5-2009

An Analysis of Data Collected from the 2007-2008 Tennessee State Report Card and the Variables Related to Science Test Results.

Julia Lorie Lamons
East Tennessee State University

An Analysis of Data Collected from the 2007-2008 Tennessee State Report Card and the Variables Related to Science Test Results

---

A dissertation

presented to

the faculty of the Department of Educational Leadership and Policy Analysis

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Doctor of Education

---

by

Julia W. Lamons

May 2009

Dr. Eric Glover, Chair

Dr. Cecil Blankenship

Dr. James Lampley

Dr. Terrence Tollefson

Keywords: Standardized Testing, Science, No Child Left Behind, Academic Achievement

ABSTRACT

An Analysis of Data Collected from the 2007-2008 Tennessee State Report Card and the Variables Related to Science Test Results

by

Julia W. Lamons

The purpose of this study was to determine whether Tennessee Comprehensive Assessment Program (TCAP) Achievement reading scores, school district per-pupil expenditures, school size, percentage of students qualifying for free and reduced-priced meals, and attendance were related to science TCAP test scores from the 2007-2008 school year. The data were gathered from an analysis of mean standardized test scores in reading and science of 8th graders in 67 school systems comprising 181 schools located throughout Tennessee. One hundred eighty-one schools configured grades 6 through 8 were used in this study. Only 177 schools had reported attendance available on the Tennessee Department of Education website.

Pearson correlations were performed between the 8th mean grade science TCAP scores and 8th grade mean reading scores, per-pupil expenditure, school size, attendance, and the percentage of students qualifying for free and reduced-priced meals. Independent-samples *t* tests were conducted to evaluate whether 8th grade mean science TCAP scores and 8th grade mean reading scores varied depending on whether the school per-pupil expenditure was above or below the state average of $8,345. Independent-samples *t* test were also conducted to evaluate whether 8th grade mean science TCAP scores and 8th grade mean reading scores varied depending on whether the school attendance percentage was above or below the state goal of 93%.

The school characteristic with the strongest association with the mean 8th grade science TCAP scores as computed by Pearson's correlation is the mean 8th grade reading TCAP scores. The values can be ranked as follows: mean 8th grade reading scores (.92) > percentage of students qualifying for free and reduced meals (-.84) > per-pupil expenditure (-.62) > attendance (.60) > school size (.23).

DEDICATION

I dedicate this study to my family:

To my son, Ayden, from whom I have learned more about life from in the past 9 months than I could ever teach him. I hope that I am able to instill in him the love of learning that I possess.

To my husband, Travis, for his constant love and support. Thanks for always believing in me.

To my parents, Larry and Helen, for their guidance. You have always encouraged me to follow my dreams.

To my sisters, Carita and Jaclyn, for encouraging me.

And to the rest of my family for always believing in me.

ACKNOWLEDGMENTS

I would like to acknowledge and thank:

First and foremost, I would like to thank the Lord. Without His guidance I would not have been able to complete this task. He has blessed my life in too many ways to count. *I can do all things through Christ who strengths me* (Philippians 4:13).

Dr. Eric Glover for his guidance and wisdom during this process. Dr. Glover's feedback, patience, and high expectations provided great support in keeping me focused on the completion of this study.

Dr. Cecil Blankenship, Dr. James Lampley, Dr. Terrence Tollefson, members of my committee, all of whom have provided valuable feedback and assistance to me for which I am grateful.

To my colleagues at work, thank you for the support and constant words of encouragement.

To all of the members of Cohort O.C., you are a great group of educators and it has been my pleasure to get to work with you.

A special thanks to Whitney Ramsey, my cohort member and friend, for her support, encouragement, and friendship throughout this entire process. You have become like a sister to me. I could not imagine anyone I would have rather gone on this journey with.

CONTENTS

Chapter

Chapter

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION TO THE STUDY

Standardized testing is a central feature in American public schools. These tests have been prevalent for close to 100 years. However, since the early 1980s, high-stakes standardized testing has spread dramatically from state to state to the dismay of many educators (Kohn, 2001). This increase was primarily attributed to the release of the report *A Nation at Risk* during the Reagan administration in 1983 (Amrein & Berliner, 2002). This report paved the way for many events that changed the face of education. These events included the revision of Public Law 94-142 in 1997 and again in 2004, now known as the *Individuals with Disabilities Education Improvement* Act, and the well known *No Child Left Behind* Act in 2001 (Hirsh, 2006). These laws impacted the way in which students were educated in the United States.

Among those who favored more stringent testing were policy makers and business leaders who were convinced that tough standards and accountability would save the nation's schools. Many leaders seemed to believe that educators would do a better job of teaching if the pressure for tougher standards increased (Kohn, 2001). Standardized tests are administered to millions of students across the United States each year for various reasons. Each spring, thousands of Tennessee students in grades three through eight take the Tennessee Comprehensive Assessment Program (TCAP) achievement test, the high-stakes test chosen by the state. The TCAP is a timed multiple-choice test that measures student knowledge in reading, language arts, mathematics, science, and social studies. The test questions are directly aligned to the state's adopted curriculum standards. Federal educational funds could be withheld from schools that performed poorly on the TCAP (Tennessee Department of Education, 2007c). As a

result, schools that are most in need of assistance might lose funding due, in part, to socioeconomic and demographic factors that are largely beyond their control (Kohn). Worse than the loss of funds, however, could be the continued narrowing of the instructional program as teachers across the nation devote their energies to preparing students for state and federal tests in lieu of more creative and worthwhile classroom activities (Berliner & Biddle, 1995; Kohn).

Legislators feel pressure from the general public to hold school systems and educators more accountable for student achievement. In recent years there has been a push for public education systems to be run more like private businesses (Kohn, 2001). In business, products are measured for quality assurance. Therefore, policymakers note that academic achievement of children should be measured forcing local school districts to place greater emphasis on increasing test scores. U.S. Secretary of Education Margaret Spellings (2007) has reported that standardized test scores led to many improvements in public schools throughout America.

Statement of the Problem

Standardized testing measures achievement in public schools. Increased accountability brought about by recent school reform placed greater emphasis on standardized test scores and created stressful challenges for educators across the country. The purpose of this study was to examine the relationship among several variables including reading scores, per-pupil expenditure, school size, attendance, and the percentage of students qualifying for free and reduced meals and to determine the relationship of those variables to science TCAP test scores of 8th grade students.

Significance of the Study

Standardized testing is a paramount issue facing public school systems. In Tennessee schools and school systems can lose significant funding if they do not perform well on the end of the year TCAP tests. With such stringent expectations impressed on local educators, it would be beneficial to know what variables possess the strongest correlation to science TCAP scores. By identifying the variables that display the strongest correlation to high standardized test scores in science, educators may be able to provide interventions that will lead to student achievement in that subject, resulting in higher standardized test scores and improved student learning.

Limitations and Delimitations

The instrument used in this study was the TCAP state report card. The study was delimited to data contained in the 2007-2008 Tennessee Schools Report Card. This study included students in only 8th grade. This study was restricted to 67 school systems and 181 schools located throughout Tennessee. The study was limited by only including schools that were configured grades 6 through 8.

Perhaps, the most obvious limitation of this study was the researcher's bias concerning the over emphasis on standardized testing. Personal bias was addressed through the use of outside experts and the careful examination of the data by the dissertation committee. A thorough examination of the literature from other research was examined while conducting this study.

Research Questions

The following research questions guided this study:

1. Is there a relationship between mean science TCAP scores of 8th graders in Tennessee and the following variables: a) mean reading TCAP scores, b) per pupil expenditure, c) school size, d) attendance, and e) percentage of students qualifying for free and reduced meals?

2. Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

3. Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

4. Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

5. Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

Definitions of Terms

1. *Acceptable Performance* - Performance that meets the standards set by a particular state.
2. *Accountability* - Accountability refers to a system of checks and balances to guarantee appropriate outcomes. Educational accountability holds schools, teachers, and students responsible for their performance. School accountability refers to a state's holding the school responsible for the performance of its students. If adequate school performance do not occur, actions by the state could include ranking the school, assigning the school to a low-performing list, or removing administrative staff. Student accountability might include grade retention or withholding a high school diploma (Morgan, Moore, Detch, & Walton, 2007).
3. *Adequate Yearly Progress* (AYP) - Benchmarks developed by states to measure learning progress. Schools were held accountable for the progress of all subgroups (U. S. Department of Education, 2007).
4. *Criterion-referencing* - Compares individual student scores to the standards set by a particular state. Criterion-referenced tests measure whether students can perform a task or show knowledge at a predetermined level of proficiency the state has set as a criterion.
5. *Extrinsic* - Motivation that comes from the expectation of outside rewards.
6. *Intrinsic* - Motivation that comes from within the individual.
7. *Standardized test* - A test administered according to standardized procedures that assesses a student's achievement by comparison with a standard.
8. *Standards* - Preset levels of quality accepted as norms.

Overview of the Study

This study report is divided into 5 chapters. Chapter 1 provided an introduction to the study, statement of the problem, significance of the study, limitations and delimitations, research questions, and definitions of terms used in the study. Chapter 2 contains a review of literature associated with the issues addressed in the study. Chapter 3 focuses on the research methodology and design including data collection and data analysis. The results are discussed in Chapter 4. Chapter 5 presents a summary of finding, conclusions, and recommendations for further research on the subject matter.

