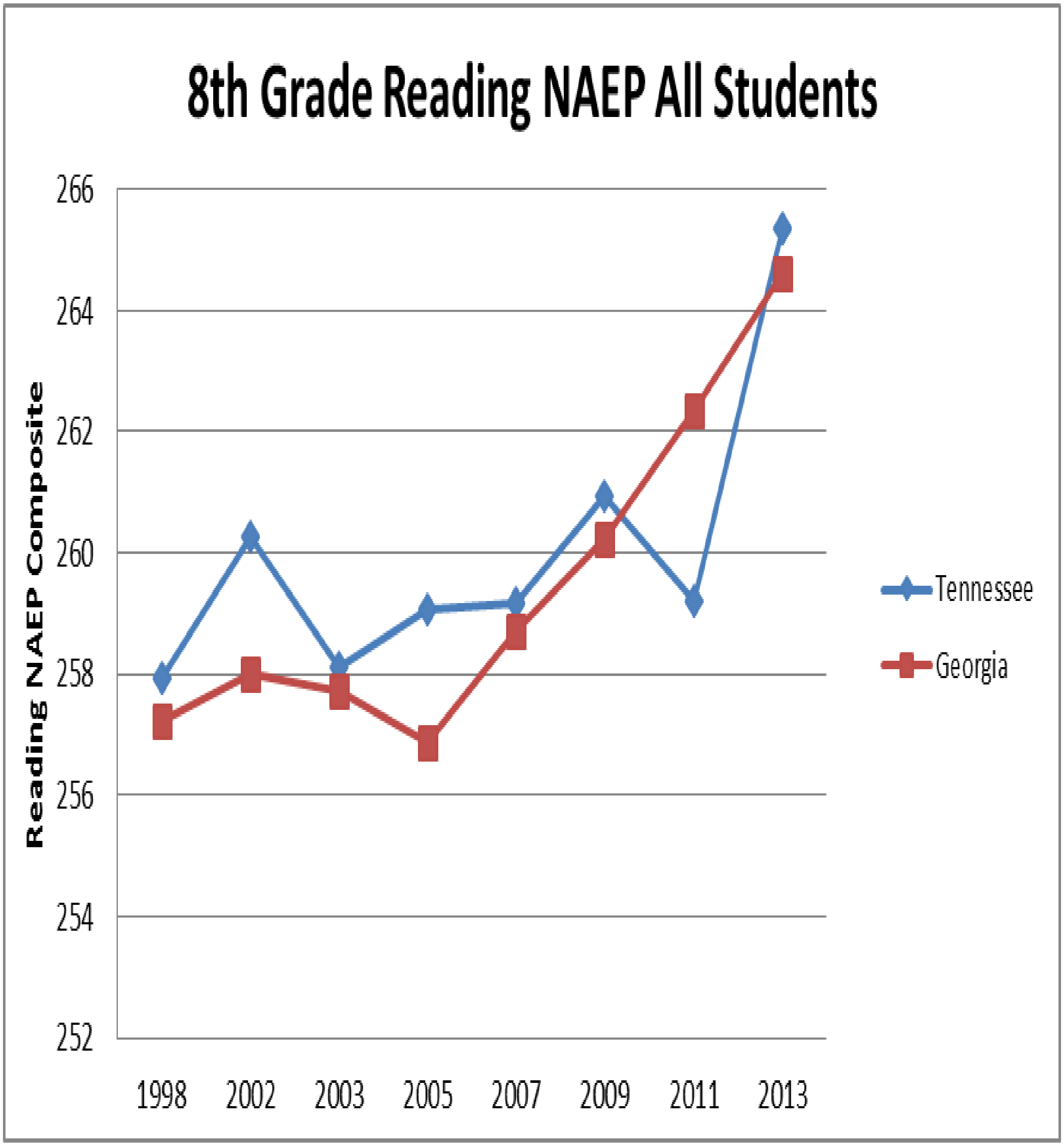
(NAEP, 2014)

Figure 13. *Reading NAEP composite All tested students in Pennsylvania and Virginia 2002-2013*



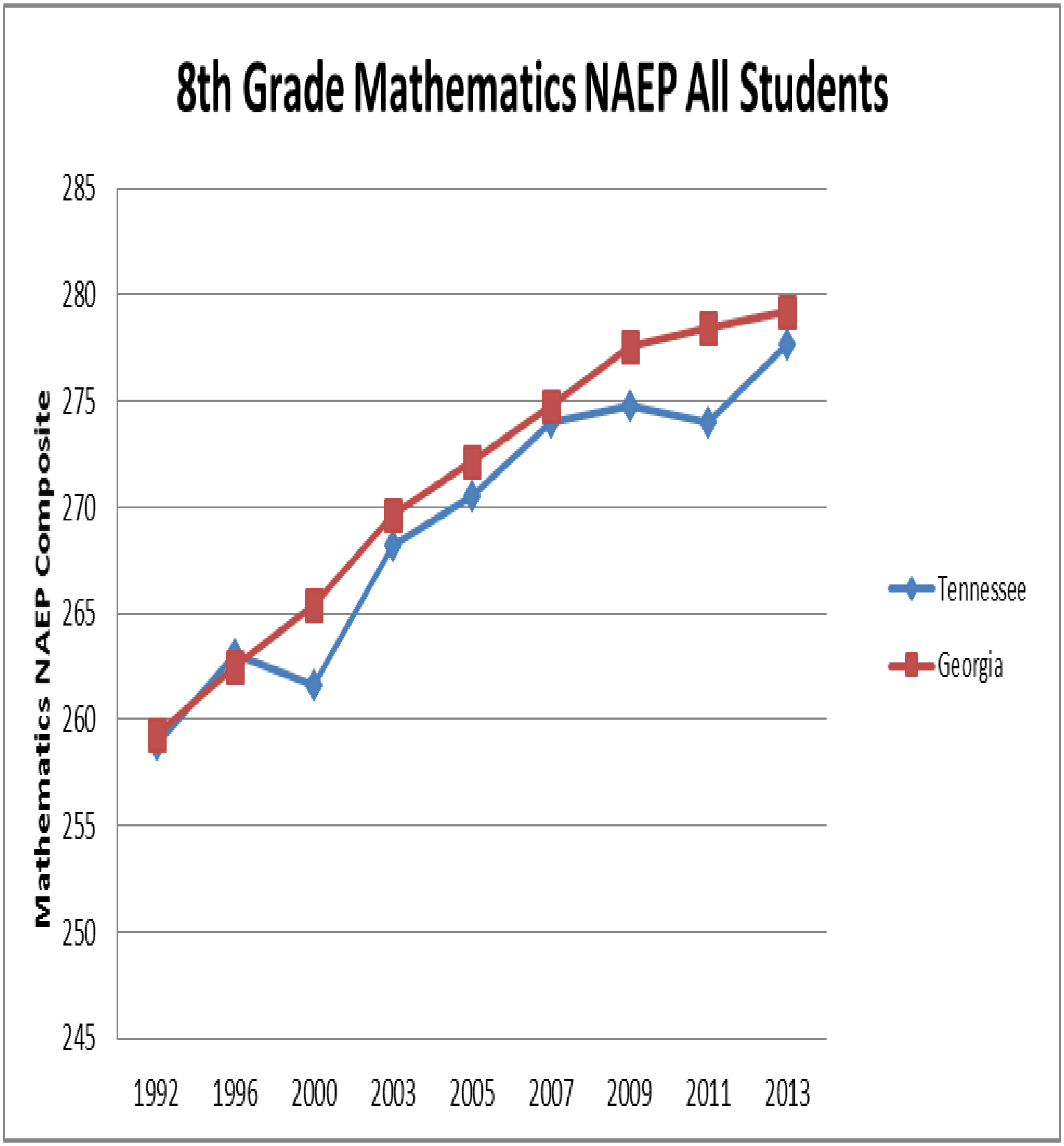
(NAEP, 2014)

Figure 14 mathematics NAEP composite All tested students in Pennsylvania and Virginia 2003-2013



(NAEP, 2014)

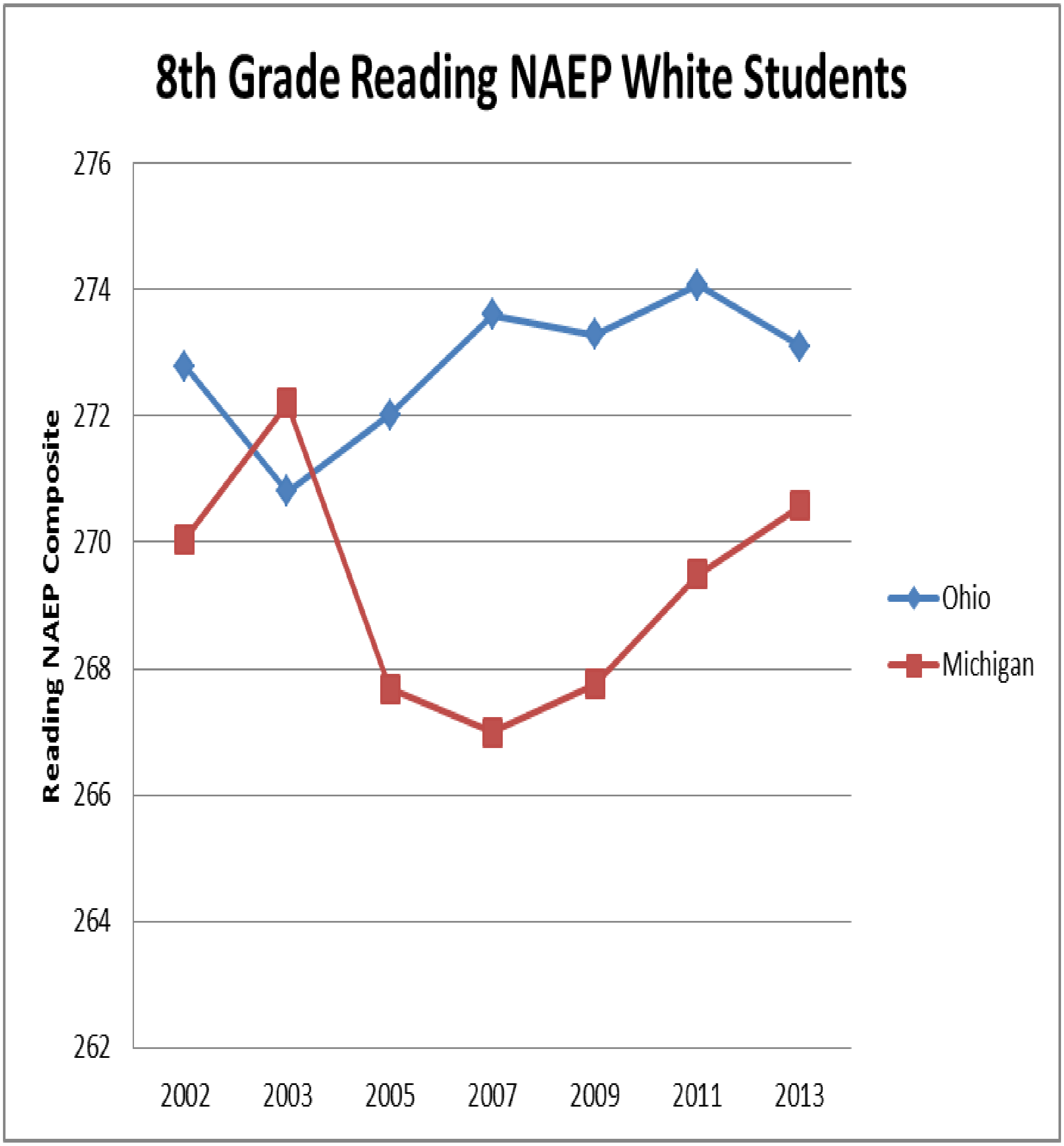
Figure 15. Reading NAEP composite All tested students in Tennessee and Georgia 1998-2013



(NAEP, 2014)

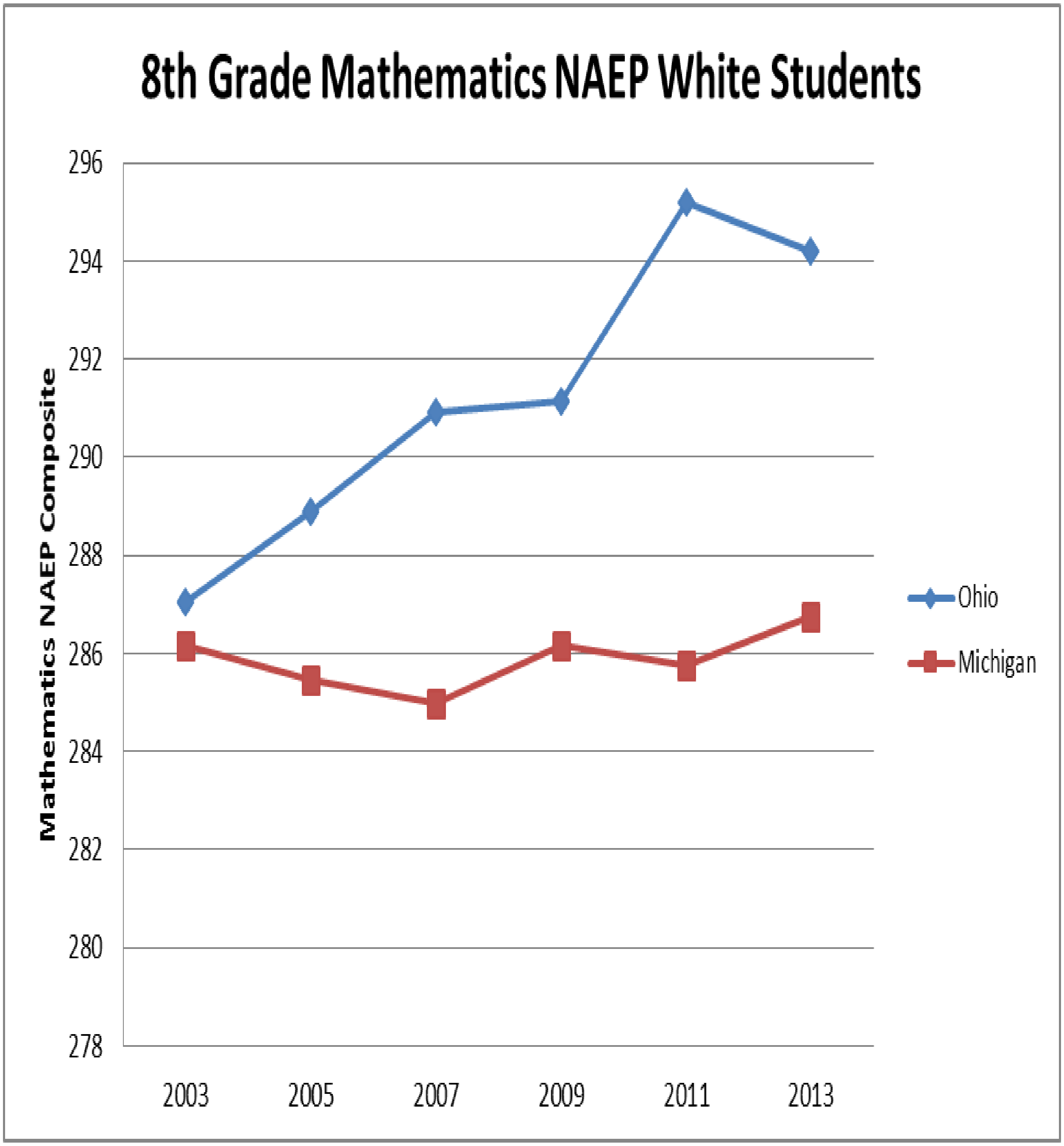
Figure 16. Mathematics NAEP composite All tested students in Tennessee and Georgia 1992-2013

APPENDIX E: RESEARCH QUESTION #3



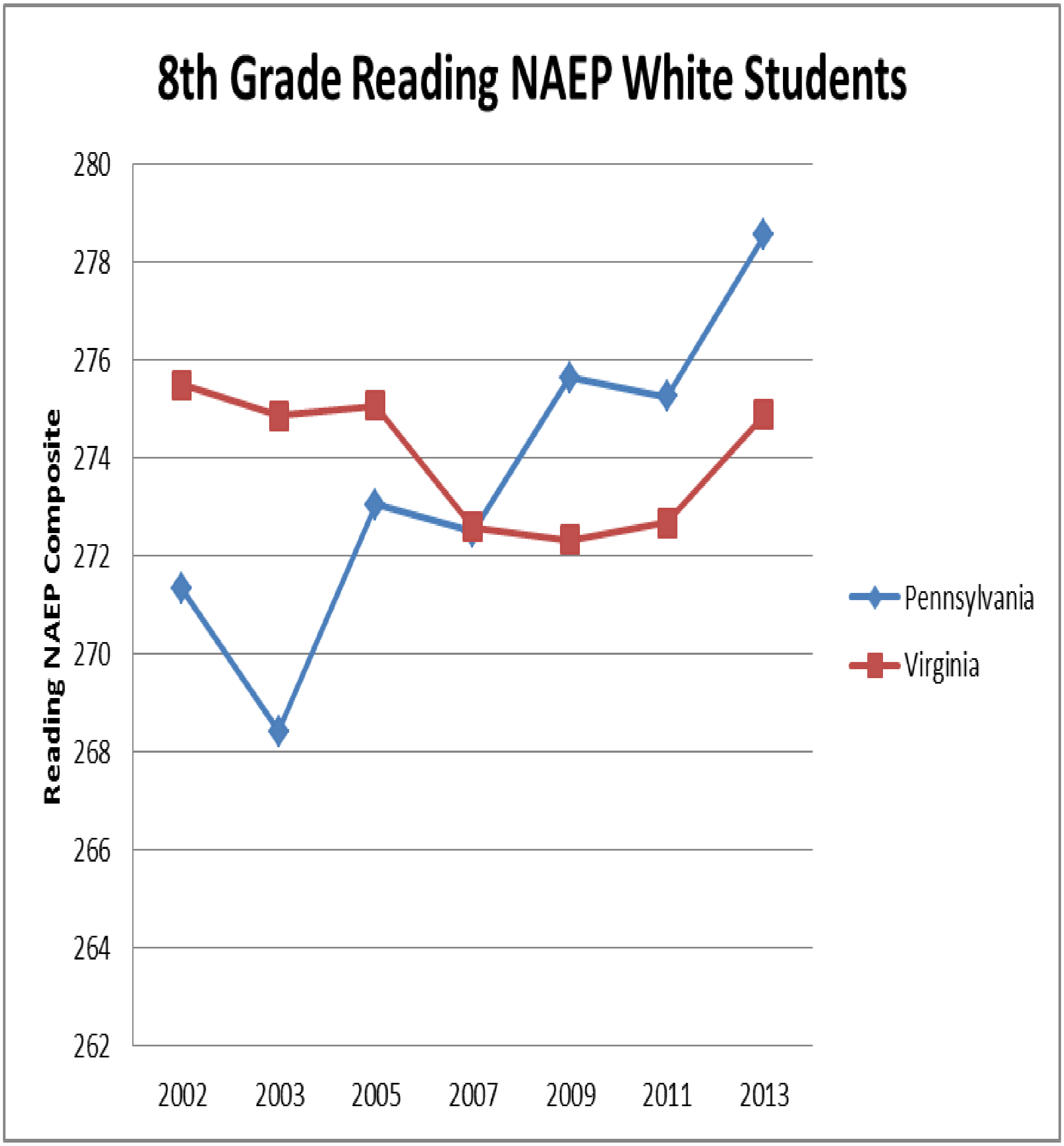
(NAEP, 2014)

Figure 17. *Reading NAEP composite All Tested White Students in Ohio and Michigan 2002-2013*



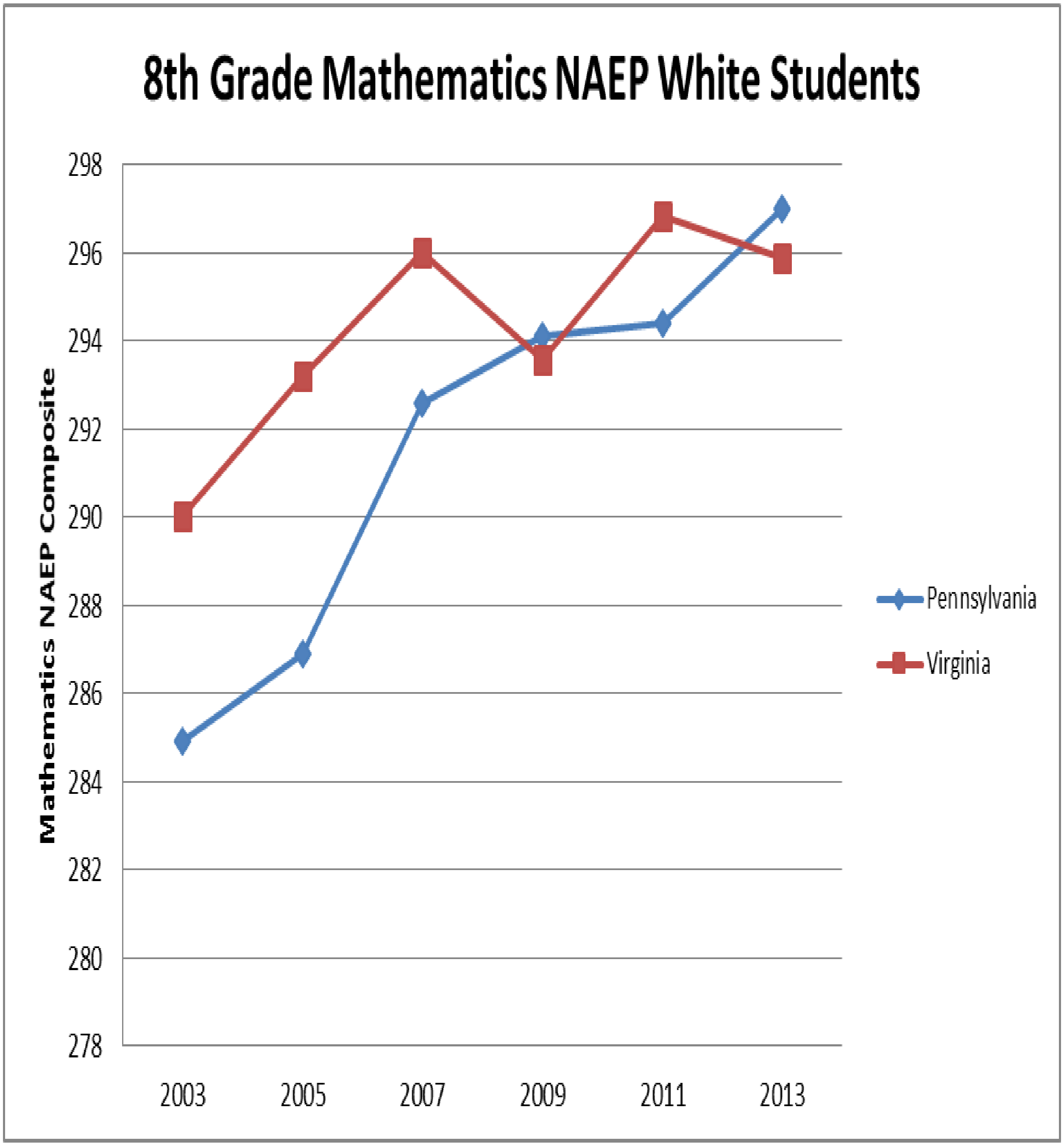
(NAEP, 2014)

Figure 18. *Mathematics NAEP composite all tested White students in Ohio and Michigan 2003-2013*



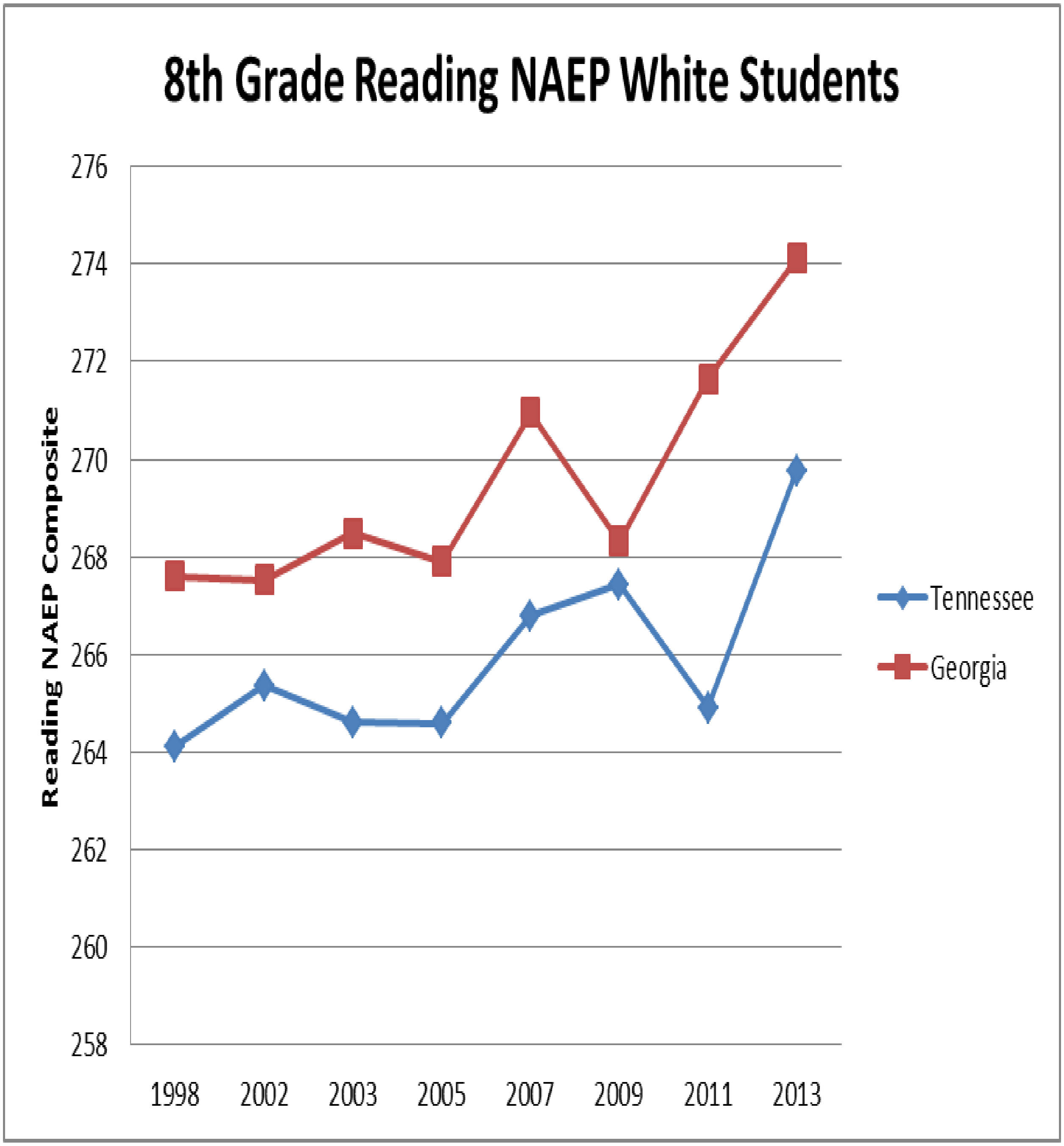
(NAEP, 2014)

Figure 19. *Reading NAEP composite all tested White students in Pennsylvania and Virginia 2002-2013*



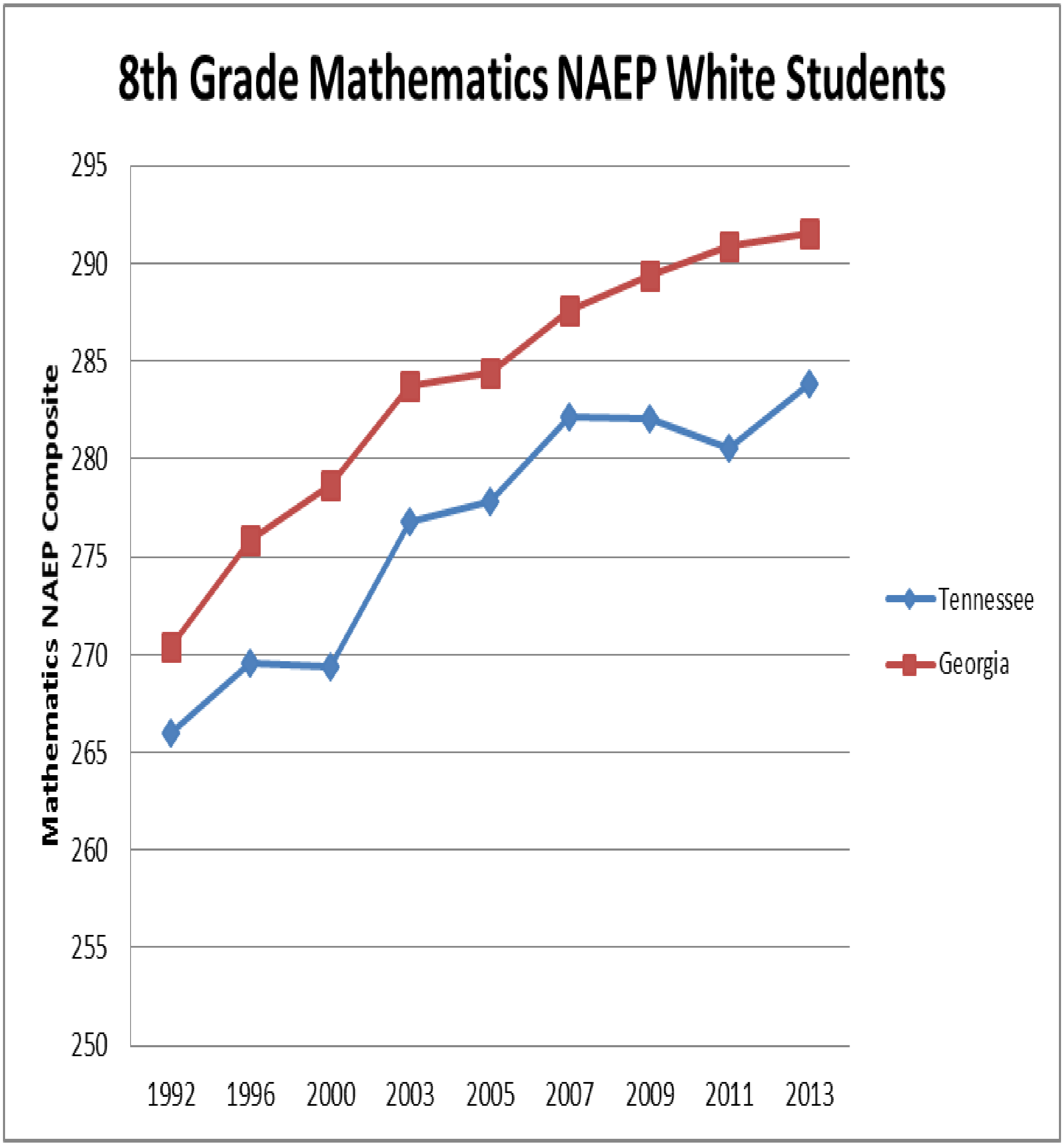
(NAEP, 2014)

Figure 20. Mathematics NAEP composite all tested students in Pennsylvania and Virginia 2002-2013



(NAEP, 2014)

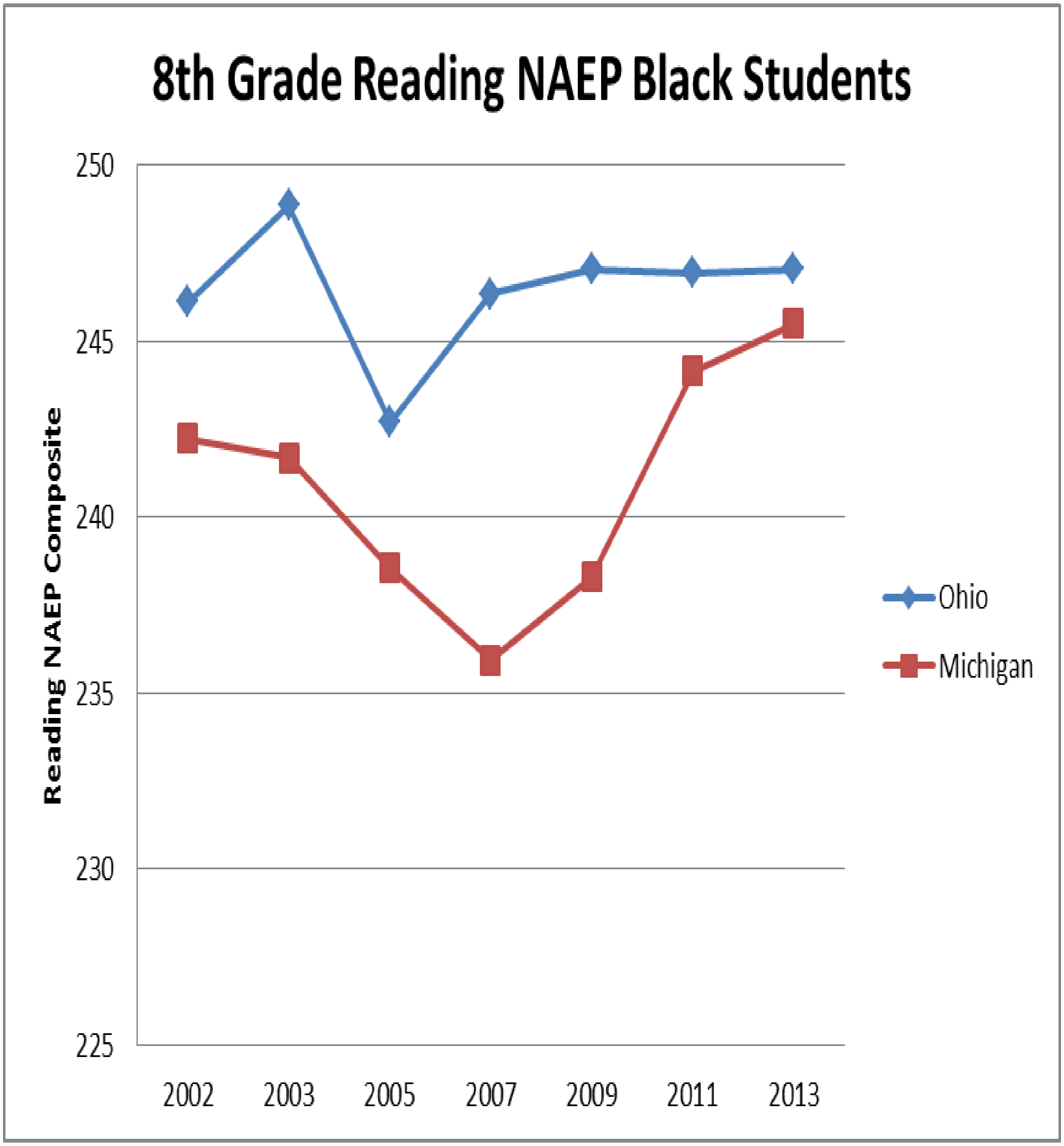
Figure 21. *Reading NAEP composite all tested White students in Tennessee and Georgia 1998-2013*



(NAEP, 2014)

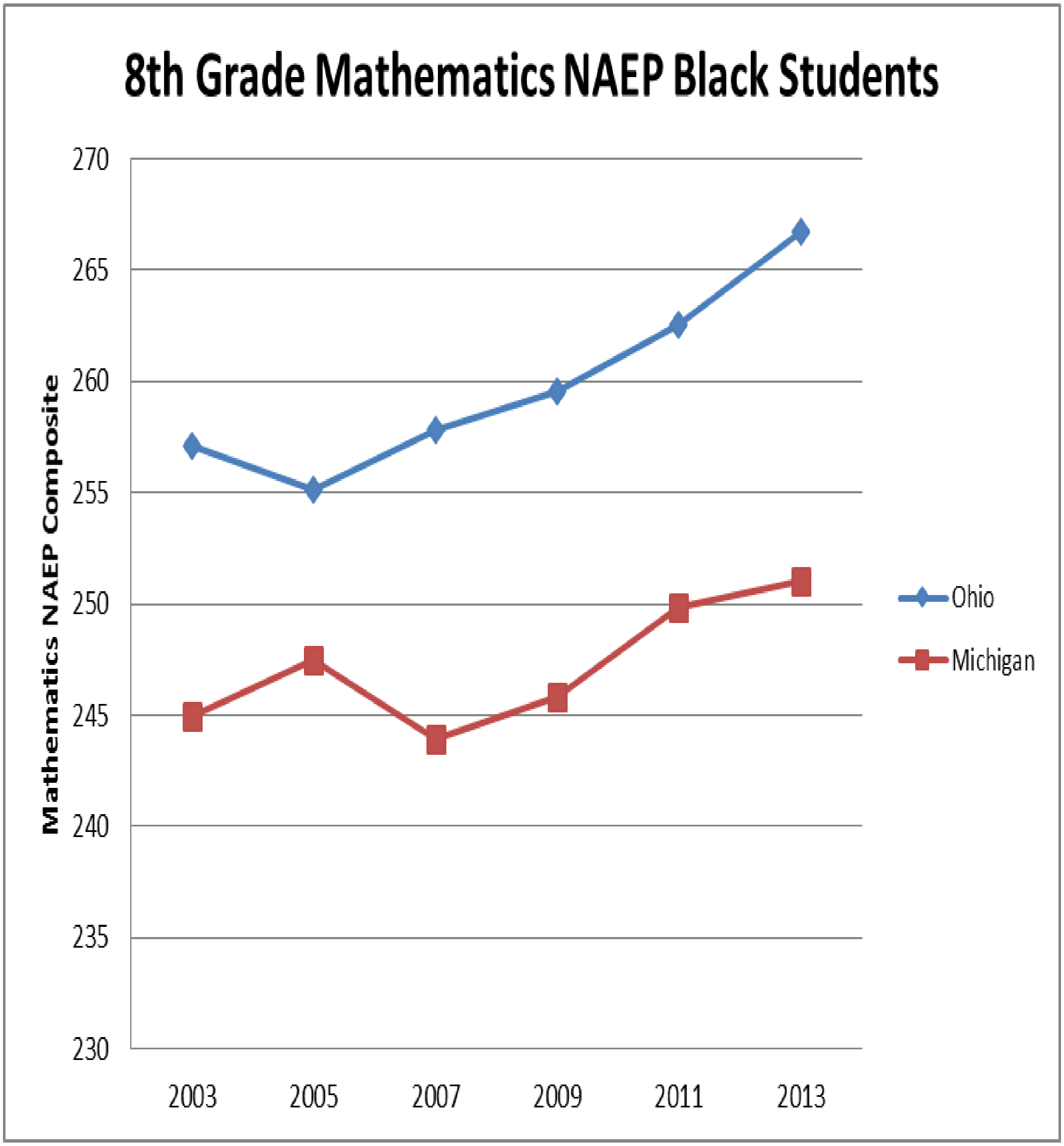
Figure 22. Mathematics NAEP composite all tested White students in Tennessee and Georgia 1998-2013