CHAPTER 2

REVIEW OF THE LITERATURE

There has been a great deal of debate on the use of standardized tests and their effects on the quality of teaching in our schools. Due to the increased pressure on schools to improve test scores, they often implement quick fix measures to satisfy the public's hunger for accountability. Standardized test results provide all stakeholders with results on basic and cognitive skills, but many educational professionals have noted that such data are not good indicators for evaluating the quality of students, their teachers, or their schools. Harrison and McAfee (2002) describe such a test as a single measure for assessing students, teachers, and schools.

Data are collected from two types of standardized tests, norm-referenced and criterion-referenced. Norm-referenced tests compare individual performance with the entire student population of that particular grade level. Criterion-referenced tests are linked to individual state education standards and goals. Unlike on norm-referenced tests, all students could possibly reach a very high standard on a criterion-referenced test because they are not being compared to a reference group but judged only on their knowledge of the content being tested (Zucker, 2003). The Tennessee TCAP achievement test is a criterion-referenced test and assessment results are reported according to a level of performance such as not proficient, proficient, or advanced (Tennessee Department of Education, 2007c). However, data collected through both types of tests cannot provide policymakers with a true indication of what takes place in a classroom (Lynd, 2000).

Flippo and Riccards (2000) reported that educational reform efforts were at the top of the nation's agenda and that policymakers were hearing calls from members of the public to improve standardized test scores, which became an important issue. Since the 1980s, this push has

resulted in the view that standardized test results equal to and are equated with student and teacher performance (Flippo & Riccards). Centering standards on instructional leadership was a positive development, but standards could not reveal what kinds of teaching and learning should take place and what instructional leadership should look like (Anderson, 2002). Modern policies have emphasized the implementation of standards and accountability with little regard for the emotional consequences to students. Anderson argued that the pressures of standardized testing can create anxiety for students, thus, causing teachers to deal with the emotions of their *stressed out, over tested* students, which, in turn, creates additional concern and discomfort in the teacher.

According to Gratz (2000), there are two primary purposes for standardized testing. The first is economic. Gratz stated that the American people fear falling behind other countries in the international game of economic achievement; thus, they push students to learn more at a faster pace as part of a preparation plan. Gratz also noted that this is not a fair game because Americans are not prepared academically and are consequently falling further behind in the international race. Therefore, the economy suffers and America loses jobs to other countries. According to the U.S. Secretary of Education, Margaret Spellings (2007), the nation would lose its economic and competitive advantage if schools did not improve. On the other hand, Bracey (2003) refuted the accusation that America was falling behind other countries by presenting information provided by the Global Competiveness Report 2001-2002, in which the Geneva-based World Economic Forum (WEF) ranked 75 nations on their global competitiveness. In that report, the United States ranked second, behind only Finland. In the 2002-2003 report, ranking 80 countries, the United States ranked first in global competitiveness, growth competitiveness, and microeconomic competitiveness.

Gratz (2000) argued that the second purpose of standardized tests address the disparity between high and low-achieving students. This creates a disadvantage for all students who might not be challenged to meet higher expectations. Ediger (2003) supported this statement by noting that school systems, schools, and teachers often narrow their curriculum to include only material that will be on the test. Students could be missing valuable information and not performing to their highest capabilities because of the pressures of standardized testing.

History of Standardized Testing

"While there are many purposes that can be served by assessment, they are all secondary to the improvement of student learning" (Young, 2006, p. 1). Standardized testing was around long before the implementation of *No Child Left Behind.* According to Young, standardized testing has a long history. The earliest recorded use of standardized testing was in 2200 BC in China. These exams tested primarily for rote memorization and were administered to those interested in being in the Civil Service, a highly respected occupation in China. In 1905, China embarked on a large education reform movement that led to the removal of standardized tests that featured rote memorization. Standardized testing emerged in Europe during the Middle Ages to solve problems with the oral exam system previously used. By 1803, many European universities required written exams as part of their entrance requirements, in addition to interviews and other measures (Young).

Standardized testing did not appear in the United States until the middle of the 19th century in the form of written exams, as opposed to the oral exams previously recorded. Harvard University implemented the first entrance exam in 1851. During World War I, the United States Army developed multiple-choice Intelligence Quotient (IQ) exams (Finder, 2004). Following

acceptance of standardized tests by the Army, testing gained popularity throughout the United States.

The Scholastic Aptitude Test (SAT) first appeared in 1941 and was supported by the idea that education gained through a 12-year American education could be measured by a single test (Berliner & Biddle, 1995). In 1957, with the Russian launch of *Sputnik*, the condition of education in the United States changed once again (Bracey, 2007) with more emphasis placed on science and math. Young (2006) stated that many began to question American educational programs and wondered why America was not first to launch a satellite. Bracey refuted the idea that America's education programs were to blame for not launching a satellite before Russia and declared that the United States could have beaten the Russians by over a year if then-President Dwight D. Eisenhower had chosen not to rush into launching a satellite.

In 1965, the *Elementary and Secondary Education* Act was passed as part of a larger movement that demanded accountability from America's public schools (Young, 2006). This Act required all schools that receive federal money to prove that they were accomplishing prescribed educational goals. The government also specified that acceptable educational goals would be measured by standardized tests. In the early 1980s, Bracey (2007) responded to a report that many of America's captains of industry had grown lazy in making steel and building cars, while the Germans, Koreans, and Japanese had zipped ahead. Immediately, the educational system was blamed once again for this so-called shortfall, which birthed a provocative document highly critical of American education entitled *A Nation at Risk.* That document heralded the call for accountability still prevalent in all facets of education. *A Nation at Risk* demanded higher performance standards, increased learning times, and increased teacher preparation. Berliner and Biddle (1995) reported that many of the claims made in *A Nation at Risk* were not supported by

evidence, that, in fact, American schools were not failing and the accusation that they were failing was driven by the government and industry groups that commissioned the publication.

TCAP/NAEP

The Tennessee Comprehensive Assessment Program (TCAP) test is a criterion-referenced test that was developed by CTB McGraw Hill. The TCAP Achievement Test is a timed, multiple choice, state-mandated exam administered annually in grades 3-8. TCAP tests basic skills and content application in reading, language arts, math, science, and social studies. The length of the test varies depending on grade level. The time limits are generous, which allows most students to complete the exam. The total time for all components of the comprehensive test battery is approximately 5 hours and tests are broken into segments for the students to take over several days (Tennessee Department of Education, 2007a). Administrators, teachers, parents, and students receive copies of the test results. The reports can be generated for districts, schools, and individual students (Tennessee Department of Education).

Even though the TCAP was the standardized test selected by the state, the National Assessment of Educational Progress (NAEP) is currently considered the *gold standard* of assessment (Pellegrino, 2007). Pellegrino cited two reasons for this title. First, NAEP has a history as a well designed and developed national indicator. Second, NAEP is viewed as a high-quality indicator of academic achievement. NAEP levels of rigor and validity surpass that of achievement tests developed by individual states (Pellegrino). The NAEP provides important information on what America's students know in various subject areas.

The No Child Left Behind Act of 2001 requires that all states receiving Title I grants participate in the NAEP assessments. In 2007, 978,368 Tennessee students took part in the NAEP assessment (USDE, 2008). Because NAEP is administered to only a sample of students in

each state, no district, school, or individual student reports are generated (USDE). The sampling ensures that the statewide results are valid indicators of how Tennessee students perform relative to other states and the nation. In 2007, the average scale score for eighth-grade students in Tennessee was 259, which was not significantly different from that of the nation's other public schools (261). Of the 50 states and other jurisdictions that participated in the 2007 eighth-grade assessment, students' average scale score in Tennessee was higher than those in 10 jurisdictions, not significantly different from those in 11 jurisdictions, and lower than those in 30 jurisdictions. The percentage of students in Tennessee who performed at or above the NAEP *Proficient* level was 26 in 2007. The percentage of students in Tennessee who performed at or above the NAEP *Basic* level was 71 in 2007 (USDE).

Education and Government Role

Education in America began with America's first settlers. The United States of America was founded in 1776 with 13 original colonies that became states. Individual rights were protected under the Constitution. The founders, led by Thomas Jefferson, were deeply influenced by the European Enlightenment. What they desired most was an educated and self-governing citizenry (Pangle & Pangle, 1993). Ten of the nation's founders were also founders of academic institutions. The course was set early in our history by such government actions as the Land Ordinance Act of 1785 and the Northwest Ordinance of 1787. In 1787, Rev. Manasseh Cutler successfully negotiated with Congress for the Ohio Company land purchase to set aside two square middles for a public university (Scott, 2006). Article three of the ordinance stated, "Religion, morality, and knowledge, being necessary to good government and the happiness of mankind, schools and the mean of education shall forever be encouraged" (Avalon Project of Yale Law School, 2005, p.1). These acts required that each township reserve a section of land to

support education. Even prior to these landmark ordinances, there was support for education and the belief that it was critical in the development of an emerging nation. The Massachusetts Law of 1647, more commonly known as the "Old Deluder Satan Law" required that: 1) Every town of 50 households appoint and pay a teacher of reading and writing, and 2) Every town of 100 households must provide a grammar school to prepare children for the university (Wyett, 1998). While this law was primarily aimed at providing religious instruction, it was also a noted landmark in providing publicly funded educational opportunities in Colonial America (Wyett).