APPENDIX F: RESEARCH QUESTION #4



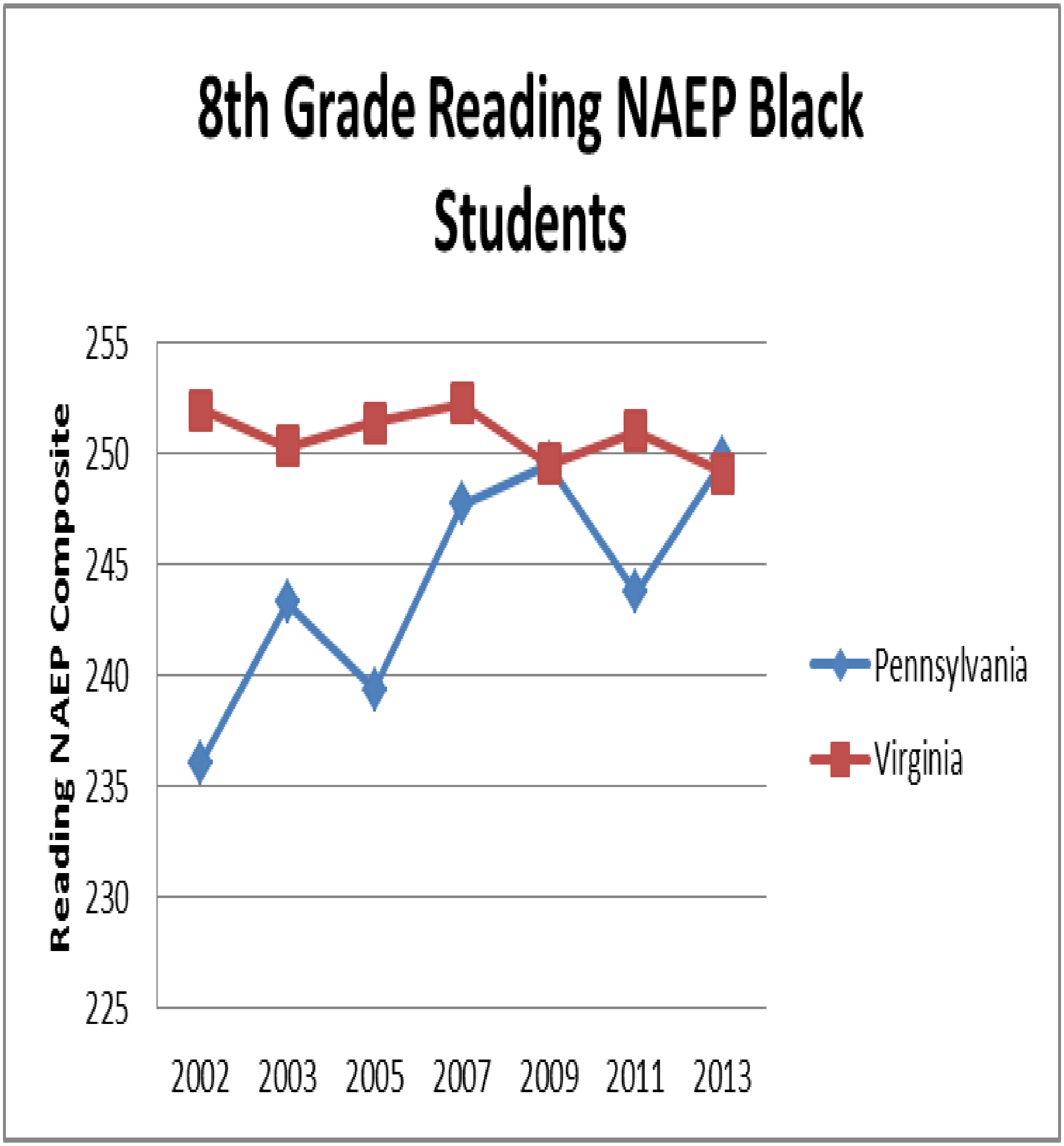
(NAEP, 2014)

Figure 23. Reading NAEP composite all tested Black students in Ohio and Michigan 2002-2013



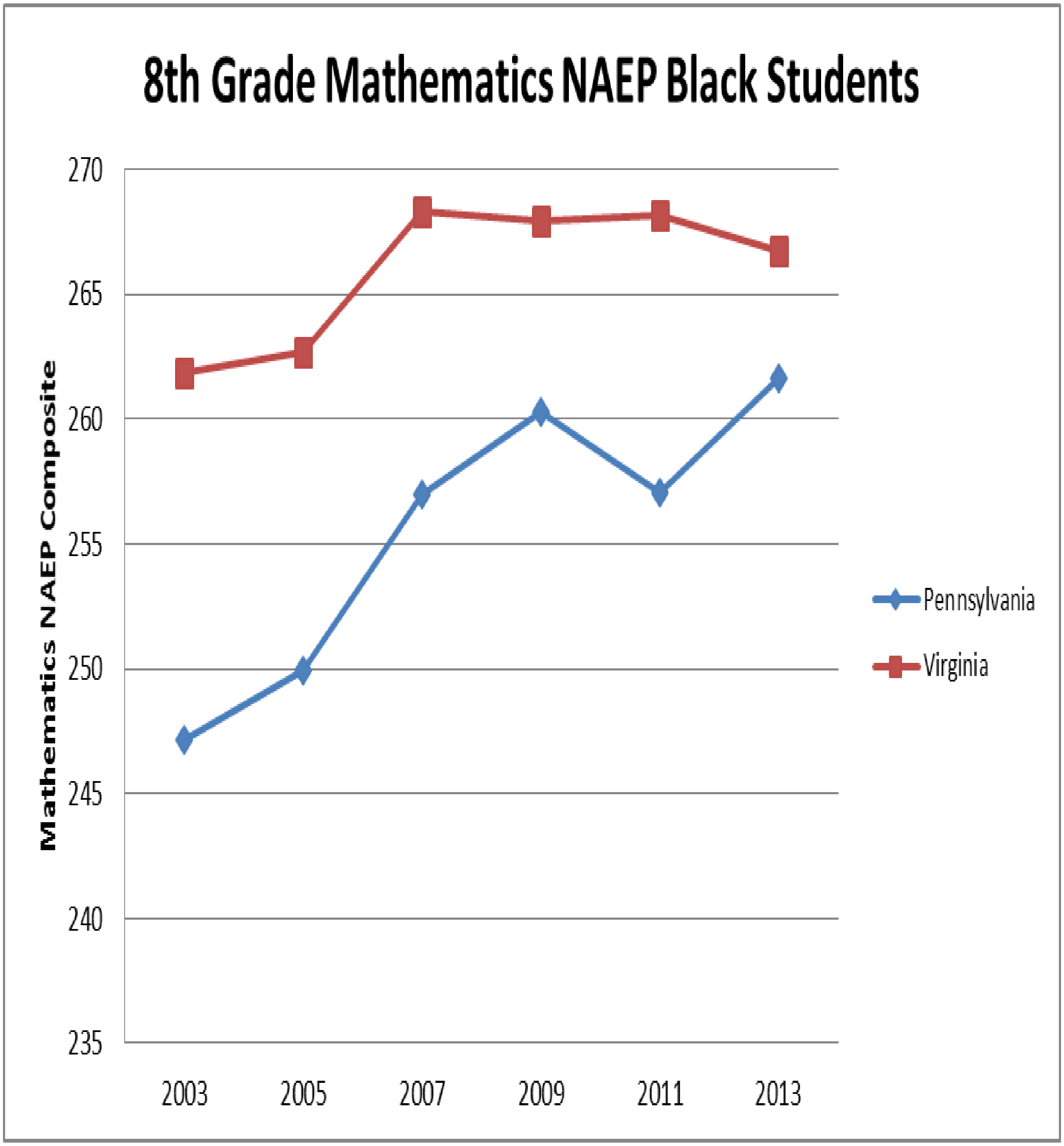
(NAEP, 2014)

Figure 24. Mathematics NAEP composite all tested Black students in Ohio and Michigan 2003-2013



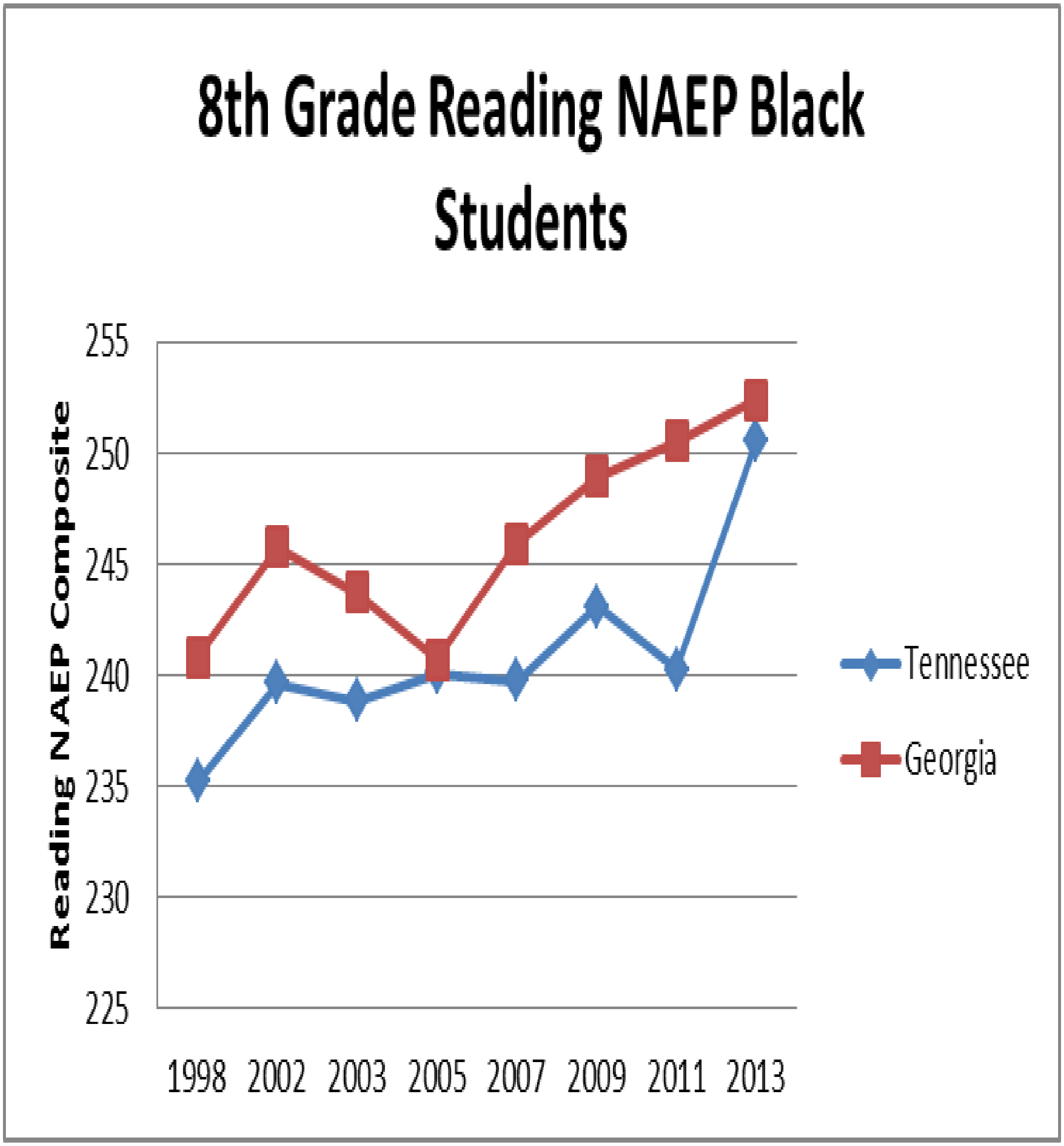
(NAEP, 2014)

Figure 25. *Reading NAEP composite all tested Black students in Pennsylvania and Virginia 2002-2013*



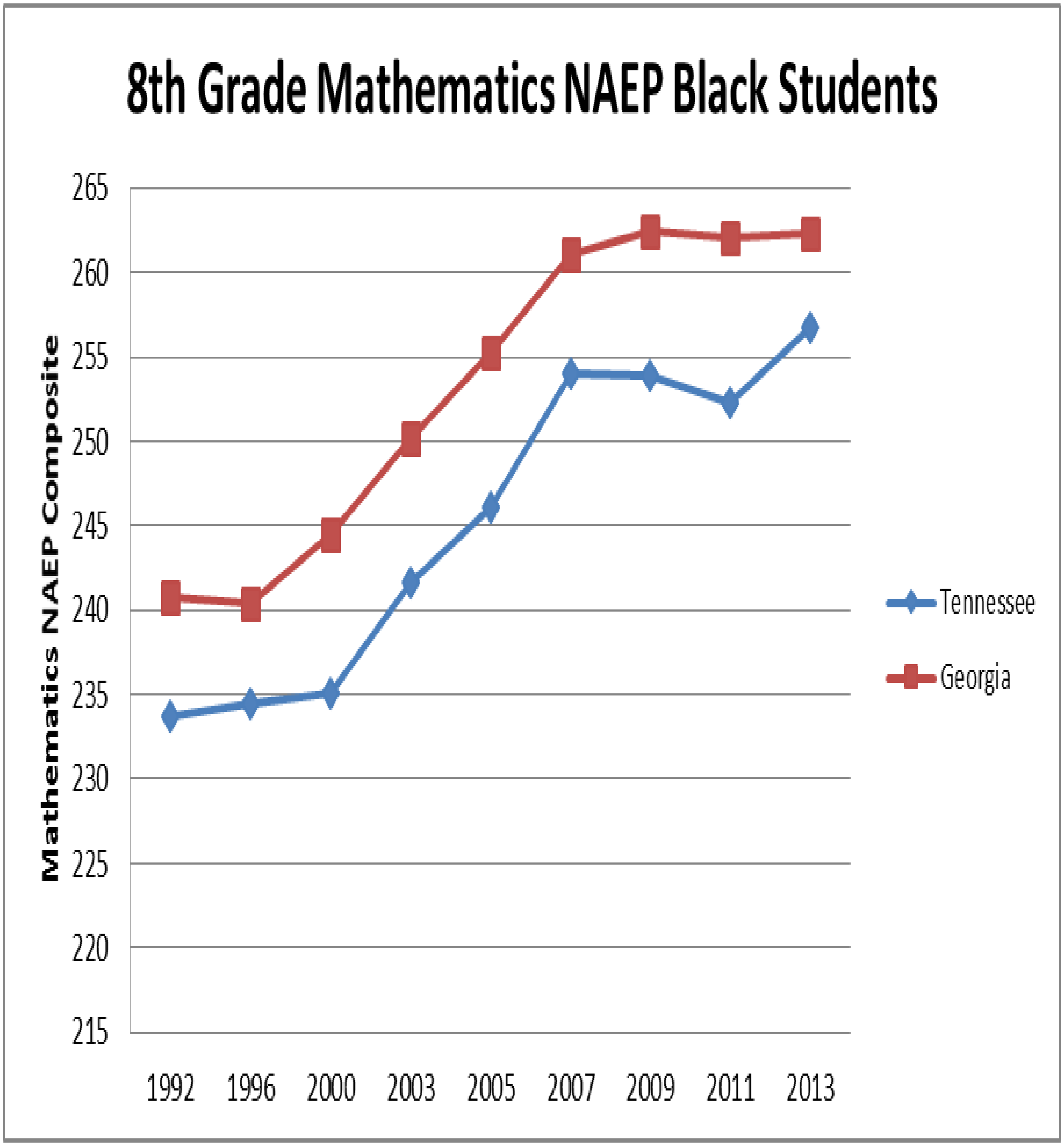
(NAEP, 2014)

Figure 26. Mathematics NAEP composite all tested Black students in Pennsylvania and Virginia 2003-2013



(NAEP, 2014)

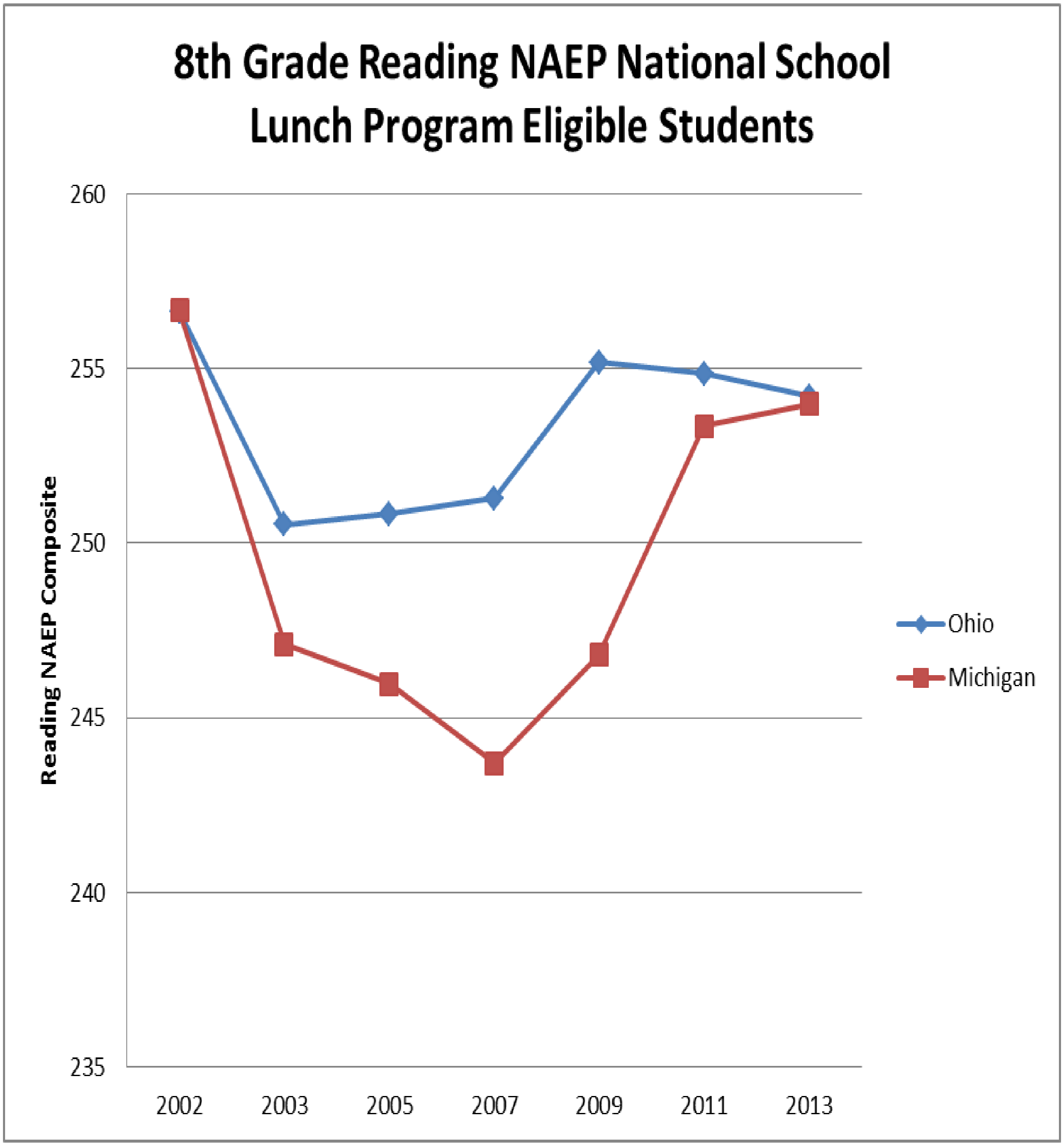
Figure 27. *Reading NAEP composite all tested Black students in Tennessee and Georgia 1998-2013*