Throughout the 1800s, education continued to develop and by mid-century, many states passed compulsory school attendance laws. Massachusetts was the first in 1852, followed by New York in 1853 (Thattai, 2001). The *Morrill Act* of 1862 provided more support for educational endeavors by providing additional federal financial support to state universities, thus resulting in more opportunities to attend college. According to Key (1996), the *Morrill Act* was one of the most significant pieces of federal education legislation.

Education flourished over the next 50 years as high schools were formed across the country. In 1922, Tennessee followed Massachusetts' example, becoming the first southern state to enact a compulsory school attendance law. All Tennessee children between the ages of 8 and 16 were required by law to attend school (*Tennessee Education Laws Annotated,* 2000). In 1947, the state's first retail sales tax was levied by the Tennessee State Legislature and 80% of the proceeds were allotted to the public schools. In 1953, all students in grades 1 through 12 were issued free textbooks (Tennessee Department of Education, 2007d).

The role of the United States government in education expanded after Congress passed the Elementary and Secondary Education Act (ESEA) of 1965 (USDE, 2007). In 1994, the federal government ordered a reauthorization of the ESEA. Thirty-six years after the

implementation of the Elementary and Secondary Education Act and just 3 days after taking office in January 2001, President George W. Bush emphasized his deep belief in public schools and announced *No Child Left Behind* (NCLB). *No Child Left Behind,* the reauthorization of ESEA, led to the most dramatic change in national school education since the inception of ESEA. President Bush described *No Child Left Behind* as his bipartisan education reform, "the cornerstone of my Administration" (Tennessee Department of Education, 2005 p.1).

The No Child Left Behind Act of 2001 was signed into law by President George W. Bush on January 8, 2002. With the passage of *No Child Left Behind*, legislators hoped to achieve the following: 1) greater accountability for results; 2) more flexibility for schools, school districts and states in how they used federal funds; 3) a wider range of educational choices for families from disadvantaged backgrounds; and 4) an emphasis on research-based teaching methods (Aldridge, 2003). As part of the Act, annual state report cards inform parents and communities about the progress being made by states and schools. Today, schools that fail to make adequate yearly progress (AYP) toward statewide proficiency goals are subject to corrective actions. Each year schools not meeting AYP are labeled as *not performing*. After being labeled as not performing for 2 years, sanctions that include loss of federal funding, termination of staff, and dissolution of school districts can be enforced. The theme of *No Child Left Behind* appeared to focus on the best interest of the student, but the legislation was plagued by controversy because of the lack of state, local, and federal funding (Rose & Gallup, 2003).

The Center on Education Policy (CEP) is an independent nonprofit research and advocacy organization has conducted comprehensive and continuous research on the effects of NCLB since 2004. CEP produces annual reports contained in the series, *From the Capital to the Classroom* (Jennings & Rentner, 2006). By surveying officials in all state departments of

education, conducting case studies of individual school districts and schools, administering questionnaires to a nationally representative sample of school districts, and generally monitoring the implementation of this national policy, the CEP compiles a multiyear review and analysis. Their latest report concluded the following major effects of NCLB on American education:

1. State and district officials reported that student achievement on state tests was rising, which was a cause for optimism.

2. Schools were spending more time on reading and math, sometimes at the expense of subjects not being tested.

3. Schools were paying much more attention to the alignment of curriculum and instruction and analyzing test score data much more closely.

4. Low-performing schools were undergoing makeovers rather than the most radical kinds of restructuring.

5. Schools and teachers made considerable progress in demonstrating that teachers met the law's academic qualifications, but many educators were skeptical this would really improve the quality of teaching.

6. Students were taking many more tests because of NCLB.

7. Schools were paying much more attention to achievement gaps and the learning needs of particular groups of students.

8. The percentage of schools on state *needs improvement* lists was steady but not growing.

9. The federal government was playing a bigger role in education.

10. NCLB requirements meant that state governments and school districts also had expanded roles in school operations but often without adequate federal funds to carry out their duties (Jennings & Rentner, 2006).

According to Secretary of Education Margaret Spellings (2007), the act is working. The most current blueprint for *No Child Left Behind* shows progress since the academic bar was raised with the implementation of *No Child Left Behind*. In 5 short years, *No Child Left Behind* changed the way Americans viewed education, evolving from an idea, to a law, to a new way of life. Spellings alluded to the fact that educators, public officials, and the media would not be engaged in conversation regarding higher educational standards without the passage of *No Child Left Behind.* However, in 2008, renewal of this law was under consideration by Congressional lawmakers. According to Sanchez (2008), many people remarked that the goals of *No Child Left Behind* were unrealistic and viewed it as an unworkable, heavy-handed mandate, claiming that many of the gains measured were, in fact, seen before the implementation of *No Child Left Behind.* Azzam (2007) cautioned that "everyone interested in NCLB needs to be very careful about reaching conclusions based on flawed or simplistic interpretations of data" (p. 92).

Attendance and Test Scores

Attendance has long been viewed as possessing a strong correlation to academic performance in the classroom (Atkinson, 2005; Moore, 2005) but little research has been conducted on the topic. Since the implementation of No Child Left Behind, research on the topic of attendance has increased. Atkinson attributed at least part of the increase to inclusion of student attendance as an indicator of adequate yearly progress (AYP). The NCLB legislation specified attendance as an *other academic indicator* for elementary and middle schools (USED, 2006).

According to Corville-Smith, Ryan, Adams, and Dalicandro (1998), school absenteeism is a multicausal problem and that to understand the behavior, variables from multiple domains warrant investigating. These domains can be categorized as personal characteristics of the absentee, family factors, and school factors. When looking at school factors, nonattendance appears to be related to strained and conflict-ridden teacher-pupil relations and school curriculum that does not stimulate student interest or seems unlikely to hold future benefits (Corville-Smith et al).

Several researchers linked student attendance to instructional teaching styles. In a study of North Carolina's physical science standardized tests, researchers found that inquiry-based instruction has led to a dramatic improvement in student attendance and participation and higher classroom grades earned by students (Tretter & Jones, 2003). Klem and Connell (2004) suggested that by creating a personalized student environment levels of engagement would increase, along with student attendance and test scores. Promoting active student engagement motivates students to attend school and remain there until graduation (National Research Council, 2003).

Reading and Standardized Testing

Every year, thousands of students prepare for and take standardized tests as a measure of student learning. Since the implementation of NCLB, which mandates testing of all students in grades 3-8 every year, many schools are targeted for *improvement* based on test results (Valencia & Buly, 2004). Post NCLB and with the rise in standardized testing, there has been growing concern that reading comprehension is being tested in other subjects such as science and math. Flick and Lederman (2002) argue that information in those areas was presented in paragraph form and test takers had to read and interpret separate items and meaningfully relate the

questions to earlier information. Ediger (2005) noted that science teachers need to be thoroughly grounded in science content and methodology and also be an instructor of reading. Teachers must understand the importance of excellent comprehension instruction in the primary grades and continue to refine them through the intermediate and middle grades (Jongsma, 2004). Being able to read, however, did not mean that the student is able to comprehend the material being read (Fleener & Bucher, 2003).

Valencia and Buly (2004) reported that although legislators, administrators, parents, and educators have been warned repeatedly not to rely on a single measure to make important instructional decisions, scores from state tests still drive school system administrators when looking for instructional programs to meet state standards. In their research, Valencia and Buly also found that many of the students failing standardized tests were the students typically found in a regular classroom who were experiencing reading difficulty but had not been identified as needing special education services or interventions. Fleener and Bucher (2003) suggested that classroom teachers, not reading specialists or special education teachers, are solely responsible for the reading instruction of these students and, ultimately, their achievement.

Per-Pupil Expenditure

Whether or not increasing expenditures for education results in student achievement has been an important issue for school policymakers (Lopus, 1990). Ever since the release of the *Coleman Report* in 1966, controversy has existed over the relationship between money and schooling (Bracey, 1993). Bracey reported that conservative critics rallied around the claim that there was no relationship between money and schooling outcomes. While the *Coleman Report* was attacked on numerous methodological grounds, the attackers themselves were unable to

establish that any relationship between school funding and student achievement existed. Test scores offer endless opportunities for misinterpretation (Bracey, 1996).

Over the years, many researchers have tried to figure out what variables contribute to higher standardized test scores. In 1983, former Secretary of Education Terrel Bell attributed the rise in SAT scores to the impact of *A Nation at Risk*, the report he commissioned that appeared in 1983 (Bracey, 1996). The problem with this assumption was the fact that the scores released in the fall of 1983 actually came from the tests taken in 1982. This meant that the rise in scores could not be attributed to the implications in *A Nation at Risk*. In 1989, Eric Hanushek of the University of Rochester supported the material released in the 1966 *Coleman Report*. Hanushek wrote in *Educational Researcher* that "there is no strong or systematic relationship between school expenditures and student performance" (p. 1) (as cited in Bracey, 1997). In modern education, economies of scale is becoming of increasing concern in the financing of school districts. In the past large school systems were able to purchase items at a lower cost. Eberts (2007) explored the role of teachers' unions in public education. One of his primary focuses addressed the way in which unions affect both student achievement and the cost of providing quality education. His findings suggested that although higher pay and fringe benefits increased educational costs, they also might have increased the quality of education if they attracted better teachers.

School Size

When seeking increased achievement, school districts need to research many variables including school size. As previously mentioned, economy of scale was becoming of more concern in the financing of school districts. The relationship between school size and efficiency depend on whether the schools were relatively small or relatively large (Butler & Monk, 1985).

Increasing school size might create more economic efficiency, but when schools become too large, they tend to become economically inefficient (Alspaugh, 1994). Gough (2002) provided advantages and disadvantages of small and large schools. He suggested small schools offer more opportunities for leadership in extracurricular activities, teachers knew the students well, and teachers developed relationships with the families of their students. He found that a major disadvantage of small schools was that they were often not able to provide students with the opportunity to experience various teaching methods; students might have the same teacher for several classes. Large schools could provide students with access to advanced courses and various teaching methods; however, they were not as personalized as small schools and leadership opportunities were much more competitive (Gough).

Tennessee's Project STAR, the only large-scale, controlled study of the effects of reduced class size, was conducted in 79 elementary schools in the state of Tennessee from 1985 to 1989. In all, about 7,500 pupils in more than 300 classrooms participated in the 4-year longitudinal study. Teachers and students were randomly assigned to one of three groups: small class (13 to 17 students per teacher), regular class (22 to 25 students per teacher), and regular-with-aide class (22 to 25 students with a full-time teacher's aide). The intervention began as the students entered kindergarten and continued through third grade (Health, 2003). The design drew heavily upon previous research findings, namely, that any benefits of small classes are likely to be realized in the primary grades, that there may be different outcomes for students based on race or economic disadvantage, and that only substantial reductions in class size are likely to have noteworthy impact (Illig, 1996). The results of Project STAR were definitive and noted:

- Children in small classes consistently outperformed children in large classes.

- Low-income students (determined by eligibility for free and reduced price lunch program) did less well than other children did regardless of location or class size.
- Children in smaller classes were less likely to be retained.
- Students in smaller classes were more likely to participate in subsequent classes.
- At each grade level (K-3) and across all school locations, students in the smaller classes made the highest scores on the norm-referenced Stanford Achievement Test (SAT) and the criterion-referenced Basic Skills First Test (BSF).
- The only consistent positive regular-aide class effect occurred in first grade.
- Teachers reported that small classes helped them better identify student needs, provide more individual attention, and cover more material effectively.
- Students who were in STAR small classes were less likely to be retained and were less likely to drop out of school (Health; Illing).

It is vital to the economy to keep children in school learning on (or above) grade levels. The findings of Project Star showed numerous short- and long-term benefits associated with class-size reduction. It was not inexpensive to implement small-class sizes; however, the cost of not reducing class sizes would likely result in greater social and economic expenditures in the long run (Pate-Bain, Fulton, & Boyd-Zaharias, 1999).

Socioeconomic Status

Students from a low socioeconomic (SES) background constitute the largest population of individuals considered at-risk of not graduating from high school. They often have had increased economic, legal, and psychological problems (Caldwell & Ginthier, 1996). Numerous reasons have been cited for students dropping out of school. Dunn and Griggs (1988) generalized these reasons into four common categories: familial factors, personal characteristics,

socioeconomic factors, and educational achievement and school behaviors. Educators had the opportunity to alter the educational achievement and school behaviors category. While low socioeconomic status was strongly correlated with low achievement, some low socioeconomic students were academically successful (Caldwell & Ginthier). Segelken (2007) stated that a major difference in middle-class and poor students was that middle-class students could slog through with effective teaching, while poorer students in under-resourced schools lacked the necessary scaffolding to grasp subject matter and they opted out.

Caldwell and Ginthier (1996) reported that the critical differences between low and high achieving students were internal variables related to motivation. Low motivation was critical in student achievement, especially for the low socioeconomic student. Encouraging students through instructional methods and strategies to promote active participation allowed students to take ownership of their learning and enabled them to develop autonomy, thereby, increasing motivation and achievement (Caldwell & Ginthier).

Test Anxiety and Pressures

As test anxiety and pressures to improve test scores mounted, high-stakes tests dictated the scope and direction of instructional programs in schools across the nation (Donegan & Trepanier-Street, 1998). Donegan and Trepanier-Street surveyed teachers and students at both elementary and secondary levels. Among the elementary teachers, 73% reported *occasionally to consistently* experiencing personal stress. Ninety-two point nine percent of secondary students and 68.1% of elementary students reported anxiety due to standardized testing. The researchers concluded that pressures to improve test scores caused a narrowing of the curriculum to those areas and skills covered on the tests and resulted in significant class time being used to teach test-taking skills (Ediger, 2003).

Perreault (2000) conducted a study in which teachers were interviewed regarding the impact that state-mandated testing had on their classes. They reported that teachers noted that their instructional programs were severely restricted after they were ordered not to teach anything that was not on the test.

Accountability and Acceptable Performance

What does it mean to say a school is doing well? Accountability clearly presents an enormous challenge because it requires the cooperation and collaboration of all stakeholders. A site-based decision-making committee (SBDM) could aid in this collaboration. This committee improved student performance through widespread participation of all stakeholder groups in a process of bottom-up decision-making (Wyman, 2000). Standards and the measurement of performance were developed with the intention of cleaning up a messy system and holding teachers and school administrators accountable for student performance. The modern aim is the same as it has always been: to systematize and standardize so that the public knows which schools are performing well and which are not. Because of these standards, there have been payments and penalties for performance (Eisner, 2001). As previously stated, school systems that perform poorly on standardized tests have been subject to many consequences. Popham (2001) concluded that standardized tests should never be used to evaluate teachers or schools. He also added that classroom instructional decisions such as the decision to promote or retain a student should never be based solely on any single test or test result. Suggestions offered by Holloway (2002) to overcome the negative effects of high-stakes testing included the following: 1) avoiding use of a single measure of student performance; 2) avoiding comparisons between schools; 3) recognizing and reporting the degree of uncertainty of tests results; and 4) ensuring that those who made decisions were proficient in interpreting test results.

Teachers experience pressure when it comes to standardized testing. Kohn (2001) argued that teachers are an easy target when it comes time to blame someone for inadequate test scores. Kohn offered teachers the following suggestions concerning standardized testing:

1. Do what is necessary to prepare your students for the test, and then get back to the *real learning*. Never forget the difference between these two objectives.
2. Do not do any more test preparation than is absolutely necessary.
3. Make sure that whatever time you spend on test preparation is creative and worthwhile.
4. One of your primary obligations as a teacher is to absorb as much pressure as you can from those above you without passing it on to those below you (Kohn).

Alternatives to Standardized Testing

Traditional standardized testing methods have been criticized for failure to assess student learning. As a counteraction, many educators use various forms of authentic assessments to supplement traditional forms of standardized tests. Alternative assessments are used in many schools throughout the United States and other industrialized countries. Chen and Martin (2000) reported that traditional forms of assessment did not adequately represent the activities children typically undertook in the classroom. They also did not support current theories of learning and cognitive development. Alternative assessments require thinking about the purposes of knowledge and what constitutes knowledge. Alternative assessments allowed many other voices, perspectives, and actions to be viewed when measuring accountability (*Standardized Testing*, 2002). For an education assessment tool to be useful, it has to clarify what it measures and measure what is most valued, improved student learning (Gandal & McGiffert, 2003).

Portfolio-based assessment has been identified as one alternative to standardized testing (*Standardized Testing*, 2002). Portfolio-based assessment can be coupled with traditional testing methods to provide a more realistic view of student abilities. Portfolio-based assessment is a purposeful and systematic collection of student work. During the school year, teachers and students gather work that displays student progress and achievement in various subjects. Portfolio-based assessments vary considerably but they all incorporate records kept by the teacher and collections of the students' work. Several advantages to this classroom-based approach have been identified, including encouraging student reflection. Reflection helps students think about what they have learned, develop their own individual learning style, and contribute to the overall goal of improving student learning. Portfolio-based assessment also provides an evaluation based on a wide range of student work completed over a long period of time rather than the snapshot of a single test (Chen & Martin, 2000; *Standardized Testing*). A criticism of this approach was that it worked best with quality teachers. Teachers must take the time to collaborate with the students, parents, and other teachers to ensure student success. Another potential problem that has been identified with portfolios was the logistics of managing the portfolios (*Standardized Testing)*.

Performance exams given to all students provide another alternative to standardized testing. The student completes a certain task, such as writing an essay, conducting an experiment, or giving an oral presentation. Performance assessment requires students to demonstrate their knowledge and skills in response to authentic activities (Chen & Martin, 2000). One advantage performance exams have over standardized tests was that they drive the curriculum. Students are able to personalize the testing and reveal their content knowledge of a particular topic in a less stressful manner. Performance exams can measure skills such as

integration of content knowledge across subject areas, which paper and pencil tests can not. Some disadvantages of performance exams are that they take time, expertise, and a great deal of money (*Standardized Testing*, 2002).