(NAEP, 2014)

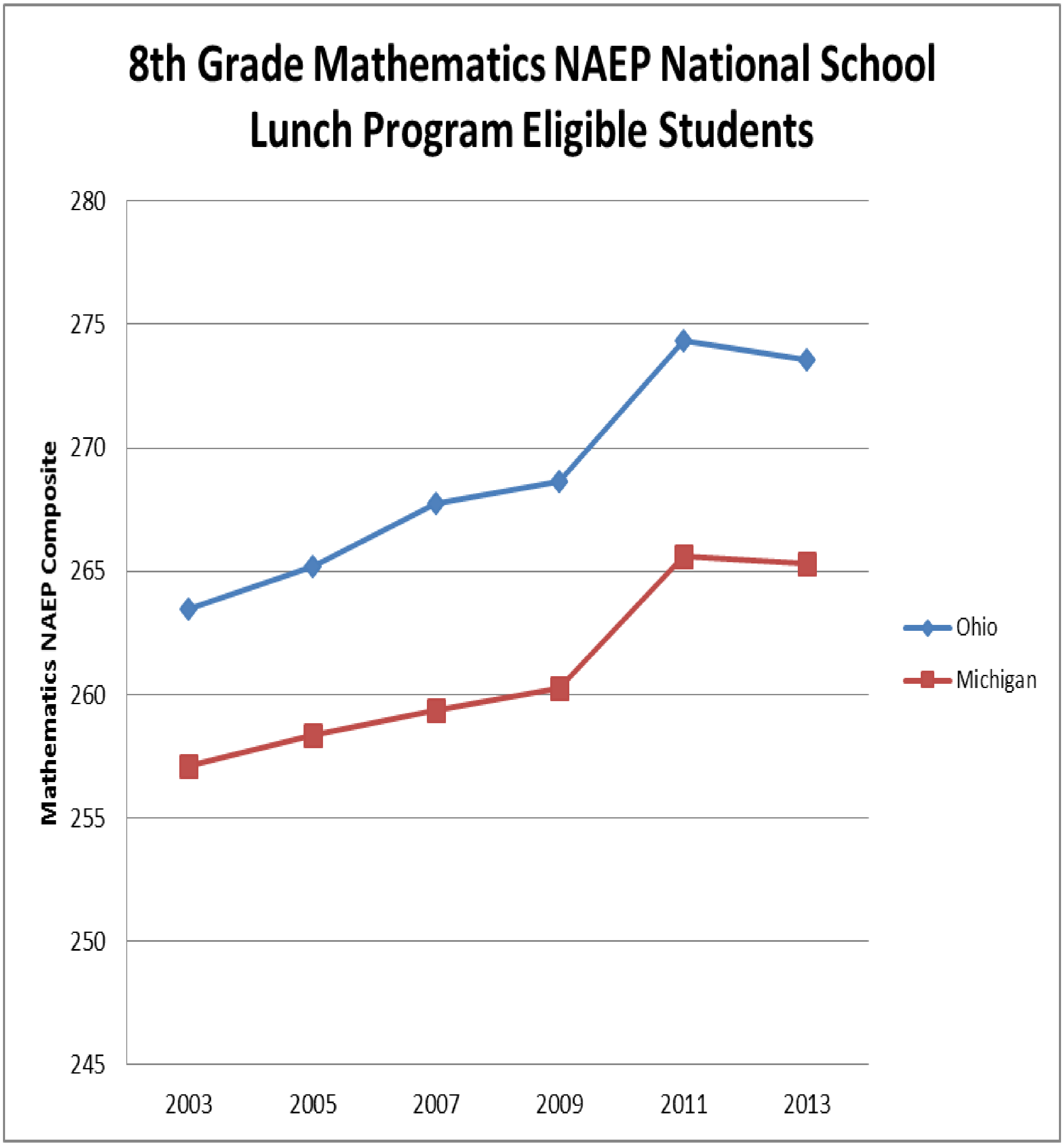
Figure 28. *Mathematics NAEP composite all tested Black students in Tennessee and Georgia 1992-2013*

APPENDIX G: RESEARCH QUESTION #5



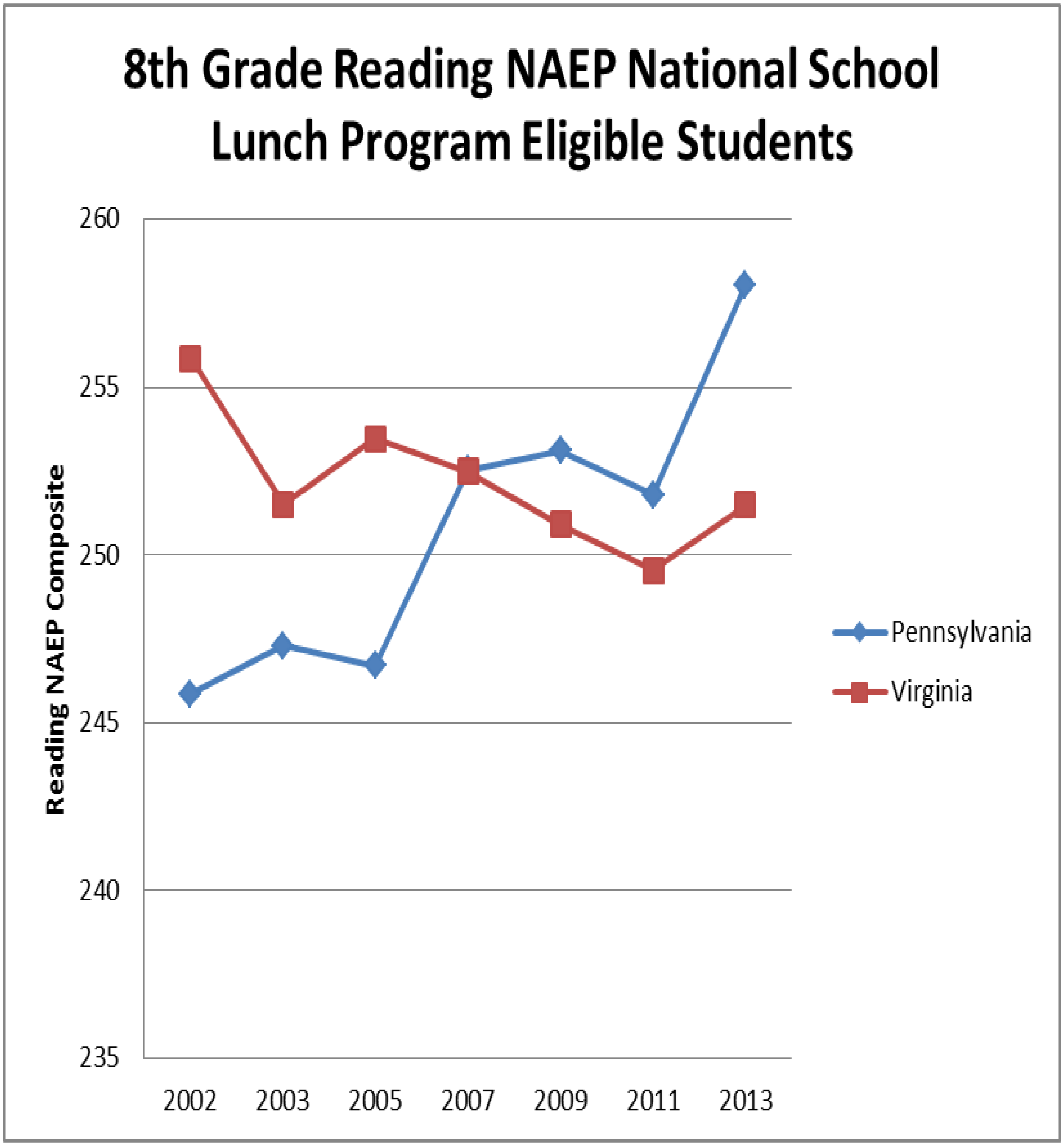
(NAEP, 2014)

Figure 29. Reading NAEP composite all tested National School Lunch Program Eligible students in Ohio and Michigan 2002-2013



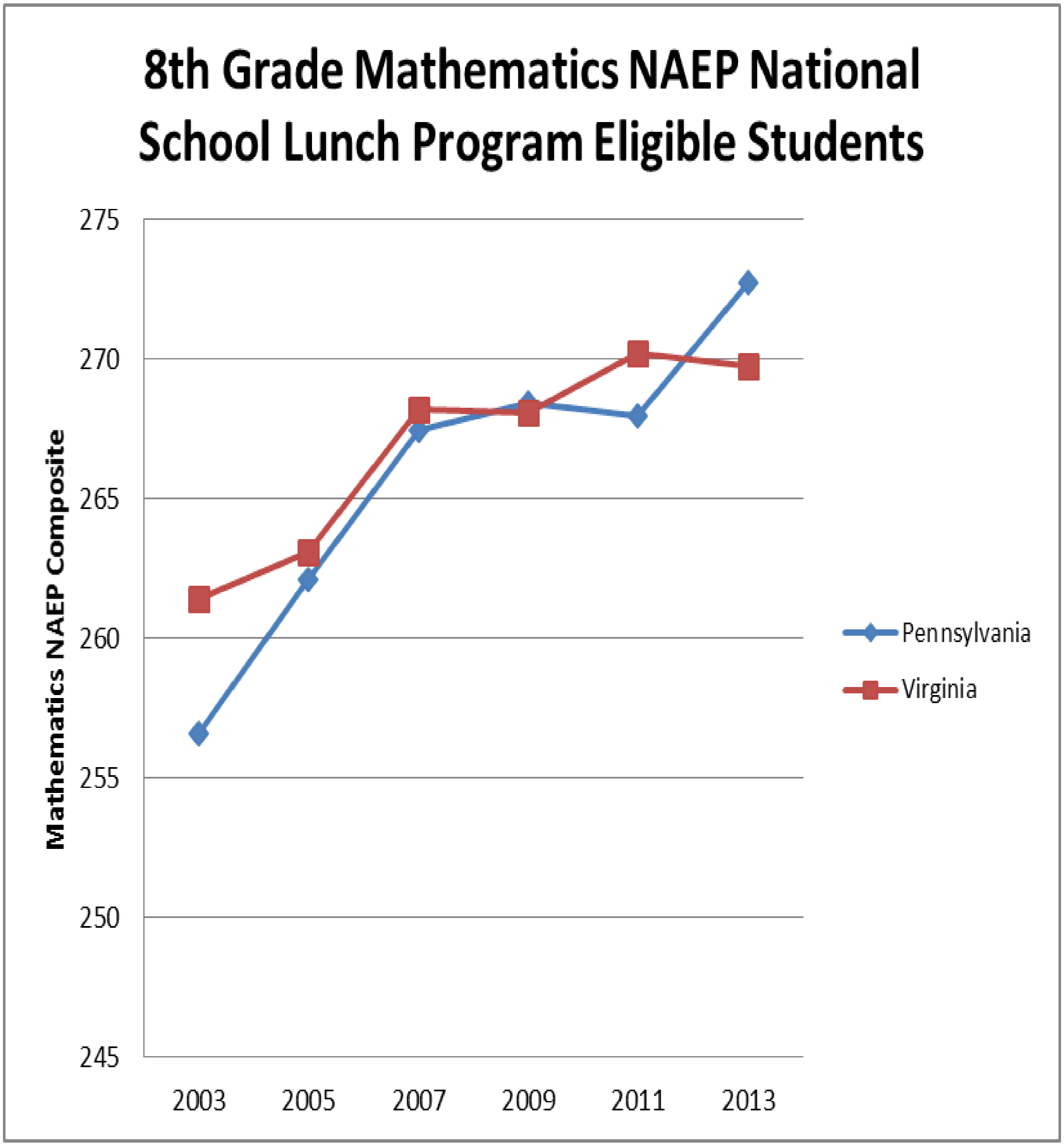
(NAEP, 2014)

Figure 30. Mathematics NAEP composite all tested National School Lunch Program Eligible students in Ohio and Michigan 2003-2013



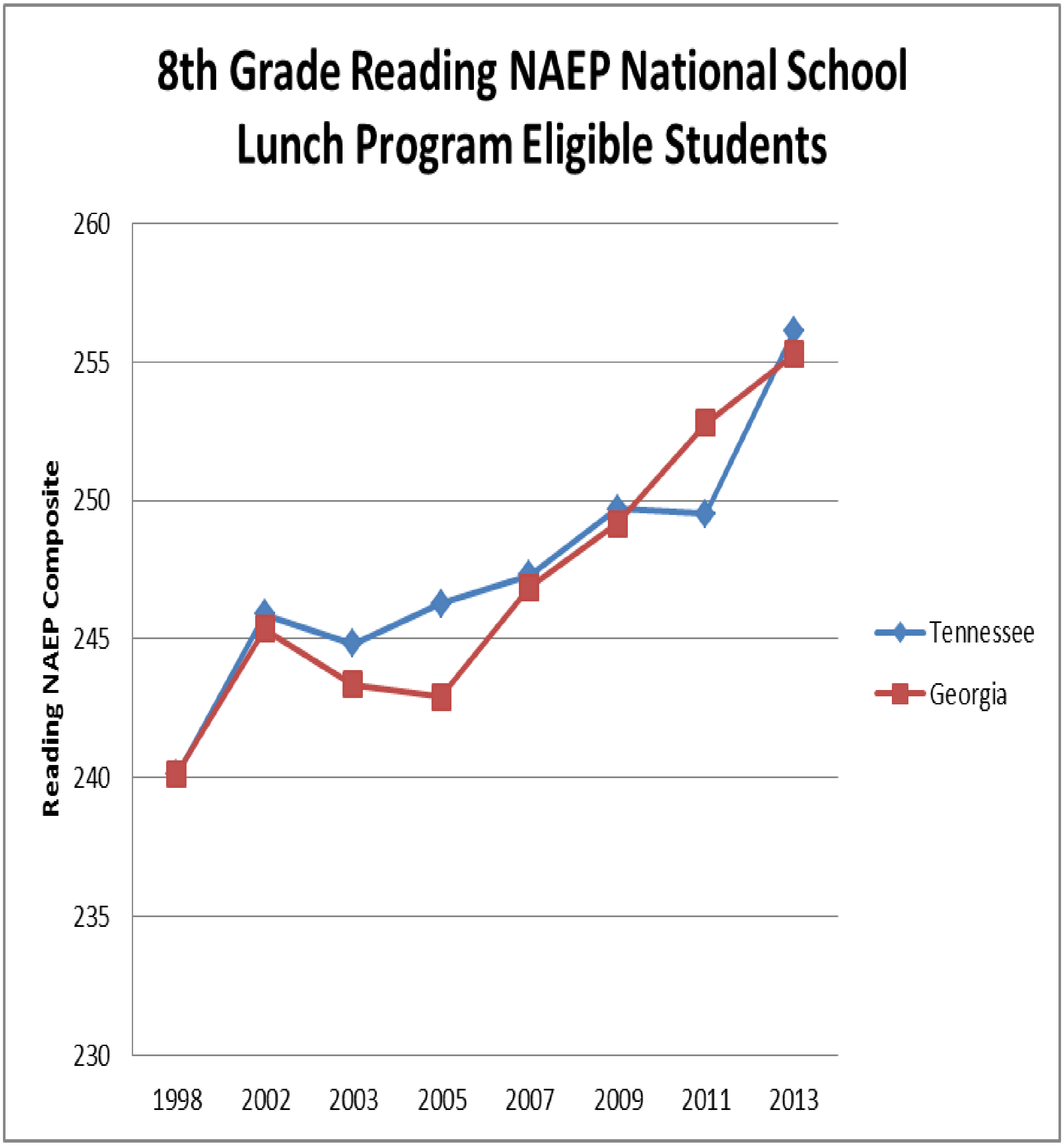
(NAEP, 2014)

Figure 31. *Reading NAEP composite all tested National School Lunch Program Eligible tested in Pennsylvania and Virginia 2002-2013*



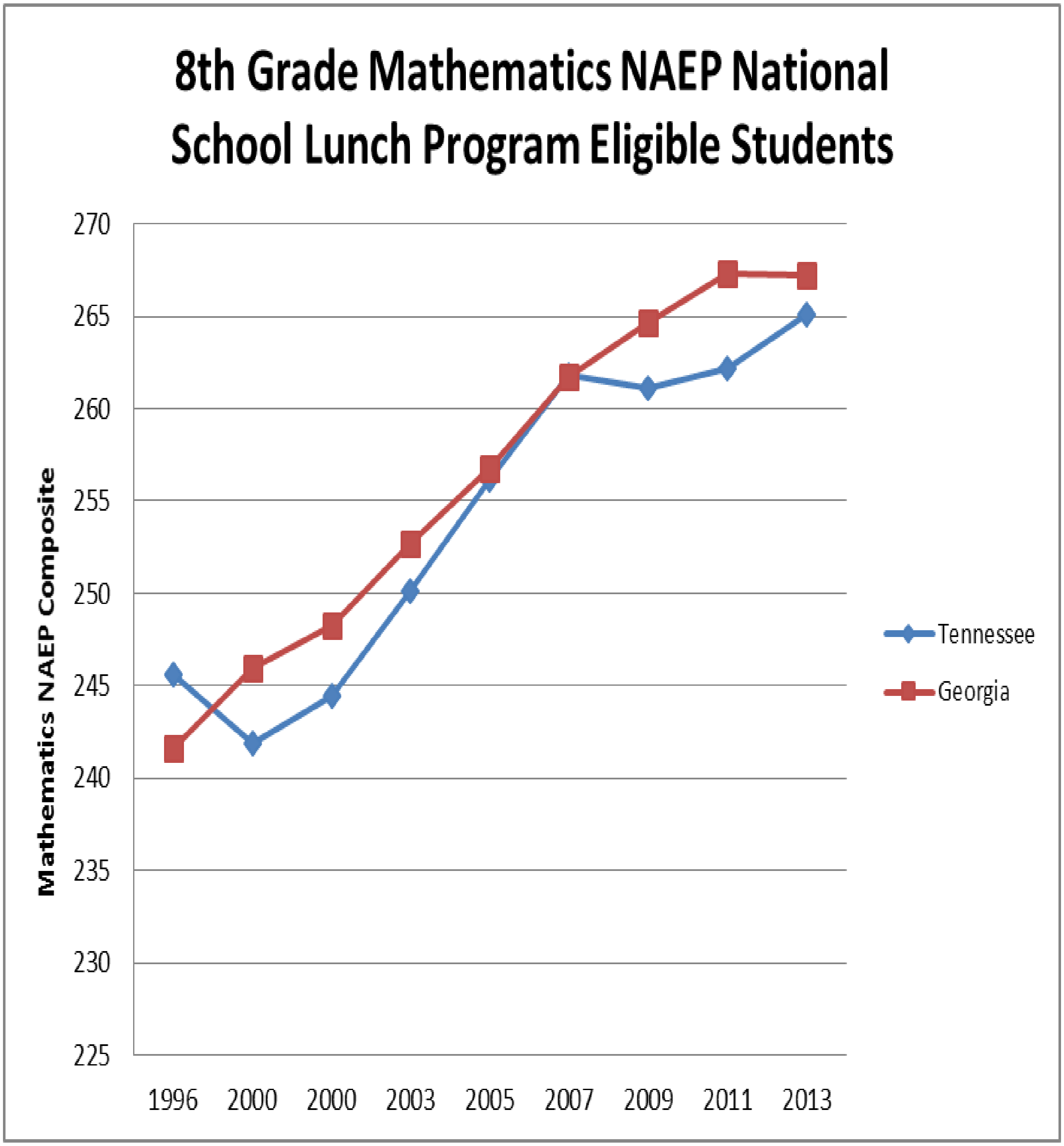
(NAEP, 2014)