An assessment known as proficiency exit standards combines portfolio-based assessment, performance exams and, sometimes, standardized tests. Proficiency standards focused on four broad areas: math, science, communication, and research. They have been generally considered more rigorous that most standardized tests (*Standardized Testing*, 2002).

Exhibitions of student work such as the science fair offer another assessment tool. Strengths in using exhibitions are that they provide a hands-on approach to learning and ways for all students to succeed independently. A major problem with the use of exhibitions is the difficulty in determining who completes the project because parents often help their children complete the assignment (*Standardized Testing,* 2002).

The Use and Misuse of Standardized Testing

Standardized tests provide school systems with valuable data that can be essential to school and student improvement. However, the same test data are often misused for teacher accountability (Holloway, 2001).

When used to improve student achievement, tests can be vital in education reform. They provide teachers with individual levels of student achievement, making it easier to offer students appropriate levels of instruction. In the medical field, tests are used to diagnose illnesses and provide proper treatment. The same procedure can be used in education. Tests can aid educators in identifying academic problem areas. Students should meet academic benchmarks as they mature and standardized testing provides schools with data that compare their students to others on a local and national level (Gandal & McGiffert, 2003). Unfortunately,

standardized tests are frequently misused. Specifically, standardized tests that are designed to draw national comparisons among students are frequently used to evaluate teachers and schools. This is an unfair assessment because there are rarely any references to a particular school's curriculum or content standards. Thompson (2001) concluded that these assessment tools were not necessarily bad, they were simply ill suited to the task.

Strategies Used in Standardized Testing Preparation

Test preparation strategies vary in many respects. Teaching students to become better test takers is a frequent approach with some advantages; however, the disadvantages include stressing drills on subject matter over higher levels of cognition (McColskey & McMunn, 2000).

Many teachers revert to quick, short-term fixes to boost their students' test scores. McColskey and McMunn (2000) suggested that a school might adopt a variety of *quick-fix* strategies to increase test scores including: reduced emphasis on non-tested material or subjects; reduced emphasis on projects or performance assessments that did not align with the kinds of items on the state tests; and arrange for teachers to attend professional workshops and in-services that focused strictly on the improvement of test scores. Teachers, schools, and school districts need to be cautioned against falling victim to these quick-fix methods (McColskey & McMunn).

Student, Parent, and Teacher Concerns Regarding Standardized Testing

As the school year begins, so does the preparation for the standardized tests that will come during the second semester. Teachers prepare their lessons to cover all material that will be tested. Teachers often experience a feeling of powerlessness and anger at having to cope with a system that rewards superficial thinking and relies on inappropriate measurement (Popham, 1999). Tests are inadequate for measuring many types of student growth. They are also misused for accountability purposes and fail to recognize the multifaceted nature of good teaching

practice. Teachers see first hand how their students are affected by high-stakes testing. They note the pressures their students face along with the mental challenges of taking the test (Cruickshank & Haefele, 2001).

Teachers are also concerned with the inability of a test to be nondiscriminatory. A student's low socioeconomic status could greatly affect his or her test results. For example, it would be inappropriate to compare a child who must sleep in an unheated home attempting to take a test on an empty stomach to a child arriving in a luxury automobile, well fed and groomed (Gibson, 2001).

In 2000, a group of secondary school teachers from British Columbia began working together to reform their own teaching practices and bring new life to their curriculum (Wassermann, 2001). One of their major concerns was assessment. Prior to beginning work on the assessment tools, they set out to identify the educational principles on which the evaluation process rested. The assessment tools they evaluated reflected Gardner's Theory of Multiple Intelligences. In his book, *Multiple Intelligences: The Theory in Practice*, Gardner (1993) proposed seven dimensions of intelligence:

1. Linguistic intelligence: the ability to understand words and how they are combined to produce useful languages. This is important for writers, poets, and journalist.
2. Logical-mathematical intelligence: the ability to see patterns, order, and relationships in seemingly unrelated events in the world around us and to engage in logical chains of reasoning. Scientists, mathematicians, and engineers come to mind when you think of this intelligence.

3. Musical intelligence: the ability to discern pitch, melody, tone, rhythm, and other qualities of musical symbolism and integrate them into intellectual activity such as reasoning. Musicians, singers, and composers fit in this category.
4. Spatial intelligence: the ability to accurately perceive and think in terms of the visual qualities of the world and its dimensions and to manipulate and transform them in creative ways. Architects, artists, and photographers fall under this intelligence.
5. Bodily-kinesthetic intelligence: the ability to control one's bodily motions, the capacity to handle objects skillfully, and the skill to combine these into a language with which to express oneself. This is important for athletes, dancers, and figure skaters.
6. Intrapersonal intelligence: the ability to access and understand the inner self. This refers to self-awareness. Individuals are comfortable with their feelings, reactions, and aspirations. They are comfortable with their personal emotions and is able to differentiate between various feeling and apply them when thinking about the world. Playwrights and writers are examples.
7. Interpersonal intelligence: the ability to notice and make distinctions among individuals. These individuals are able to read the moods, temperaments, motivations, and intentions of other individuals. Individuals possessing high levels of interpersonal intelligence find it useful in educational leadership.

(p. 239)

In 1999, Gardner introduced an eighth intelligence, naturalist. Naturalist intelligence includes appreciation of and sensitivity to the natural world. People with a highly developed naturalist intelligence are adept at classifying plants and animals (Rubado, 2002).

Applying Gardner's theory to assessment enabled students to show what they had learned by allowing them to tap into their own unique talents and skills. Teachers found that, by assessing students in this manner they were able to gain considerable data about student performance. In the end, the teachers decided that assessment tools should:

- provide teachers with information about the extent of students' understanding of complex concepts;
- give students some choices in their evaluation options, allowing them to reveal their areas of strengths;
- call for careful preparation, gathering, and application of concepts of an extended period of study;
- give students time to rethink and rework their material, to reinforce the material taught;
- emphasize success, rather than focus on failure;
- give teachers more information over a period of time, instead of relying on a single test;
- allow teachers to gather information about each individual students' strengths and weaknesses, so that they can provide individual help to individual learners;
- allow teachers to individualize their students learning needs; and
- emphasize self-evaluation as an important means of promoting students' critical awareness of themselves as learners (Wassermann, 2001).

Many parents are concerned with the emphasis placed on standardized testing. Carol Holst, a Texas mother-turned activist and leader of the Parents United to Reform TAAS (Texas

Assessment of Academic Skills), claimed her fourth-grade son was so worried about testing that he could not sleep. She pointed out that her son's school had no science or social studies curriculum because those subjects were not covered on the TAAS (Ohanian, 2001). A group of parents in Gwinnett County, Georgia, took the fourth-grade practice test and scored miserably. They questioned the test and wanted to know why well-educated adults performed so poorly on the test. They formed the Concerned Parents of Gwinnett County to fight the tests. The parents are now under investigation by the Gwinnett County police who wanted to know how the parents received the test. The parents refused to provide any information (Ohanian).

Additional Problems Associated with Standardized Testing

Additional problems are associated with standardized testing. Principals and teachers are constantly under pressure to make sure their students perform well on tests, which often causes schools to emphasize drill and practice activities. Administrators may purchase test preparation materials with funds that would be better allocated to other areas of instruction. The fact that these techniques are used to prepare students to take the test undermines the validity of the scores. Administrators, teachers, and students might be compelled to cheat due to the tremendous pressure to succeed (Amrein & Berliner, 2003).

High-stakes testing causes many other problems for schools as well. One of the most significant is that districts narrow the curriculum they offer to allow more time for instruction in areas that will be tested. Even within the particular curriculum areas that are tested, many subareas may be dropped if they are unlikely to appear on the test. This deprives students of reaching their full potential due to the narrowed curriculum (Ediger, 2003).

In modern society, there seems to be the general consensus that if it moves – test

it. Edward Fiske, former editor of the *New York Times,* criticized standardized tests for several reasons, as follows:

> They assume a single, correct answer to problems. They don't allow for complex answers or multiple ways of arriving at them.
>
> Standardized tests recognize how good students are at recognizing information, not at how well they can generate information, synthesize information, solve problems, or think independently.
>
> Because the majority of tests are timed, more value is placed on thinking quickly than on thinking profoundly.
>
> Standardized tests focus mostly on basic skills and do not focus on higher-order thinking skills.
>
> Standardized tests emphasize isolated learning not the integration of facts and ideas (as cited in Wassermann, 2001).

Tests are altered regularly making the collected data very unreliable. Too many important decisions are based on the results of poorly designed and inappropriately used instruments of measure. Everyone involved in the education process suffers from the anxiety and the consequences of a narrowed curriculum brought about by standardized tests (Fraenkel & Wallen, 2003).

Ways in which to accommodate students with special needs presents another serious problem faced by high-stakes testing. Parents of students with disabilities are concerned with the difficulty that their children will almost surely have in qualifying for a diploma that requires them to pass tests that address certain skills beyond their abilities. There are at least two major risks to be considered when dealing with students with special needs and the emphasis placed on

high-stakes exit exams. One of these is that special education students might drop out of school because they see no reward in staying. Thus, they might remove themselves from school and miss opportunities to learn that are not measured by a test. A second major risk posed by high-stakes testing is that teachers at the elementary level are required to administer "benchmark exams," another factor of leading to teaching to the test. Many might fear that the inclusion of students with special needs would reduce classroom scores (McDermott & McDermott, 2002).

Meeting the needs of students who are not fully proficient in the English language creates another major problem to be addressed when viewing standardized testing. Some dual-language tests were devised to alleviate this problem, but results showed that structural differences across the two versions still remained (Khaliq & Sireci, 2002). A study that investigated the way in which standardized testing methods impacted learners differently across varying language proficiency levels confirmed that performance anxiety varied inversely with language proficiency. Environmental pressures were also found to impact learning unevenly across the language spectrum (Rotenberg, 2002).

Neill (2003) concluded that high-stakes testing was not the answer to our academic shortfalls. Simply demanding higher scores from students would not fix the country's academic problems. Research conducted since the early 1990s has shown that high-stakes testing has not produced improvements in educational outcomes, while states that do not have high-stakes graduation exams were more likely to improve their average scores on the National Assessment of Educational Progress. They are also more likely to show improvement on college entrance exams such as the ACT.

Student Motivation

Students are faced with many pressures at school, but testing is likely near the top of the list of such pressures. From the time students enter school until the time they graduate, they are bombarded with tests. Amrein and Berliner (2003) suggested that student learning and motivation did not improve as a consequence of tests. Many federal legislators recommend attaching rewards and consequences to rigorous high-stakes testing to *motivate the unmotivated* to learn. The *unmotivated* are usually identified as the low socioeconomic students in urban schools (Amrein & Berliner, 2003). Yet, other research revealed that applying sanctions and rewards to tests not only affects student performance, but that students became less intrinsically motivated to learn and less likely to engage in critical thinking. Research found that high-stakes testing caused teachers to take greater control of the learning experiences of their students, denying their students the opportunities to direct their own learning. Teachers no longer encourage students to be creative in their thinking; they teach their students how to take tests (McColskey & McMunn, 2000). Attaching stakes to tests alienated students from their own learning experiences. The assumption that high-stakes testing motivates students is seriously flawed. In fact, such tests have decreased student motivation and led to lower student retention and higher dropout rates (Amrein & Berliner, 2003).

As students aged, their dislike of testing grows (Amrein & Berliner, 2003), which causes more teenagers to exit traditional, formal high school to earn a General Educational Development (GED) credential. Attaching stakes to tests apparently obstructed students' paths to becoming lifelong, self-directed learners and alienated them from their own learning experiences in school (Shelton & Biddle, 2000).

There was disagreement about which type of motivation, intrinsic or extrinsic, worked best for students. Ediger (2001a) found that the teacher plays a key role in the classroom with extrinsic motivation but that intrinsic motivation must also exist for optimal learning to take place. He also stated that every classroom is different and that teachers should experiment with both intrinsic and extrinsic motivation to decide which approach or combination of intrinsic and extrinsic works best with their particular group of students (Ediger, 2001c).

During test preparation, teachers are forced to allot regular instruction time to practice tests. Repeated practice of basic material tends to frustrate students and causes added stress before students engage in the actual test. The drill and practice routine becomes annoying both for those who have mastered the basics and for those who have not. Students become well aware of the test and its implications during this process. Anxiety levels rise in students as they feel the tests may keep them from doing well or possibly even from graduating from high school. Some students feel that it is unfair that 12 years of work may be jeopardized depending on the results of one test (Hughes & Bailey, 2001).

Good teacher and parent relations are essential for students to reach their full potential. Promoting parent participation in their children's schooling has several advantages including an increase in test scores (Ediger, 2001b). A good relationship between parents and teachers helps children learn better at home and in school and provides teachers with the practical, fund-raising, and emotional support they sorely need. However, in many classrooms parents are viewed as more of a hindrance than a help. When this is the case, everyone involved suffers. Teachers should be open to parent interests; closing the door and trying to control interaction with parents is in the worst interest of the students and teachers. Students are more motivated in both the classroom and at home when good teacher and parent relations exist (Hargreaves, 2001).

Summary

The passage of No Child Left Behind in 2002 has resulted in significant changes to the American education system. Schools, school systems, and states could lose federal funding if they are identified as not making AYP. Tennessee, like many other states, uses criterion-referenced tests to meet the student assessment requirements of NCLB. Increased accountability as a result of NCLB results in greater emphasis being placed on standardized test scores and creates stressful challenges for educators across the country. Standardized testing could provide the state, school systems, schools, and individual teachers with valuable data essential to school and student improvement when used properly. Unfortunately, standardized tests are frequently misused. Educators need proper training on how to interpret standardized test data and use it for its intended purpose, ensuring a successful educational experience for all students.

CHAPTER 3

METHODS AND PROCEDURES

Introduction

The purpose of this study was to determine whether Tennessee Comprehensive Assessment Program (TCAP) Achievement reading scores, school district per-pupil expenditures, school size, percentage of students qualifying for free and reduced-priced meals, and attendance were related to science TCAP test scores from the 2007-2008 school year. The data were gathered from an analysis of mean standardized test scores in reading and science of 8th graders in 67 school systems comprising 181 schools located throughout Tennessee.

Chapter 3 describes the methodology and procedures that were used in this study. The chapter is organized into the following sections: research design, population, instrumentation, procedures, data analysis, and summary.

Research Design

For over a decade, Tennessee has released a State Report Card containing test scores and demographics. This study used data collected from the Tennessee 2007-2008 TCAP test.

This quantitative study was designed to determine the correlation of 8th grade mean TCAP reading scores, attendance, school size, percent of students qualifying for free and reduced-priced meals, and per-pupil expenditure with 8th grade mean science TCAP test scores.

Population

The population included 181 schools located in 67 Tennessee school systems. The 67 Tennessee school systems used in this study included: Memphis, Shelby, Knox, Rutherford, Beldsoe, Sullivan, Hardeman, Williamson, Giles, Hamilton, Benton, Blount, Morgan, Dickenson, Roane, Polk, Union, Stewart, Kingsport, Sevier, Rhea, Wilson, Tipton, Franklin, Dyer, Oneida, Bradley, Henry, Lauderdale, Lenoir, Lewis, Lexington, Macon, Monroe, Weakley, Humphrey's, Meigs, Warren, White, Paris, Elizabethton, Greene, Hawkins, Cleveland, Anderson, Coffee, Crockett, DeKalb, Dyersburg, Hickman, Tullahoma, Hamblen, Loudon, Robertson, Greeneville, Union, Hardin, Bedford, Houston, Humboldt, Campbell, Jefferson, Trousdale, Sumner, Davidson, Montgomery, and Maury. The population was limited to those public schools throughout Tennessee that were configured to provided public education services to sixth, seventh, and eighth grade students. This number was comprised of 181 schools located throughout the 67 school systems listed above.

Procedures

Approval to conduct the study was obtained from the Institutional Review Board (IRB) at East Tennessee State University. The Tennessee State Department of Education published a report card for each public school district and school that was accessed from the state web page (Tennessee Department of Education, 2007b).

After approval was granted, the 2008 school report cards were printed from the Tennessee State Department of Education website for each of the schools studied. The schools that met the criteria for the study were selected and categorized by their grade level. The information was then coded. The mean 8th grade reading and science scores and NCLB

information were obtained from the report cards. The information was entered into the Statistical Program for the Social Sciences (SPSS) software. The data were analyzed with a Pearson *r* correlation. Independent-samples *t* test were used to answer the last 4 research questions. The statistics were analyzed to determine if the null hypotheses for each of the 5 research questions should be rejected.

The data were obtained from the state report card. Each year, the Tennessee State Department of Education publishes the report card for each public school and district. The report card serves as a means of accountability for the state, enacted by the passage of the state's Education Improvement Act of 1992.

Instrumentation and Data Collection

The instrument used in this study was part of the Tennessee Comprehensive Assessment Program (TCAP) test administered to each student as a requirement during the 2007-2008 school year. The TCAP test is a criterion referenced test developed by CTB McGraw Hill. The data used in this study were readily available for review on the Tennessee Department of Education website (Tennessee Department of Education, 2007a).

The TCAP Achievement Test is a timed, multiple choice, state-mandated exam administered in grades 3-8 that tests basic skills and content application in reading, language arts, math, science, and social studies. The length of the tests varies depending on grade level. The time limits are generous and allow most students time to finish. The total time for all components of the comprehensive test battery was approximately 5 hours, and tests were broken into segments for the students to take over several days (Tennessee Department of Education, 2007a).

The test answers were machine scored. Results were reported to parents, teachers, and administrators and provided information on how well students performed on the content areas being tested (Tennessee Department of Education, 2007a).

Research Questions

The following research questions and null hypotheses guided the study's analyses.

Research Question 1

Is there a relationship between mean science TCAP scores of 8th graders in Tennessee and the following variables: a) mean reading TCAP scores, b) per-pupil expenditure, c) school size, d) attendance, and e) percentage of students qualifying for free and reduced-priced meals?