Figure 32. Mathematics NAEP composite all tested National School Lunch Program Eligible students in Pennsylvania and Virginia 2002-2013



(NAEP, 2014)

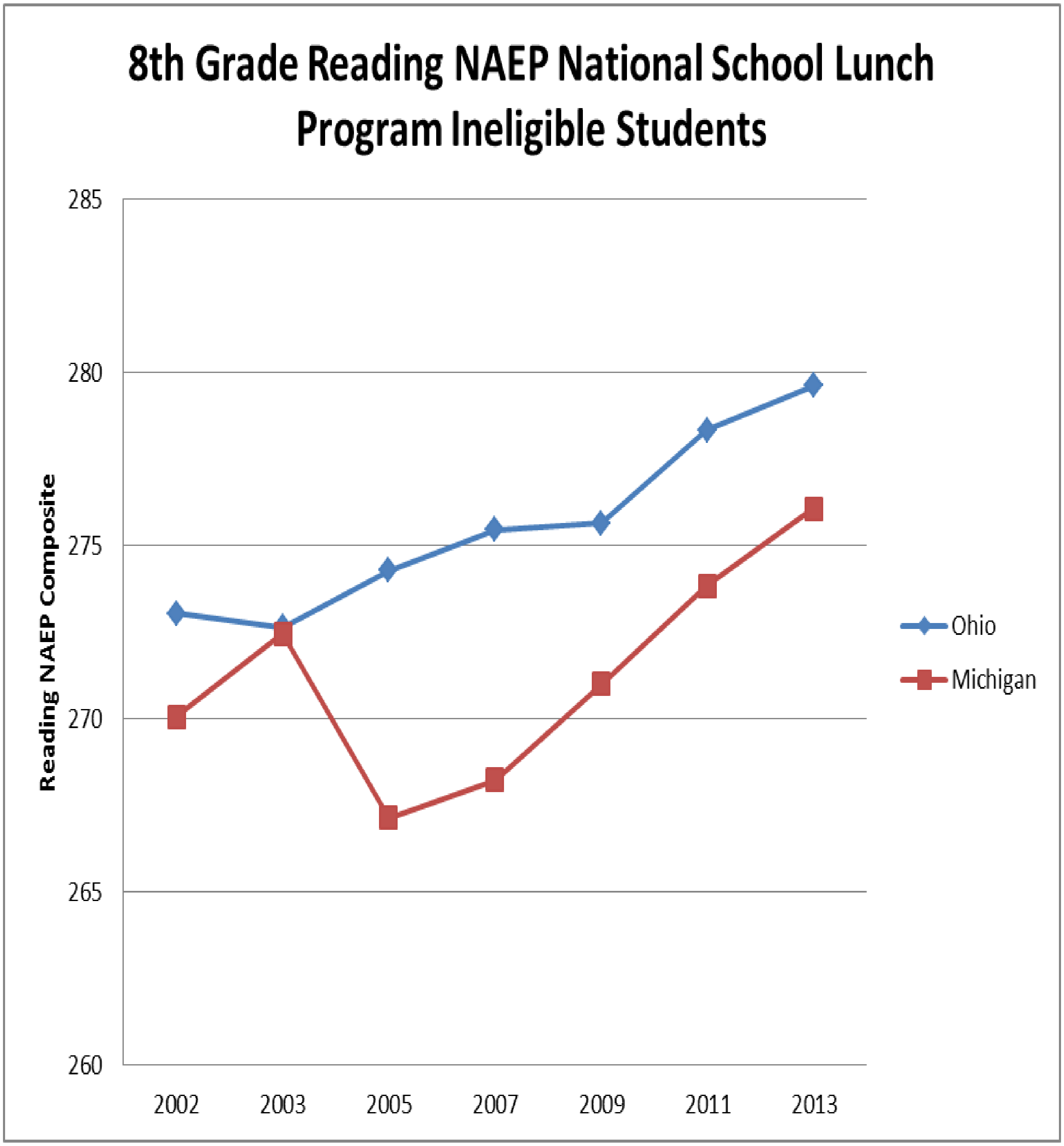
Figure 33. *Reading NAEP composite all tested National School Lunch Program Eligible students in Tennessee and Georgia 1998-2013*



(NAEP, 2014)

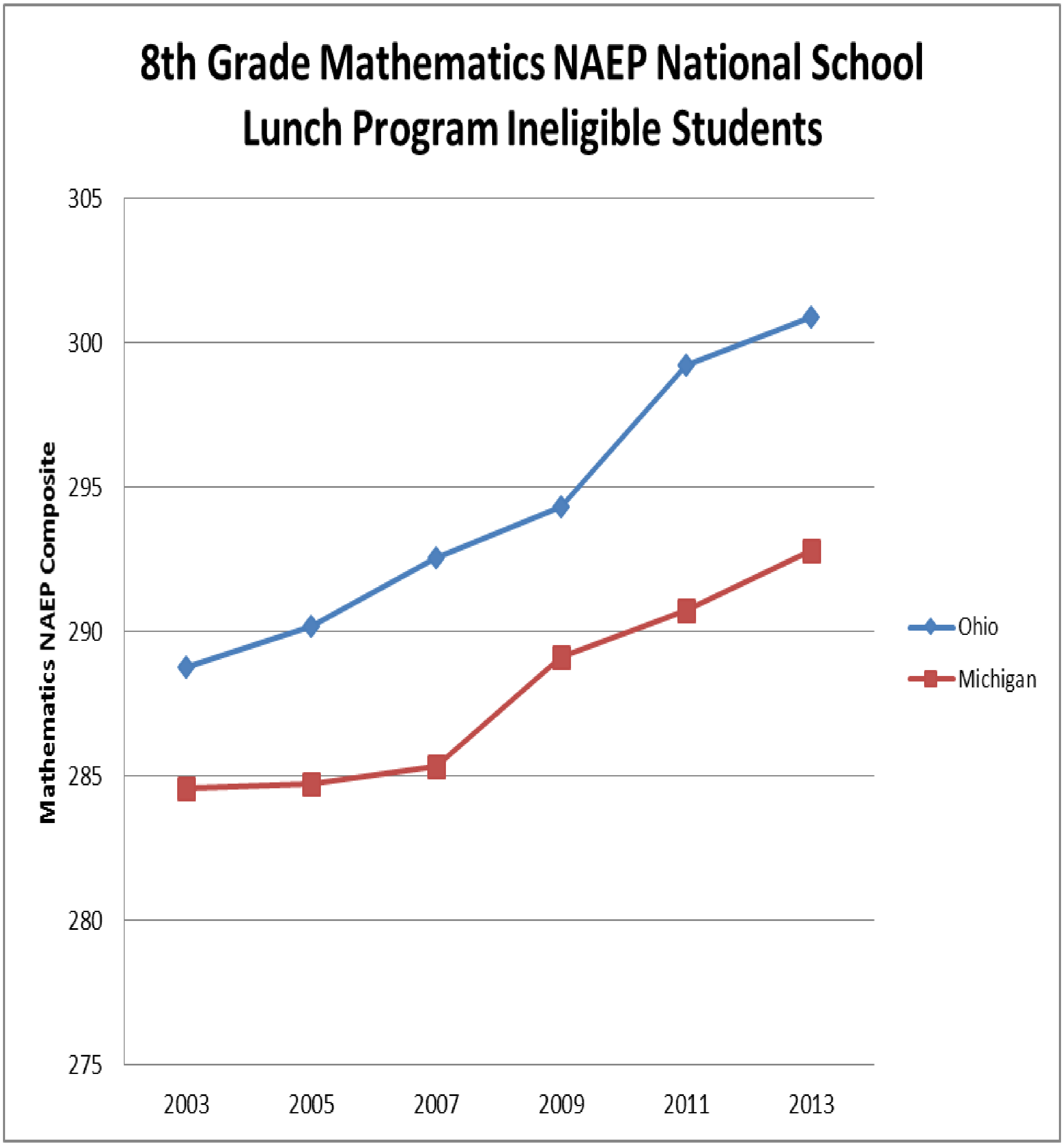
Figure 34. Mathematics NAEP composite all tested National School Lunch Program Eligible students in Tennessee and Georgia 1998-2013

APPENDIX H: RESEARCH QUESTION #6



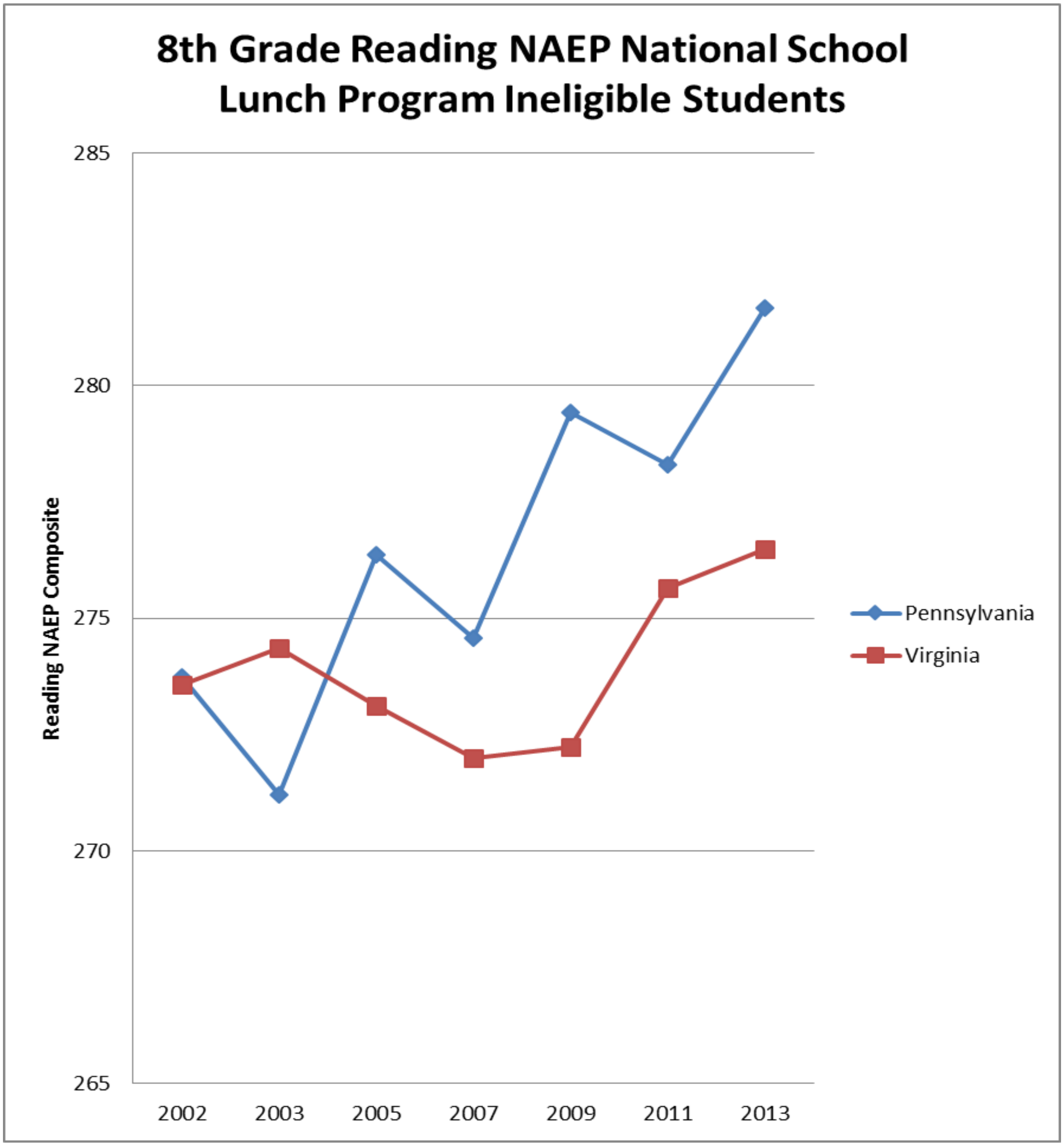
(NAEP, 2014)

Figure 35. Reading NAEP composite all tested National School Lunch Program Ineligible students in Ohio and Michigan 2002-2013



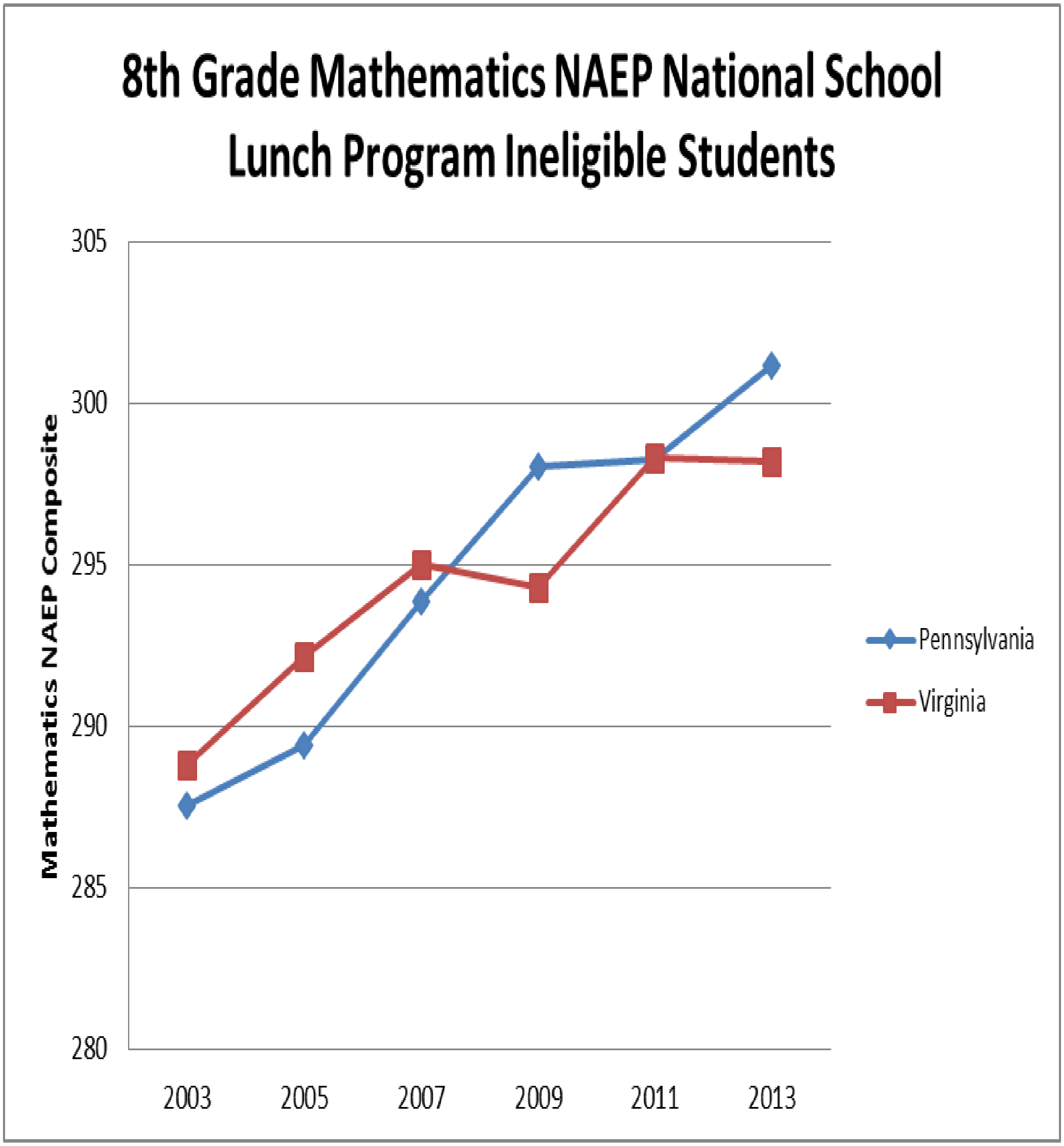
(NAEP, 2014)

Figure 36. Mathematics NAEP composite all tested National School Lunch Program Ineligible students in Ohio and Michigan 2003-2013



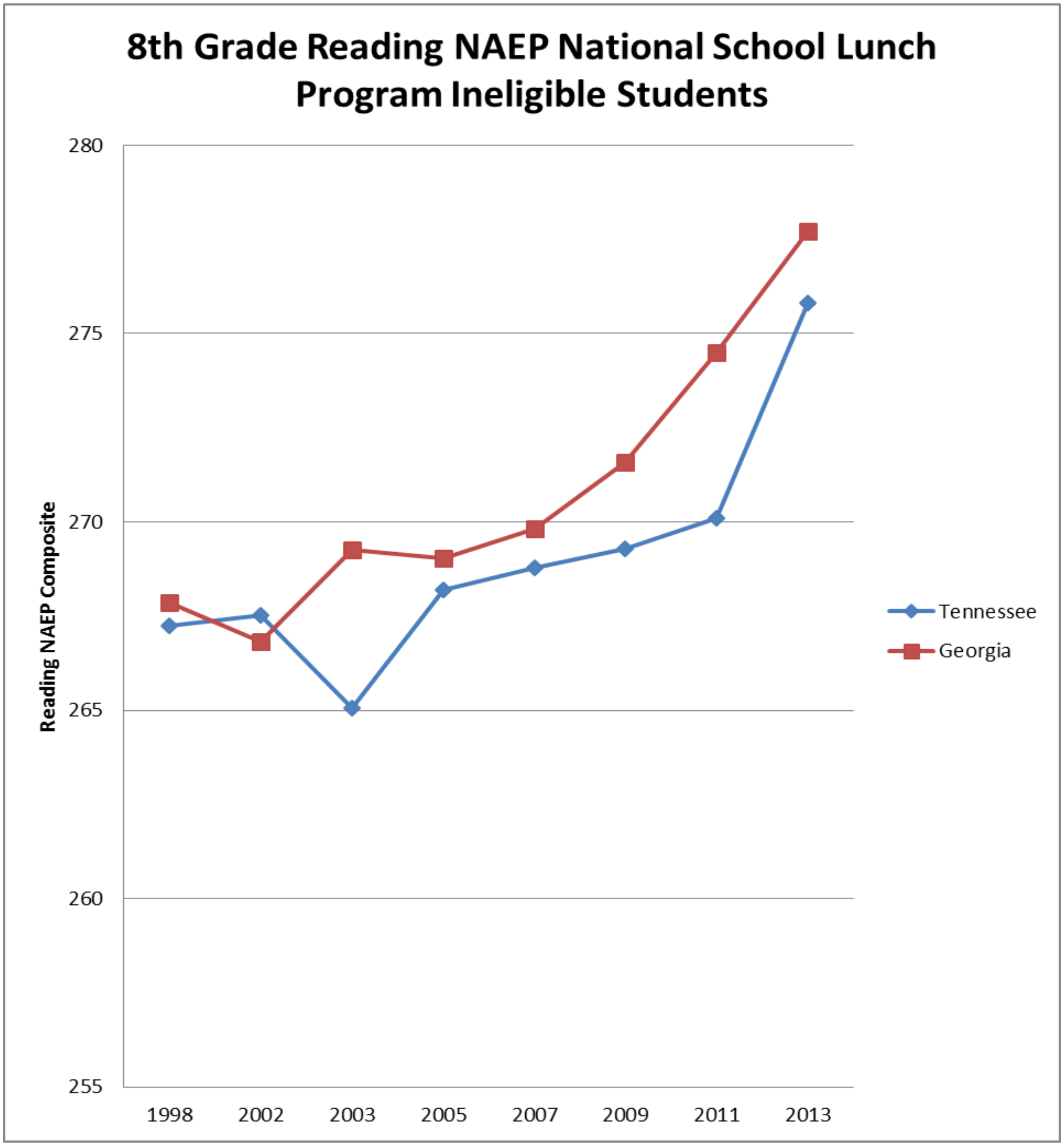
(NAEP, 2014)

Figure 37. Reading NAEP composite all tested National School Lunch Program Ineligible students in Pennsylvania and Virginia 2002-2013



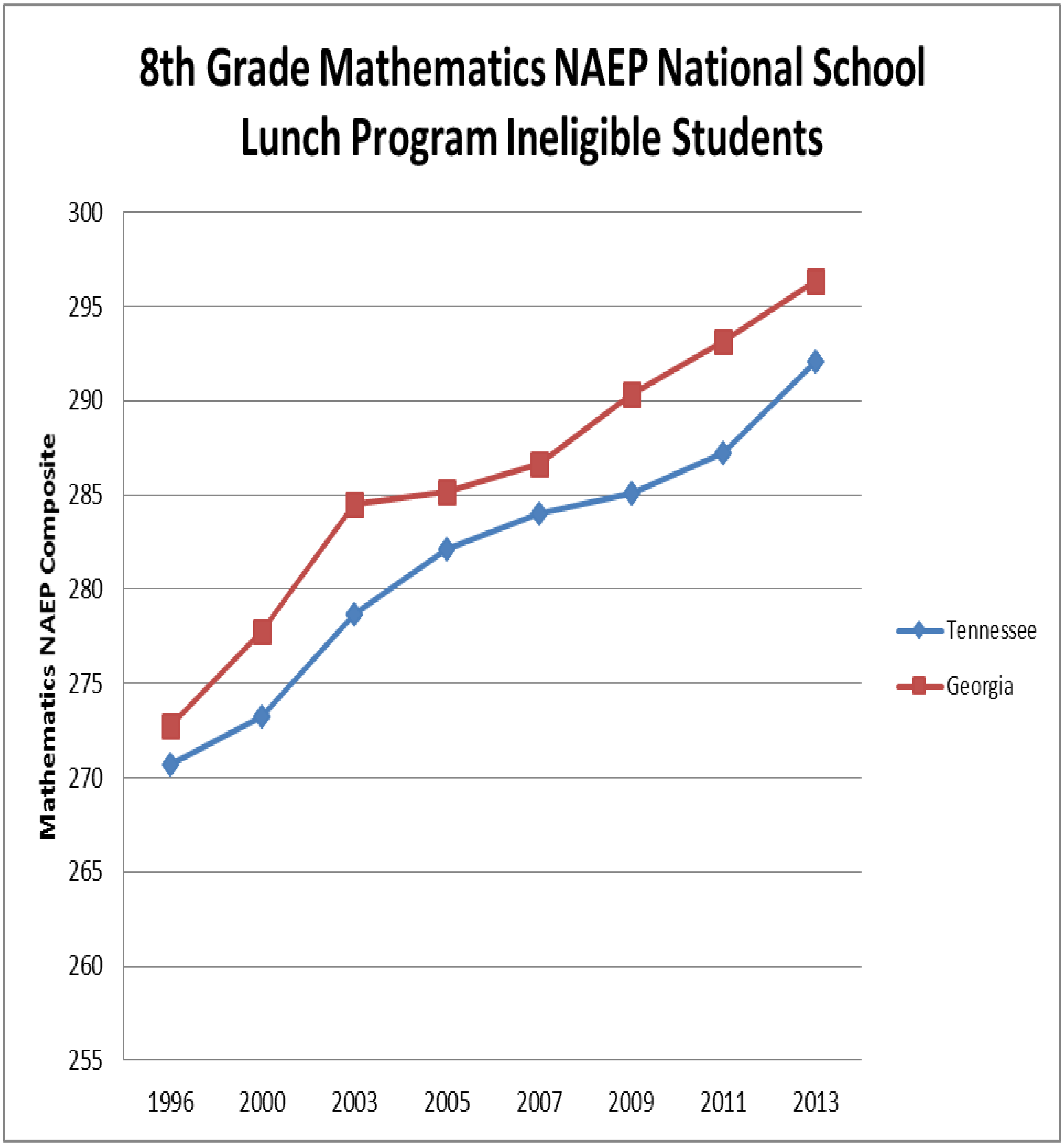
(NAEP, 2014)

Figure 38. Mathematics NAEP composite all tested National School Lunch Program Ineligible students in Pennsylvania and Virginia 2003-2013



(NAEP, 2014)

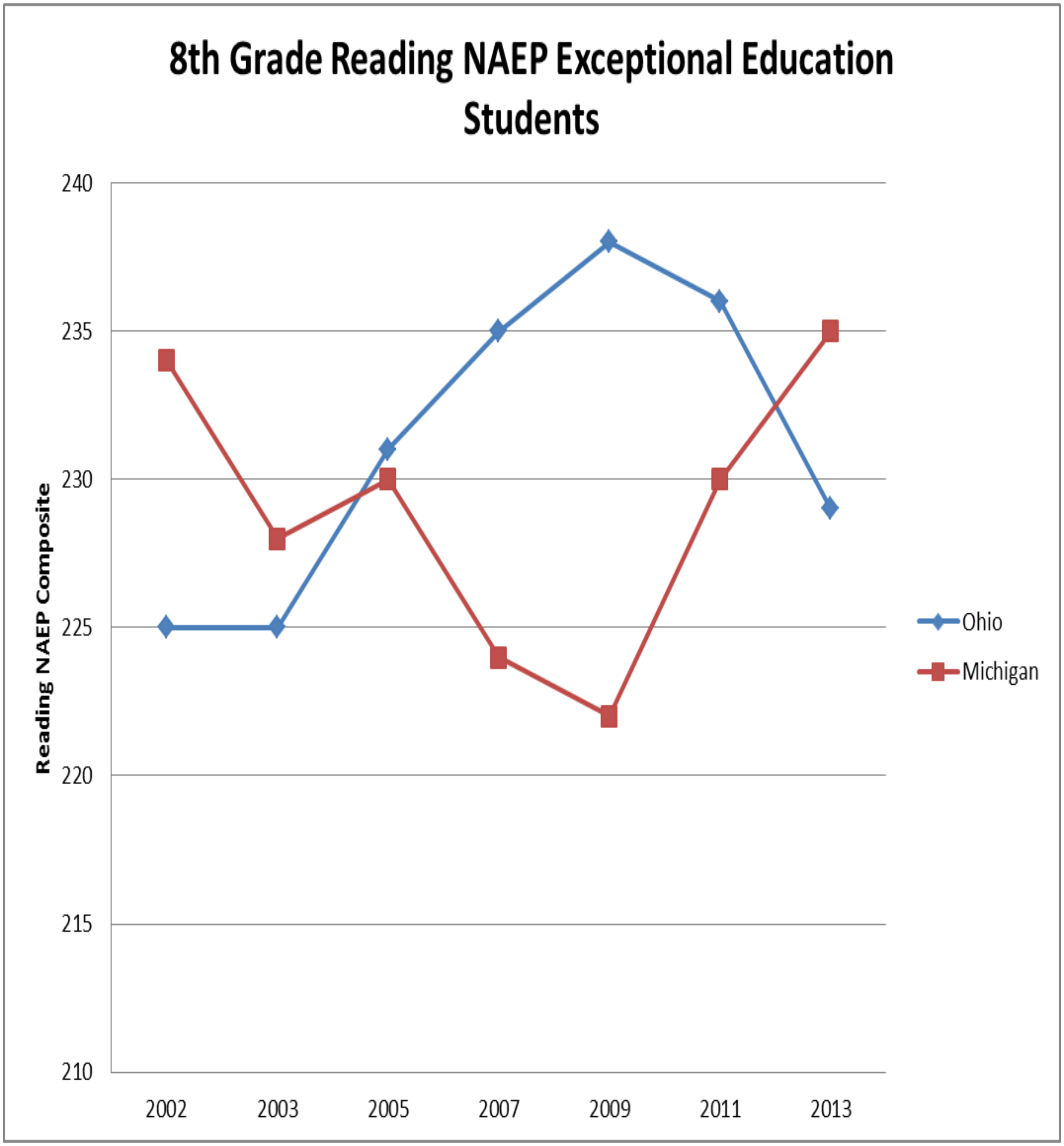
Figure 39. Mathematics NAEP composite all tested National School Lunch Program Ineligible students in Tennessee and Georgia 1998-2013



(NAEP, 2014)

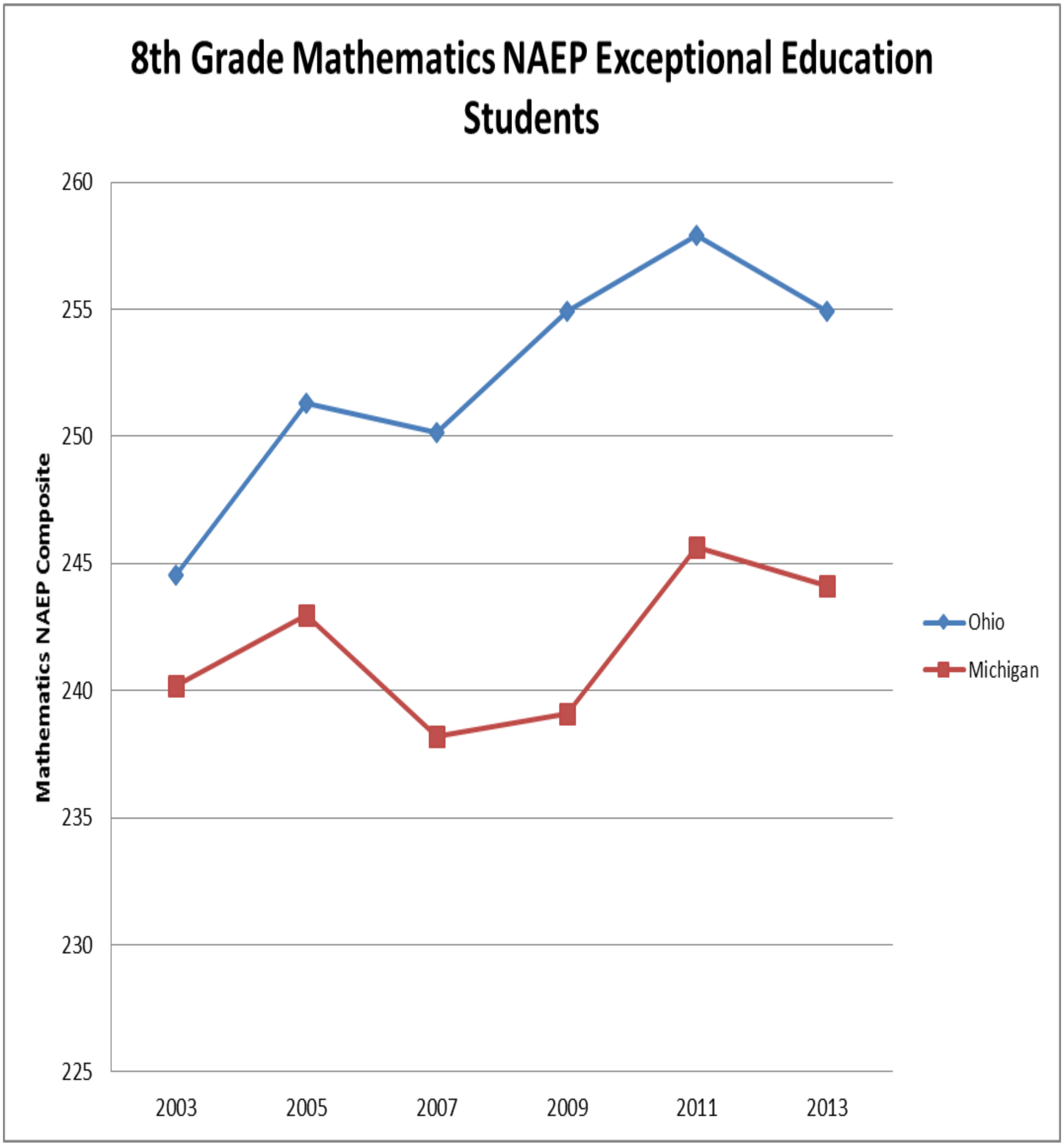
Figure 40. Reading NAEP composite all tested National School Lunch Program Ineligible students in Tennessee and Georgia 1996-2013

APPENDIX I: RESEARCH QUESTION #7



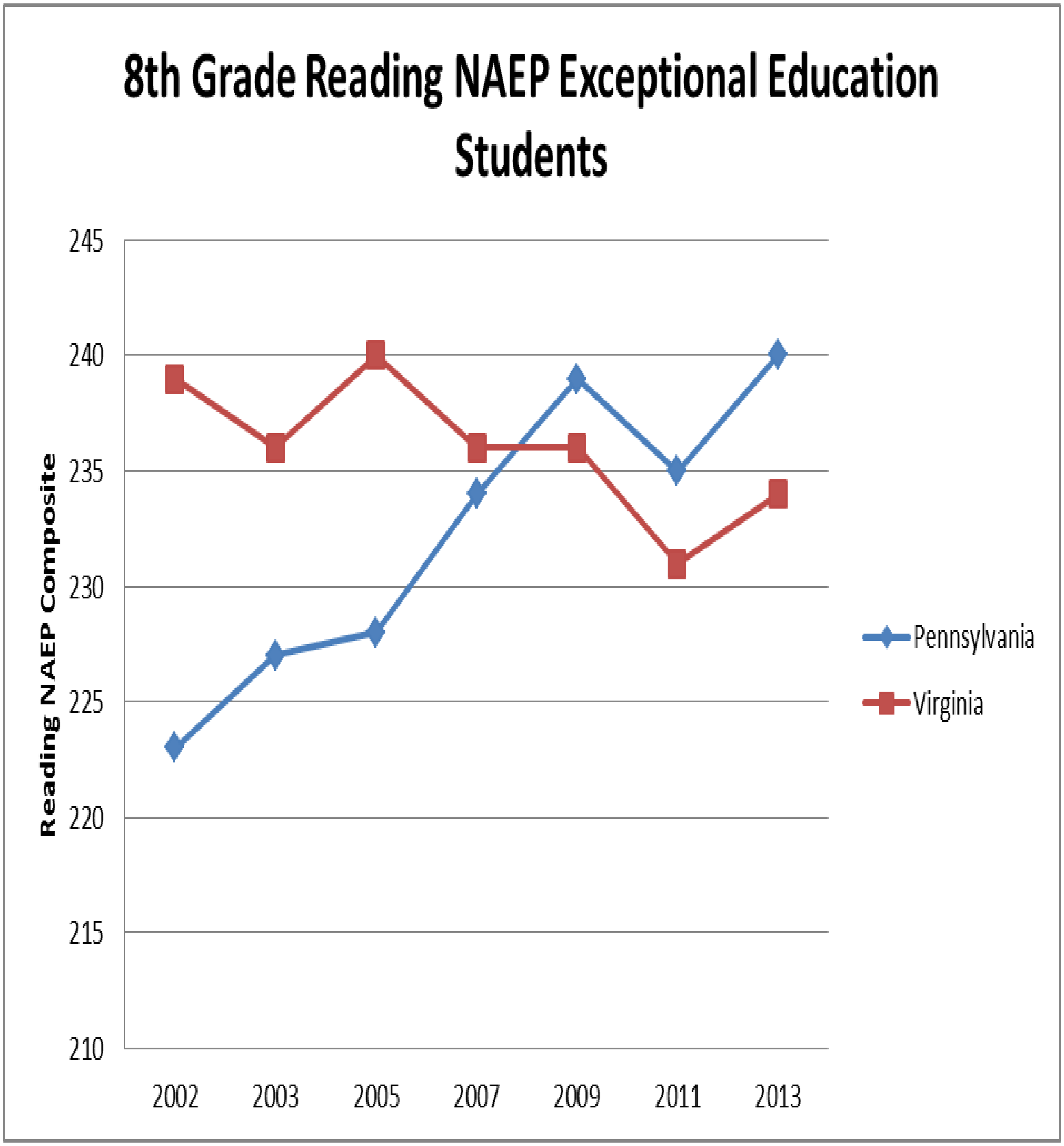
(NAEP, 2014)

Figure 41. Reading NAEP composite all tested Exceptional Education students in Ohio and Michigan 2002-2013



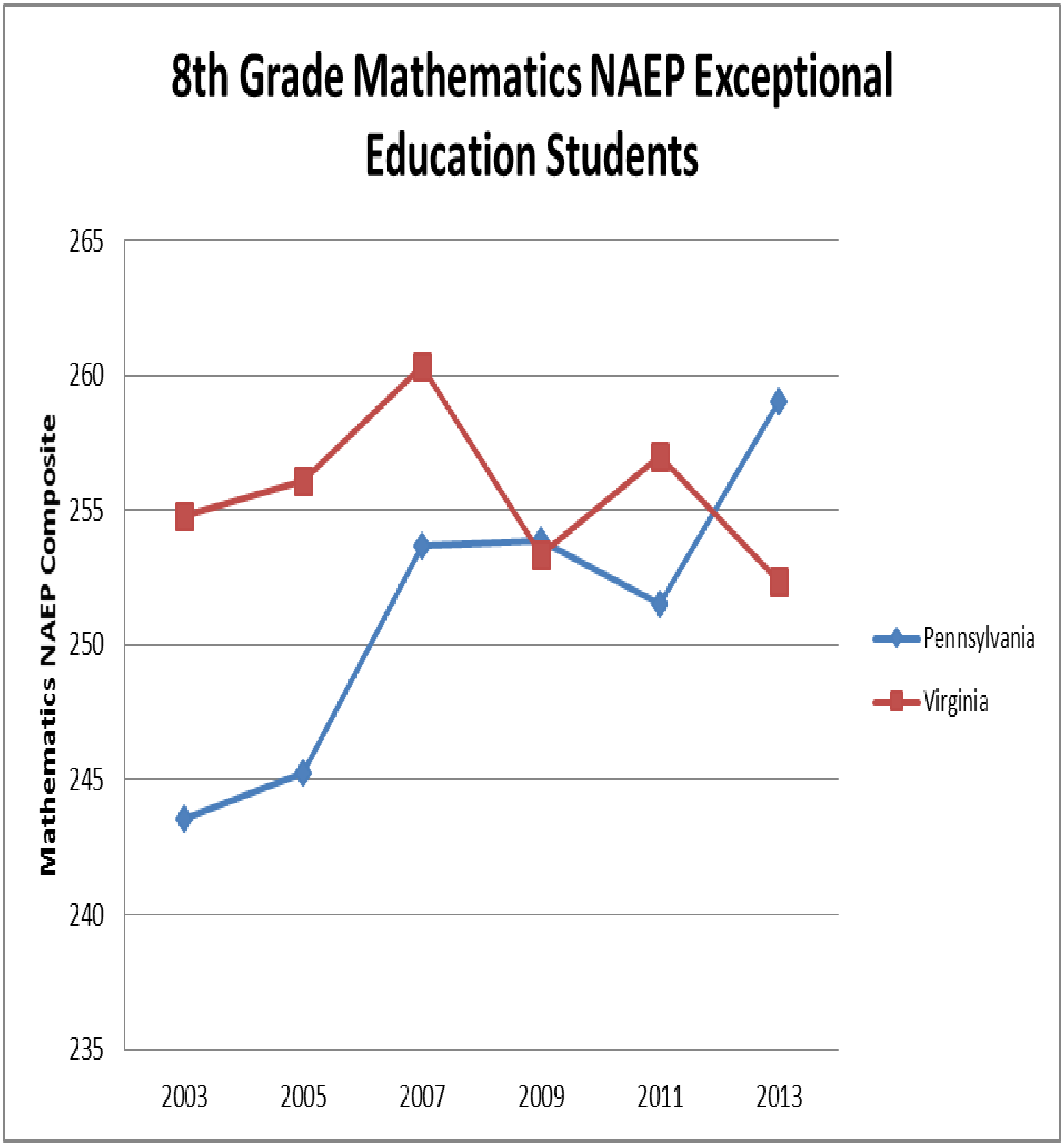
(NAEP, 2014)