Ho: 1_1

There is no relationship between mean reading TCAP scores and mean science TCAP scores of 8th graders in Tennessee.

Ho: 1_2

There is no relationship between per-pupil expenditure and mean science TCAP scores of 8th graders in Tennessee.

Ho: 1_3

There is no relationship between school size and mean science TCAP scores of 8th graders in Tennessee.

Ho: 1_4

There is no relationship between attendance mean science TCAP scores of 8th graders in Tennessee.

Ho: 1_5

There is no relationship between percentage of students qualifying for free and reduced-priced meals and mean science TCAP scores of 8th graders in Tennessee.

To answer this research question, a Pearson's correlation coefficient at an alpha level of .05 was conducted to determine the strength and direction of the relationship between mean science TCAP scores and the following variables: mean reading TCAP scores of 8th grade students in Tennessee, per pupil expenditure, school size, attendance, and percent free and reduced-priced meals. Using the Bonferroni approach to control for Type I error across the correlations, a *p* value of less than .01 (.05 / 6 = .008) was required for significance.

Research Question 2

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

Ho: 2

There is no difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345.

To answer this question an independent-samples *t* test was conducted to evaluate whether or not schools with above average per-pupil expenditures scored better on the science TCAP test than the schools with below average per-pupil expenditures.

Research Question 3

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

Ho: 3

There is no differnce between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345.

To answer this question an independent-samples *t* test was conducted to evaluate whether or not schools with above average per-pupil expenditures scored better on the reading TCAP than the schools with below average per-pupil expenditures.

Research Question 4

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

Ho: 4

There is no difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentages above or below the state goal of 93%.

To answer this question an independent-samples *t* test was conducted to evaluate whether or not schools reporting attendance above 93% scored better on science TCAP than schools reporting attendance below 93%.

Research Question 5

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

Ho: 5

There is no difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentages above or below the state goal of 93%.

To answer this question an independent-samples *t* test was conducted to evaluate whether or not schools reporting attendance above 93% scored better on reading TCAP than schools reporting attendance below 93%.

Data Analysis

The Statistical Program for the Social Sciences (SPSS) was used to analyze the data using descriptive and inferential statistics. Data for each group being studied were collected and organized into a data file. Means were calculated for each group. The statistical results were used to decide whether the null hypothesis should be rejected or retained.

Summary

The research design, methodology, and procedures used in this study were presented in Chapter 3. The population and selection procedures were described. The study used quantitative procedures to determine if a correlation existed between mean science TCAP scores of 8th graders in Tennessee and the following variables: mean reading scores, attendance, per-pupil expenditure, and percent free and reduced-priced meals. The population used for this study consisted of 181 schools located in 67 Tennessee school systems. The study consisted of 5 research questions with 9 null hypotheses. An analysis of the data is provided in Chapter 4. Chapter 5 presents a summary of findings, conclusion, and recommendations for further research on the subject.

CHAPTER 4

ANALYSIS OF DATA

After the passage of No Child Left Behind legislation in 2002, many states began using criterion-referenced tests to meet the student assessment requirements of the law. In Tennessee, students in grades 3-8 participated in the Tennessee Comprehensive Assessment Program (TCAP). The increased level of accountability and subsequent focus on standardized test scores was the impetus for the current study. The analysis of data concentrated on school characteristics most closely associated with science TCAP scores.

The purpose of this study was to determine whether Tennessee Comprehensive Assessment Program (TCAP) Achievement reading scores, school district per-pupil expenditures, school size, percentage of students qualifying for free and reduced-priced meals, and attendance were related to science TCAP test scores from the 2007-2008 school year. The data were gathered from an analysis of mean standardized test scores in reading and science for 8th graders in 67 school systems comprising 181 schools located throughout Tennessee. One hundred eighty-one schools in Tennessee are configured grades six through eight. However, only 177 schools had attendance numbers available on the Tennessee Department of Education website.

The 5 research questions presented in Chapter 1 were used to guide the study. The 9 hypotheses presented in Chapter 3 were used to test the data. Discussion of the findings for each follows. Descriptive statistics used in this study are provided in Table 1.

Table 1

Descriptive Statistics Used in the Study

	N	Minimum	Maximum	M	SD
8th grade mean reading NCE	181	35.90	77.70	59.38	7.25
% Free and Reduced-Price meal	181	1.00	98.00	50.94	24.04
% Attendance	177	80.60	97.10	94.52	1.60
Per-Pupil Expenditure	181	$6,538.00	$10,355.00	$8,285.02	$1,033.38
# of Students	181	53.00	1,342.00	652.17	269.25
8th grade mean science NCE	181	23.60	77.10	53.33	10.00

Research Question 1

Is there a relationship between mean science TCAP scores of 8th graders in Tennessee and the following variables: a) mean reading TCAP scores, b) per-pupil expenditure, c) school size, d) attendance, and e) percentage of students qualifying for free and reduced-priced meals?

The most efficient way to answer this question was to produce a table showing Pearson correlation values. The closer the value was to either -1 or 1, the stronger the relationship. All data used in the study were school-level from the Tennessee State Department of Education's 2008 Report Card.

Correlation coefficients were computed among the 6 school variables. Using the Bonferroni approach to control for Type I error across the 6 correlations, a *p* value of less than .008 (05 / 6 = .008) was required for significance. The results of the correlational analyses are presented in Table 2.

Table 2

Pearson Correlation Values Between2007-2008 Mean Science NCE Scores and 2007-2008 Mean Reading NCE Scores, Per-Pupil Expenditure, School Size, Attendance, and Percentage of Students Qualifying for Free/Reduced-Priced Meals

	8th grade mean reading NCE	Per-pupil Expenditure	# of students	Attendance	% free and reduced priced-meals
8th grade mean reading NCE		-.58 *	.27*	.61*	-.88*
8th grade mean science NCE	.92*	-.62*	.23*	.60*	-.84*

* Correlation was significant at the 0.08 level.

The school characteristic in Table 2 with the strongest association to the mean 8th grade science TCAP scores as computed by Pearson's correlation was the mean 8th grade reading TCAP scores. The values could be ranked as follows: 1) mean 8th grade reading scores (.92), 2) percentage of free/reduced-priced meals (-.84), 3) per-pupil expenditure (-.62), 4) attendance (.60), and 5) school size (.23). The negative values for per-pupil expenditure seemed surprising at first glance but further analysis might explain. Schools that served inner city students tended to have a higher percentage of students qualifying for free and reduced-priced meals and offered higher than average wages in order to attract teachers. Thus, the average per-pupil expenditure in Tennessee was $8,345; however, Memphis spent $10, 366. The same urban school districts generally have TCAP scores below the state mean.

Ho: 1_1 There is no relationship between mean reading TCAP scores and mean science TCAP scores of 8th graders in Tennessee.

A Pearson correlation coefficient was computed to determine whether there was a correlation between mean reading TCAP scores and mean science TCAP scores for 8th graders in Tennessee. A *p* value of less than .008 (.05/6 = .008) was required for significance and the Bonferroni approach controlled for Type I error. The correlation between mean reading TCAP scores and mean science TCAP scores of 8th graders in Tennessee was significant, $r(179) = .92$, $p<.008$. The finding of a positive correlation resulted in the rejection of the null hypothesis. Figure 1 provides a descriptive picture of this relationship.

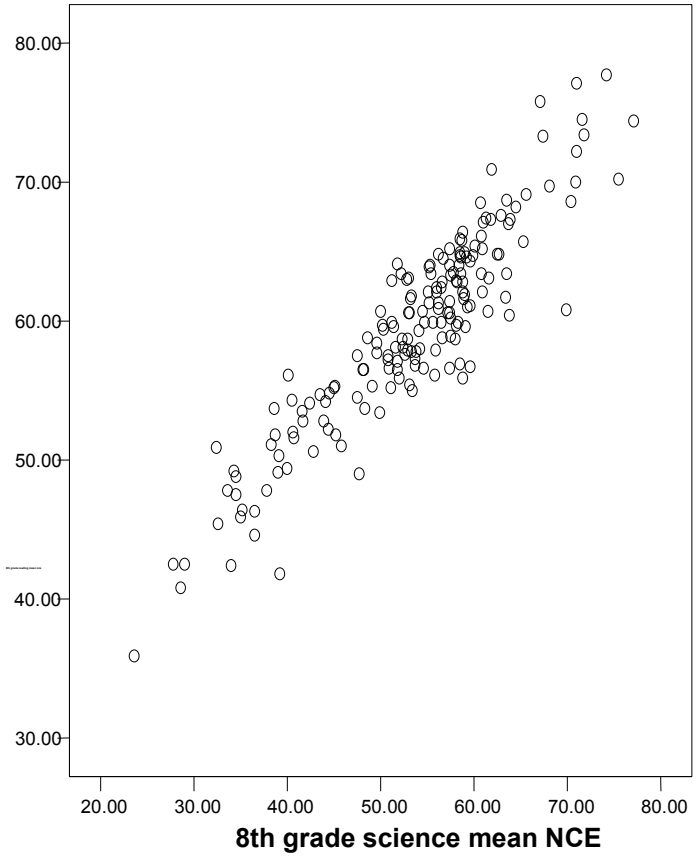

Figure 1. Scatter Plot Showing the Relationship Between 8th Grade Mean Reading NCE Scores and 8th Grade Mean Science NCE Scores from 2008 TCAP Report Data.