Figure 42. *Mathematics NAEP composite all tested Exceptional Education tested students in Ohio and Michigan 2003-2013*



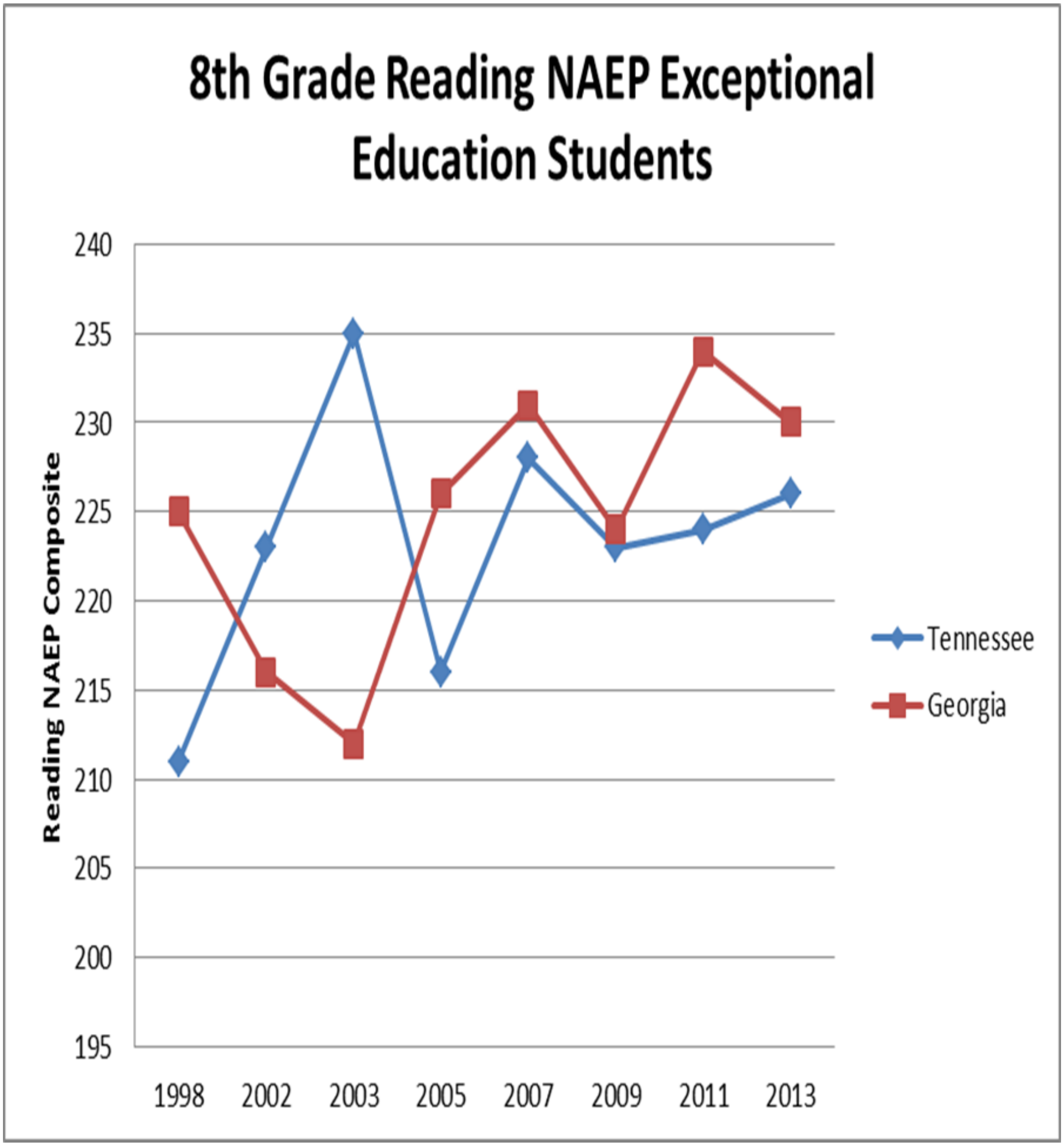
(NAEP, 2014)

Figure 43. *Reading NAEP composite all tested Exceptional Education students in Pennsylvania and Virginia 2002-2013*



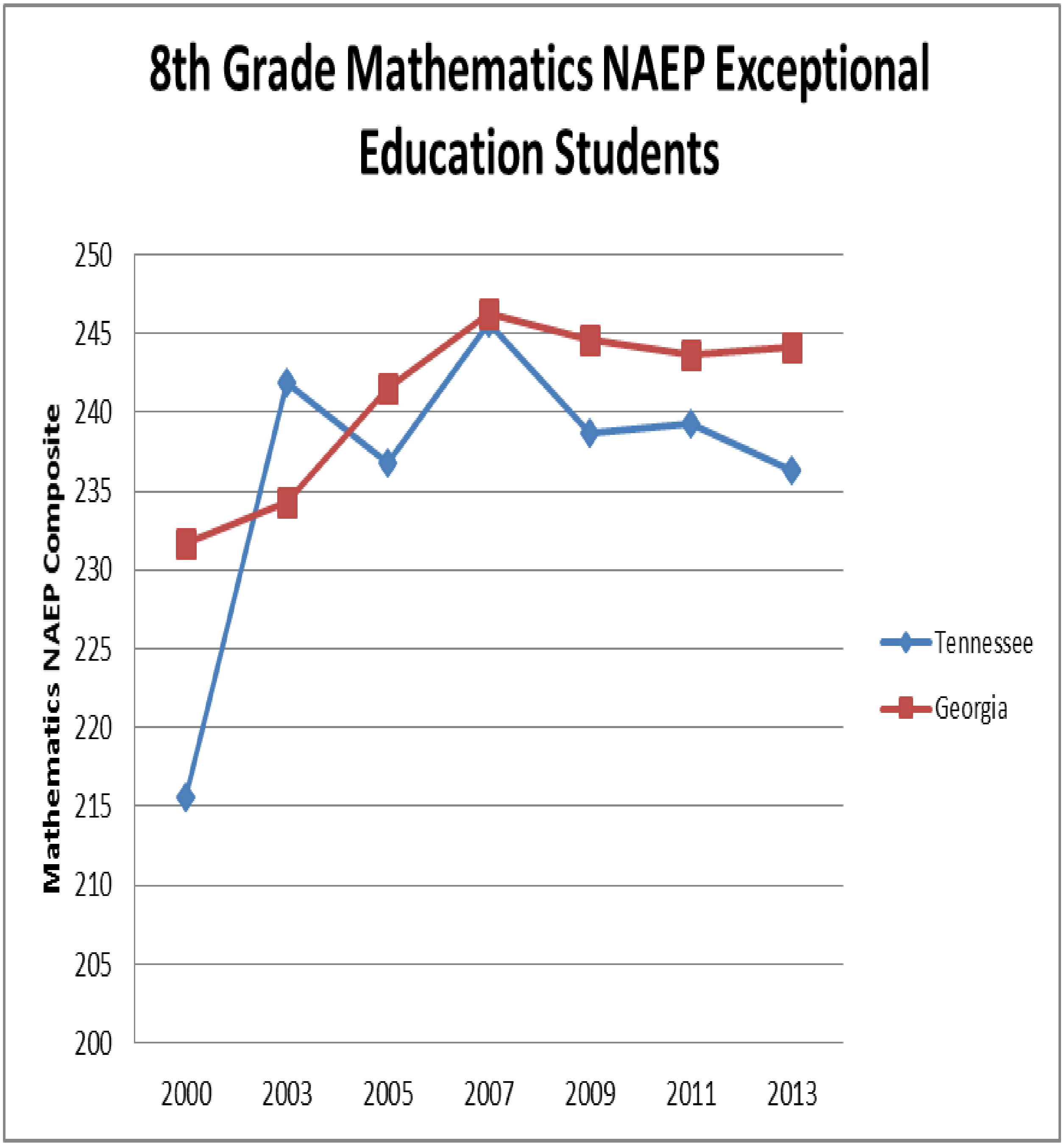
(NAEP, 2014)

Figure 44. Mathematics NAEP composite all tested Exceptional Education tested students in Pennsylvania and Virginia 2003-2013



(NAEP, 2014)

Figure 45. Reading NAEP composite all tested Exceptional Education students in Tennessee and Georgia 1998-2013



(NAEP, 2014)

Figure 46. *Mathematics NAEP composite all tested Exceptional Education tested students in Tennessee and Georgia 2000-2013*

APPENDIX J: INFORMED CONSENT LETTER

Date:

Dear State Department of Education Representative:

Thank you for taking time to offer some insight as to the workings of the Value Added Assessment Model in your state. I have contacted you for an interview to complete my study on the impact of the Value Added Assessment Model on K-12 education. My objective is to obtain an understanding of the role of the Value Added Assessment Model in states that have utilized Value Added Assessment Model for more than three years. I have a few short questions and, perhaps, a few follow-up questions that should take no more than 10 minutes to answer. They are accompanying this informed consent letter.

This process is voluntary, and there are no known risks. Assisting with this study may benefit future research and help develop best practices for education accountability. The interview will be recorded but only for purpose of insuring that the researcher is accurate in reporting the information resulting from the interview.

If you have questions about this research, please contact Dan Carter at (321) 266-8481 or my faculty supervisor, Dr. Rosemarye Taylor, Professor of Educational Leadership in the College of Education and Human Performance in the **School of Teaching, Learning, and Leadership** at the University of Central Florida. Her contact information is the following: phone number 407-823-1469, email Rosemarye.Taylor@ucf.edu.

Research at the University of Central Florida involving human participants is carried out under the oversight of the Institutional Review Board (IRB). Questions or concerns about research participants' rights may be directed at UCF IRB Office at University of Central Florida, Office of Research and Commercialization, 12443 Research Parkway, Suite 302, Orlando, FL 32826-3252. The phone number is 407-823-2901.

By agreeing to participate in this interview you are providing your informed consent.

Best Regards,
Dan Carter, Principal Investigator
Doctoral Candidate, University of Central Florida
321-266-8481
370 Narragansett St.
Palm Bay, Fl 32907

APPENDIX K: IRB LETTER

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

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

To : **James D. Carter**

Date : **October 22, 2014**

Dear Researcher:

On 10/22/2014 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:	MULTIPLE-CASE STUDY AND EXPLORATORY ANALYSIS OF THE IMPLEMENTATION OF VALUE-ADDED TEACHER PERFORMANCE ASSESSMENT ON EIGHTH GRADE STUDENT ACHIEVEMENT IN OHIO, PENNSYLVANIA AND TENNESSEE
Investigator:	James D. Carter
IRB ID:	SBE-14-10697
Funding Agency:	
Grant Title:	
Research ID:	N/A

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.

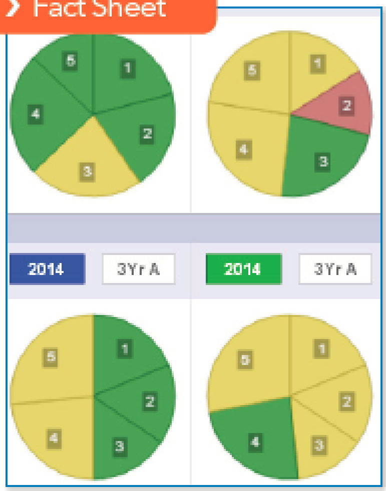
On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

Kanielle Chap

IRB Coordinator

APPENDIX L: EVAAMS

Fact Sheet



What does SAS® EVAAS® for K-12 do?

EVAAS is an educational resource for educators and policymakers. It uses test scores from a variety of assessments and follows individual students over time to provide two kinds of school improvement tools: reflective value-added reports to assess the effectiveness of districts, schools and teachers; and proactive projection reports on individual students' future performance.

Why is SAS® EVAAS® for K-12 important?

EVAAS provides insight and guidance to maximize all students' academic potential, regardless of their achievement level or background. Value-added measures reflect how much educators are accelerating student progress. Individual student trajectories facilitate customized, proactive planning for students' college and career readiness. With these tools, educators can provide more effective instruction to students, regardless of their achievement level.

For whom is SAS® EVAAS® for K-12 designed?

Teachers, principals, administrators and policymakers can all use customized reporting in EVAAS to be more effective and responsive in meeting individual school and district-wide goals.

SAS® EVAAS® for K-12

Reliably assess and predict student performance

With growing competition in the global economy, students need to be prepared for the future's high-performance careers. To support all students' postsecondary opportunities, EVAAS is based on the underlying belief that all students can learn and deserve to make appropriate academic progress each year, regardless of their prior achievement levels.

EVAAS provides accurate and comprehensive reporting at the state, district, school, classroom and student levels. This allows flexibility to respond to factors in varying education climates, such as testing changes, policy decisions, missing test data, and the mobility of teachers and students.

As school districts strive for continuous improvement with limited resources, administrators need to develop their most valuable human capital investment – teachers. With shrinking budgets and growing class sizes, teachers are challenged to effectively differentiate instruction to ensure that high achievers are challenged, middle achievers are encouraged, and low achievers are supported with appropriate instructional strategies. Easily understandable reporting gives teachers, guidance counselors, principals and other administrators the reliable information they need to have meaningful conversations regarding appropriate instructional strategies for each student.

After more than 20 years of development and practice, EVAAS is the most comprehensive reporting package of value-added metrics available in the educational market because it addresses all these challenges. Used in thousands of districts – statewide in Tennessee, Ohio, Pennsylvania and North Carolina – EVAAS assesses the influence of districts, schools and teachers on student progress rates and enables personalized educational planning at the student level. Through the breadth of reporting and interactive nature of web delivery, districts and schools can set their own policy goals and assess their success.

Educators are using EVAAS to:

- Differentiate instruction and provide early response to interventions.
- Differentiate professional development for educators based on their individual needs.
- Facilitate data-driven school improvement planning.
- Maximize scarce resources by evaluating the impact of programs on student growth.
- Inform human capital decisions such as teacher allocation, evaluation and award models.
- Identify and increase student enrollment in rigorous coursework.
- Increase the percentage of students who are proficient and college- or career-ready.
- Prevent dropouts and increase graduation rates through early identification of at-risk students.
- Facilitate data-driven parent-teacher conversations.
- Support policy changes and aid in state and federal accountability.

Benefits

- **Include more students for more reliable progress estimates.** EVAAS accommodates student and teacher mobility by not excluding students with missing test scores from the analysis. By avoiding the bias that can occur from systematically excluding those students, this approach provides the most reliable estimate of effectiveness for a district, school or teacher.
- **Include more data for more precise progress estimates.** Multivariate longitudinal models ensure that educators are not held accountable based on a single test score. EVAAS uses up to five years of data for an individual student by simultaneously analyzing scores across all grades and subjects. This process minimizes the influence of measurement error associated with individual assessments and increases the precision of the estimates. Value-added models are continually refined to increase precision while reducing measurement error and bias.
- **Accommodate team teaching.** Teacher value-added reporting accommodates team teaching and other teaching arrangements, where more than one teacher has instructional responsibility for a student's learning.
- **Give policymakers flexibility in choosing assessments.** EVAAS can analyze tests that are on differing vertical or horizontal scales. This flexibility around scaling ensures continuity of reporting when tests change over time, giving policymakers more options in their testing. The criteria for inclusion in EVAAS analyses are: reliability, high correlation with curricular standards, and sufficient stretch in the reporting scale to measure the achievement of both very low- and very high-achieving students in a grade and subject. Any assessment meeting these criteria can be used in the modeling, including some interim or short-cycle assessments. The ability to refresh the EVAAS reporting multiple times within the school year gives educators the instructional tools needed to formatively adjust their practices to best meet the needs of all students.
- **Gain the benefits of a SAS-hosted environment.** Because EVAAS is a hosted application, client partners do not need to make investments in hardware, software or staff. SAS manages sensitive data for a variety of global customers on secure servers with all of the required system backups in place to ensure the safety and privacy of sensitive data.
- **Advance knowledge through training.** EVAAS provides professional development specific to the needs of each education partner. The learning opportunities include interactive webinars that introduce users to the application and describe how to use the reporting to address real-world educational problems. In addition, interactive multimedia modules are embedded within the application to explain the reporting. Online help guides users through interpreting and applying the reports, and provides instant answers to many common questions. Plus, access to the EVAAS educator support team is available through a Contact Us link in the application.
- **Gain the advantages of decades of experience.** The EVAAS team is composed of former educators, professional statisticians, policy analysts and web specialists with as many as 20 years of experience delivering value-added modeling in a production environment across more than 20 states.
- **Have confidence in an externally validated methodology.** The reliability of EVAAS reporting has been reviewed and confirmed by prominent experts such as four US Department of Education Peer Review Committees, the Government Accountability Office, the RAND Corporation and WestEd.

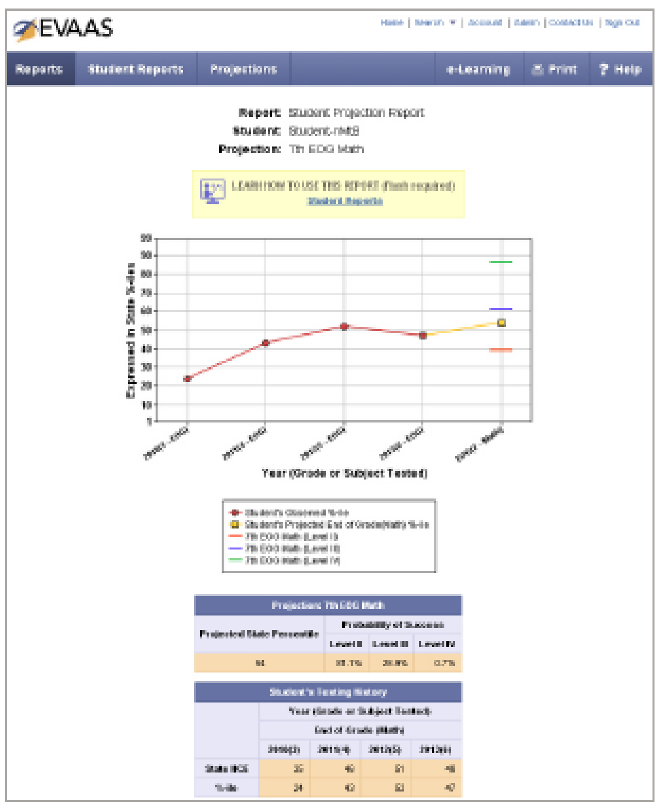


Figure 1: A sample student-level projection.

Overview

Reliability of information has been, and always will be, the primary driver behind the development of all EVAAS models. By following the progress of individual students over time, EVAAS provides balanced reporting that supports all students' growth while fairly and transparently assessing the effectiveness of educators.

Reflectively Identify Strengths and Opportunities for Improvement

Insights derived from analyzing years of data lead to intelligent resource and instructional decisions that improve teaching and learning.

EVAAS provides value-added analysis grade-to-grade and subject-to-subject for districts, schools and teachers, and reporting provides valuable feedback about curriculum implementation and the success of the instructional strategies.

Predict Student Success

EVAAS reports on students' predicted success probabilities at numerous academic milestones – such as the upcoming grade level proficiency, high school graduation requirements and various college success indicators, including advanced placement exams, the SAT and the ACT. Individual student reporting is accessible to every teacher who is authorized by local policy to view the student-level information.

Make Comparisons to Gain Insight

Interactive scatterplots enable users to visualize the impact of various progress metrics and the students who are served. For example, student achievement or growth can be compared to a variety of socioeconomic and demographic variables to assess whether all subgroups experience optimum educational outcomes. This powerful feature allows users to select the x and y axes variables to plot; drill down to detailed performance information at the district, school and teacher level; and zoom into any area within the scatterplot for further analysis.

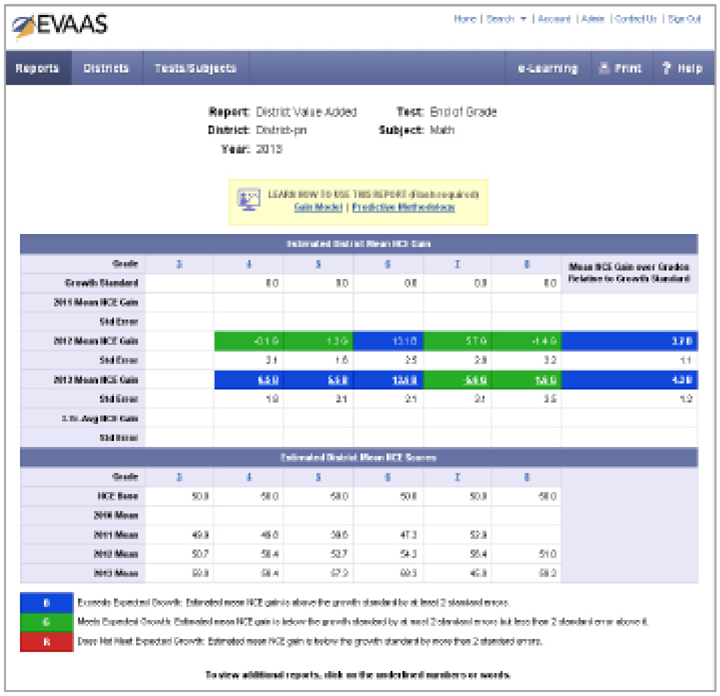


Figure 2: A sample value-added report for a district.

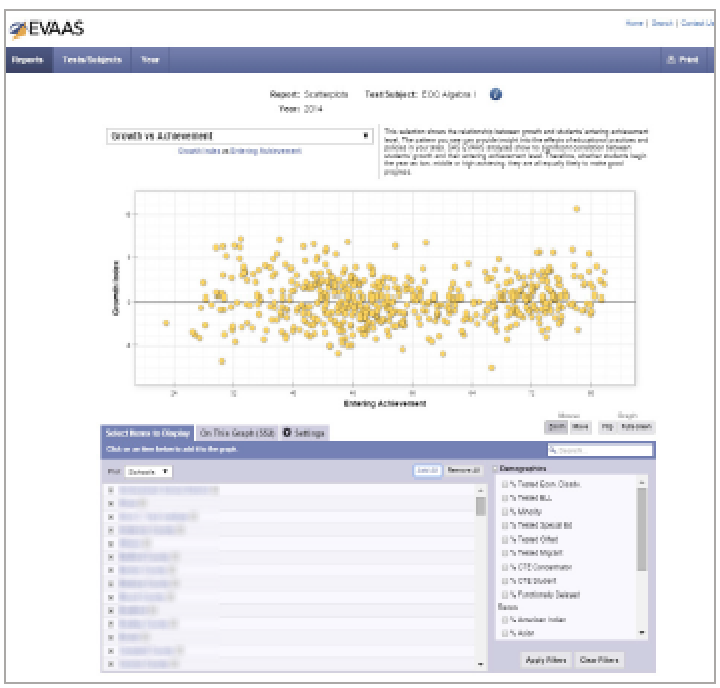


Figure 3: A scatterplot showing the relationship between growth and student achievement.

Identify Students at Risk for Underachievement

By viewing the academic preparedness of incoming students as well as the growth patterns within classrooms, educators have the opportunity to plan effectively, differentiate instruction and help students progress year to year. Diagnostic reports reveal at-risk students, thereby enabling earlier intervention.

Use Custom Reports

With custom reports, educators are empowered with the flexibility to display information in a variety of forms such as dashboards that report college readiness or upcoming grade-level proficiency, or custom student lists to monitor the effectiveness of programs. EVAAS can also build customized reports into the web application for frequent use based on specific needs.

Export Data

Authorized users can export both value-added and projection results, and download them into a variety of formats for independent use and research.

Verify Student-Teacher Linkages

The EVAAS roster verification process enables teachers, school administrators and district administrators to verify that teachers are linked accurately to students. By enabling educators to specify proportions of instructional responsibility for each student, EVAAS ensures that students' academic growth is tied to individual teachers in the fairest and most accurate way possible.

EVAAS applies the statistical rigor necessary to provide precise and reliable information on students' growth by addressing these challenges:

- Including students with missing test scores without introducing major biases that come from either eliminating the data from students or by using overly simplistic imputation procedures.
- Exploiting all of the longitudinal data for each student, even when the historical data is on differing scales.
- Providing educational policymakers flexibility in the use of historical data when tests have changed over time.
- Managing the data challenges associated with student and teacher mobility.

Build Confidence With a Secure, Proven Process

Upon receipt of the student testing data through a secure transfer process, the EVAAS team performs analysis of the data, including read-in, merging and quality control processes. The flexible and transparent EVAAS modeling is robust enough to accommodate a variety of adjustments in support of district or state policy changes.

Reports can be delivered in as little as four to six weeks after receiving clean and usable data. The district, school and teacher value-added estimates and student projections are then provided through a secure, hosted website and can be customized to fit a particular state or district.

Get more information about SAS EVAAS for K-12 at sas.com/evaas.

APPENDIX M: CENTER FOR GREATER PHILADELPHIA PERMISSION TO USE MAP

Katie Schlesinger <katie.schlesinger@gmail.com>
Thu 10/9/2014 9:48 AM

Hi Dan,

Yes, you may certainly use the map from our website. .

Thanks,
Katie

Katie Schlesinger
Project Manager, Operation Public Education
katie.schlesinger@gmail.com
215.898.8713 (office)
713.628.5152 (cell)

On Wed, Oct 8, 2014 at 11:00 PM, carter.dan <carter.dan@knights.ucf.edu> wrote:
Hello,

I am a Doctoral Student at the University of Central Florida. I am writing my dissertation on the Value Added Model. I am writing you to request permission to use your map as a figure (with citation of course) in my dissertation. The map gives the reader a great visual as a snapshot for VAA across the United States. The map I am referring to is located at the web address below.

http://www.cgp.upenn.edu/ope_nation.html

Thank you,
Dan Carter

LIST OF REFERENCES

107th Congress. (2002, January 8). Public Law – 107 -110. Retrieved November 5, 2013, from http://www.ed.gov/policy/elsec/leg/esea02/107-110.pdf

126th Ohio House of Representatives General Assembly. (2005, December 13). House Bill 107. Retrieved July 21, 2014, from http://www.legislature.state.oh.us/analysis.cfm?ID=126_HB_107&ACT=As%20Enrolled&hf=analyses126/05-hb107-126.htm#_ftn1

Anderman, E. M., Anderman, L. H., Gimbert, B. G., & Yough, M. S. (2010). Value-added models of assessment: implications for motivation and accountability. *Educational Psychologist*, 45(2), 123-137. doi:10.1080/00461521003703045

Azzam, A. (2004). NCLB: Up close and personal. *Educational Leadership*, *62*(2), 87.

Bejoian, L. M., & Reid, K. D. (2005). A disability studies perspective on the bush education agenda: The no child left behind act of 2001. *Equity and Excellence in Education, 38*(3), 220-231. Retrieved from ProQuest.

Bracey, G. W. (2007). The first time 'everything changed'. *Phi Delta Kappan*, *89*(2), 119-136.

Bratton, S. E., Horn, S. P., & Wright, S. P. (1996). Using and interpreting Tennessee's value-added assessment system: A primer for teachers and principals. Knoxville, Tn: University of Tennessee, Value Added Research and Assessment Center.

Burns, S., Wang, X., Henning, A., & National Center for Education Statistics, (2011). NCES handbook of survey methods. NCES 2013-609. *National Center for Education Statistics*, Retrieved from EBSCO*host*.

Center for greater philadelphia. (n. d.). Retrieved Novemeber 23, 2013, from http://www.cgp.upenn.edu/ope_nation.html

Chubb, J. E., & Moe, T. M. (1990). *Politics markets America's schools*. Washington, DC: The Brookings Institution.

Chudowsky, N., Chudowsky, V., & Center on Education Policy. (2010). State test score trends through 2008-09, part 1: rising scores on state tests and NAEP. *Center on Education Policy*, Retrieved from EBSCO*host*.

Coleman, J. S., Campbell, E. Q. , Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R.L. (1966). Equality of educational opportunity. Washington, DC: U.S. Government Printing Office.

Common Core State Standards Initiative (CCSSI). (2012). *About the standards*. Retrieved September 8, 2012, from http://www.corestandards.org/about-the-standards

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed mode surveys*: *The tailored design method* (3rd ed.). Hoboken, New Jersey: John Wiley & Sons, Inc.

Dillon, S. (2010, August, 31). Formula to Grade Teachers' Skill Gains Acceptance, and Critics. *The New York Times.* Retrieved from http://www.nytimes.com/2010/09/01/education/01teacher.html?pagewanted=all&_r=0

Feuer, M. J., Holland, P. W., Green, B. F., Berenthal, M. W., & Cadelle Hemphill, F. (Eds). (1999).*Uncommon measures: Equivalence and linkage among educational tests*. Washington, DC: National Academy Press.

Goldhaber, D. D., Goldschmidt, P., & Tseng, F. (2013). Teacher value-added at the high-school level: Different models, different answers?. *Educational Evaluation & Policy Analysis*, *35*(2), 220-236.

Goldhaber, D., & Hansen, M. Race, gender, and teacher testing: How informative a tool is teacher licensure testing?. *American Educational Research Journal*, March 2010, 218-251. Retrieved from EBSCO*host*.

Gordon, R., Kane, T. J., & Staiger, F. O. (2006). Identifying effective teachers using performance on the job. *Hamilton Project: Discussion Papers*, (1), 1-35.

Hamilton, L. S., & McCaffrey, D. F. (2007). Value-Added Assessment in Practice: Lessons from The Pennsylvania Value-Added Assessment System Pilot Project. Santa Monica California. RAND Corporation.

Hershberg, T., Simon, V. , & Lea-Kruger, B. (2004). The revelations of value-added. *School Administrator, 61*(11), 10-12, 14.

Hise, M. M. (1994) *Goals 2000: Educate America Act. Fordham Law Review*. 63(45). Retrieved from EBSCO*host*.

Jerald, C. D., & Education, T. (2009). The value of value-added data. K-12 Policy. *Education Trust*, Retrieved from EBSCO*host*.

Johanningmeier, E. V. (2010). "A nation at risk" and "Sputnik": compared and reconsidered. *American Educational History Journal*, *37*(2), 347-365.

Jones, L. (1996). A history of the National Assessment of Educational Progress and some questions about its future. *Educational Researcher, 25*, 15-22. doi: 10.2307/1176519

Jones, L., & Olkin, I. (2004). The nation's report card: Evolution: and perspectives. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Kirst, M., & Jung, R. (1980). The Utility of a Longitudinal Approach in Assessing Implementation: A Thirteen Year View of Title I, ESEA. *Educational Evaluation and Policy Analysis.*

Kober, N., Rentner, D., & Center on Education Progress. (2011). More to Do, but less capacity to do it: States' progress in implementing the recovery act education reforms. Center On Education Policy.

Konstantopoulos, S. (2014). Teacher Effects, Value-Added Models, and Accountability. Teachers College Record, 116(1), 1-21.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation & Policy Analysis*, *25*(3), 287-298.

Linn, R. L., California Univ. & National Center for Research on Evaluation. (2003). Accountability: Responsibility and Reasonable Expectations. CSE Report.

Lomaz, E.D., & Kuenzi, J.J. (2012). Value-added modeling for teacher effectiveness. Congressional Research Service. CRS Report 7-5700.

Manna, P., & Ryan, L. (2011). Competitive grants and educational federalism: President Obama’s race to the top program in theory and practice. *Publius: The Journal of Federalism*, *41*(3), 522-546. Retrieved from EBSCO*host*.

Marzano, R. J. (2003). What works in schools: Translating research into action. Alexandria, VA: Association of Supervision and Curriculum Development.

McCaffrey, D. F., Lockwood, R. J., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability*. RAND Corporation. P.O. Box 2138, Santa Monica, CA 90407-2138. Retrieved from ProQuest.

Mondale, S., & Patton, S. (2001). The story of American public education. Boston: Beacon Press.

Murphy, J. T. (1971). Title I of ESEA: The politics of implementing federal education reform. *Harvard Educational Review*.

National Center for Education Statistics (1999). NAEP 1999 – Trends in academic progress: Three decades of student performance. Washington, DC: United States Department of Education.

National Center for Education Statistics (2001). The Nation’s Report Card 2000.Washington, DC: United States Department of Education.

National Center for Education Statistics. (2010). National assessment of educational progress. Retrieved January 5, 2012, from http://nces.ed.gov/nationsreportcard/about/state.asp

National Clearinghouse for Bilingual Education. (2001). No Child Left Behind Act of 2001. Title I: Improving the Academic Achievement of the Disadvantaged. 107th Congress, 1st Session.

National Commission on Excellence in Education (1983). A Nation At Risk: The Imperative For Educational Reform. An Open Letter to the American People. A Report to the Nation and the Secretary of Education.

No Child Left Behind Act of 2001. Title I: Improving the Academic Achievement of the Disadvantaged. 107th Congress, 1st Session.

Office of Education. (1970). Improving Education through ESEA: 12 Stories. Elementary and Secondary Education Act of 1965.

Orfield, G. (1969). The Reconstruction of Southern Education: *The Schools and the 1964 Civil Rights Act*. New York: Wiley-Interscience.

Partnership for Assessment of Readiness for College and Careers (PACC). (2012a). *About PARCC*. Retrieved September 8, 2012, from parcconline.org/about-parcc

Partnership for Assessment of Readiness for College and Careers (PACC). (2012b). *Implementation*. Retrieved September 8, 2012, from http://www.parcconline.org/implementation

Resnick, D. P. (1980). Minimum competency testing historically considered. *Review of Research in Education*, *8*, 3-29.

Riley, B. E. (1977). *Accountability in Education: A Recurring Concept*. Newcastle: England, Texas.

Riley, R. W. (1996). Improving America's schools. *School Psychology Review*, *25*(4), 477-84.

Sanders, W. (1998). Value-added assessment. *School Administrator*, *55*(11), 24-7. Retrieved from Education Full Text database.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAMS): Mixed-model methodology in educational assessment. *Journal Of Personnel Evaluation In Education, 8*(3), 299-311.

SAS EVAAS for K-12. (2014). Fact Sheet: *Reliability assess and predict student performance*. Retrieved July 29, 2014, from http://www.sas.com/content/dam/SAS/en_us/doc/factsheet/education-eVAAMs-104850.pdf

Schafer, W., Liu, M., & Wang, H. (2007, August). Content and grade trends in state assessments and NAEP. *Practical Assessment, Research and Evaluation*, *12*(9), 1-25.

Schwartz, R. B., & Robinson, M. A. (2000). Goals 2000 and the standards movement. *Brookings Papers On Education Policy*, 173-214.

Shaul, M., & Ganson, H. (2005). The no child left behind act of 2001: The federal government's role in strengthening accountability for student performance. *Review of Research in Education*, *29*, 151-165.

Smarter Balanced Assessment Consortium. (2012). *Member States.* Retrieved September 8, 2012, from http://www.smarterbalanced.org/about/member-states

Steeves, K., Bernhardt, P., Burns, J. P., & Lombard, M. K. (2009). Transforming American educational identity after Sputnik. *American Educational History Journal*, *36*(1), 71-87.

Superfine, B. M. (2005). The politics of accountability: The rise and fall of goals 2000. *American Journal of Education, 112*, 10-43.

Taylor, J., Stecher, B., O'Day, J., Naftel, S., Le Floch, K., & U.S. Department of Education. (2010). State and local implementation of the "no child left behind act". Volume IX--Accountability under "NCLB": Final Report. *US Department of Education*, Retrieved from EBSCO*host*.

Tyack, D., & Cuban, L. (1995). Tinkering toward utopia: A century of public school reform. Cambridge, Mass.: *Harvard University Press.*

United States Congress (1965a). Elementary and secondary education act of 1965, Retrieved from EBSCO*host*.

United States Congress (1965b). Improving America's schools Act of 1964. Report on H.R. 6 Together with minority, supplemental, and additional views committee on education and labor, House of Representatives, 103rd Congress, 2nd session, Retrieved from EBSCOhost.

United States Department of Agriculture (2012, August). National school lunch program: Fact sheet. Retrieved September 27, 2012 from http://www.fns.usda.gov/cnd/lunch/AboutLunch/NSLPFactSheet.pdf

United States Department of Education (2002, November 26). *Education department issues final regulations for No Child Left Behind Act*. Retrieved December 27, 2013 from http://www.mynaep.com/network/archive/legislation_Nov.asp

United States Department of Education. (2004). Executive summary of the no child left behind act of 2001. Retrieved November 5, 2013, from http://www.ed.gov./print/nclb/overview/intro/execsumm.html

United States Department of Education. (2009). Transcript of race to the top technical assistance workshop in Baltimore, Maryland. Washington, DC, Retrieved from EBSCO*host*.

United States Department of Education. (2014). *The nation's report card*. Retrieved Feburary 15, 2014, from http://nces.ed.gov/nationsreportcard/faq.asp#ques13

United States Department of Education. (2012). *Improving basic programs operated by local educational agencies.* Retrieved May 7, 2012, from http://www2.ed.gov/programs/titleiparta/index.html

United States Department of Labor. (1964). *Title VI, civil rights act of 1964.* Retrieved May 7, 2012, from http://www2.ed.gov/programs/titleiparta/index.html

Wells, J., & Westat, I. (2011). Teacher responses to pay-for-performance policies: survey results from four high-poverty, urban school districts. *Online Submission*, Retrieved from EBSCO*host*.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed). Thousand Oaks, California: Sage Publications.