Each dot in the scatter plot represents a Tennessee school configured 6-8. The horizontal axis indicates the mean 2007-2008 8th grade science NCE for each school. The vertical axis indicates the mean 2007-2008 8th grade reading NCE for each school. If the dots were perfectly aligned, the *r* value would equal one. The *r* value of .92 represents a strong relationship.

Ho: 1_2 There is no relationship between per-pupil expenditures and mean science TCAP scores 8th graders in Tennessee.

A Pearson correlation coefficient was computed to determine whether there was a correlation between per-pupil expenditure and mean science TCAP scores for 8th graders in

Tennessee. A p value of less than .008 (.05/6 = .008) was required for significance and the Bonferroni approach controlled for Type I error. The correlation between and mean science TCAP scores of 8th graders and school per-pupil expenditure in Tennessee was significant, $r(179) = -.62$, $p<.008$. The finding of a negative correlation resulted in the rejection of the null hypothesis. These negative values implied that the more money spent on teacher and pupils, the lower the mean science TCAP NCE. However, middle schools located in large metropolitan areas, such as Memphis, that had high percentages of students qualifying for free and reduced-priced meals spent more on their pupils and teachers. Figure 2 provides a descriptive picture of this relationship.

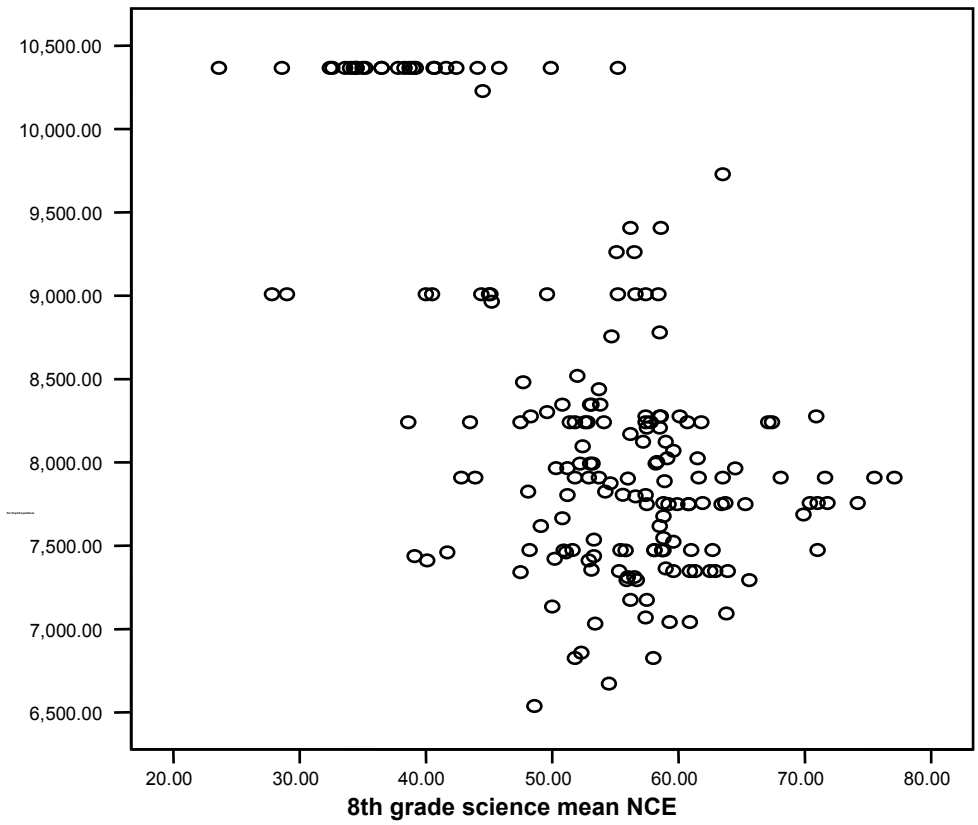

Figure 2. Scatter Plot Showing the Relationship Between School Per-Pupil Expenditure and 8th Grade Mean Science NCE Scores from 2008 TCAP Report Data.

Each dot in the scatter plot represents a Tennessee school configured 6-8. The horizontal axis indicates the mean 2007-2008 8th grade science NCE for each school. The vertical axis indicates the per-pupil expenditure for each school. The range of per-pupil expenditure for this study's range was wide from $6,538 (Bedford County) to $10,366 (Memphis), equaling a difference of $3,828. If the dots were perfectly aligned, the *r* value would equal one. The *r* value for this relationship is -.62.

Ho: 1_3 There is no relationship between school size and mean science TCAP scores of 8th graders in Tennessee.

A Pearson correlation coefficient was computed to determine whether there was a correlation between mean science TCAP scores for 8th graders and school size in Tennessee. A *p* value of less than .008 (.05/6 = .008) was required for significance and the Bonferroni approach controlled for Type I error. The correlation between school size and mean science TCAP scores of eighth graders in Tennessee was significant, $r(179) = .23$, $p<.008$. The finding of a positive correlation resulted in the rejection of the null hypothesis. Figure 3 provides a descriptive picture of this relationship.

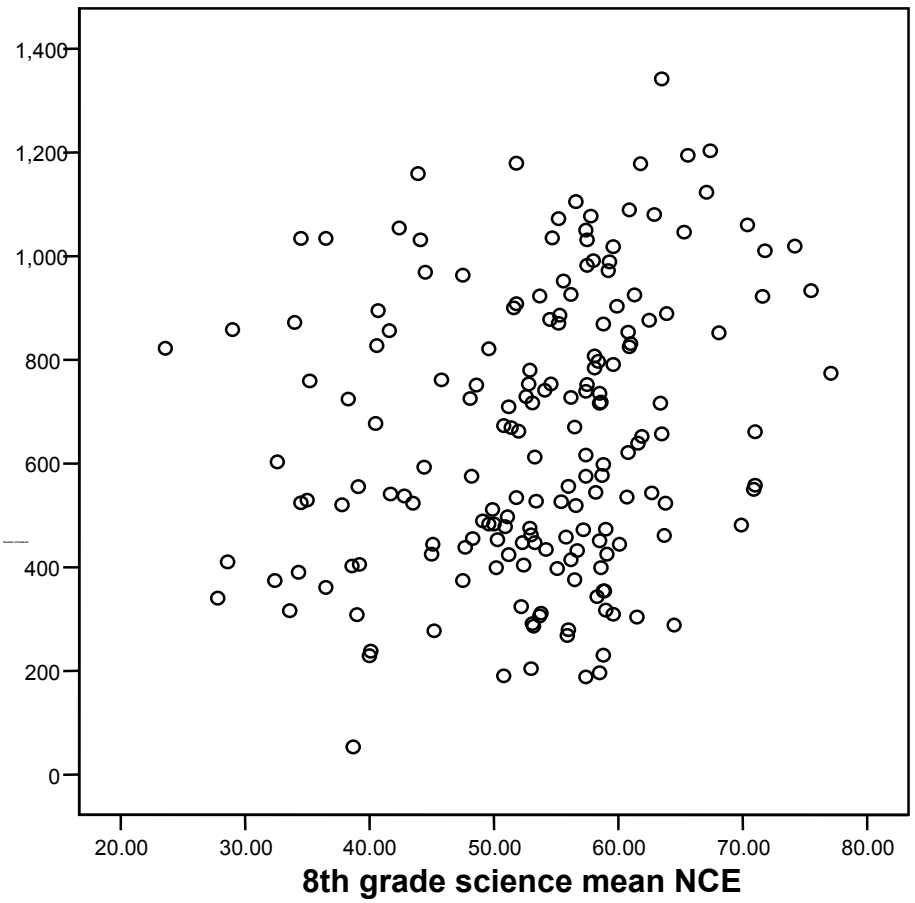

Figure 3. Scatter Plot Showing the Relationship Between School Size and 8th Grade Mean Science NCE Scores from 2008 TCAP Report Data.

Each dot in the scatter plot represents a Tennessee school configured 6-8. The horizontal axis indicates the mean 2007-2008 8th grade science NCE for each school. The vertical axis indicates the school size for each school. Student enrollment ranged from 1,342 students (Arlington Middle, Shelby County) to 53 students (Ida B Wells, Memphis). If the dots were perfectly aligned, the *r* value would equal one. The *r* value for this relationship is .23.

Ho: 1_4 There is no relationship between attendance and mean science TCAP scores of 8th graders in Tennessee.

A Pearson correlation coefficient was computed to determine whether there was a correlation between mean science TCAP scores of 8th graders and attendance in Tennessee. A *p* value of less than .008 (.05/6 = .008) was required for significance and the Bonferroni approach controlled for Type I error. The correlation between attendance and mean science TCAP scores of 8th graders in Tennessee was significant, $r(175) = .60$, $p<.008$. The finding of a positive correlation resulted in the rejection of the null hypothesis. Figure 4 provides a descriptive picture of this relationship.

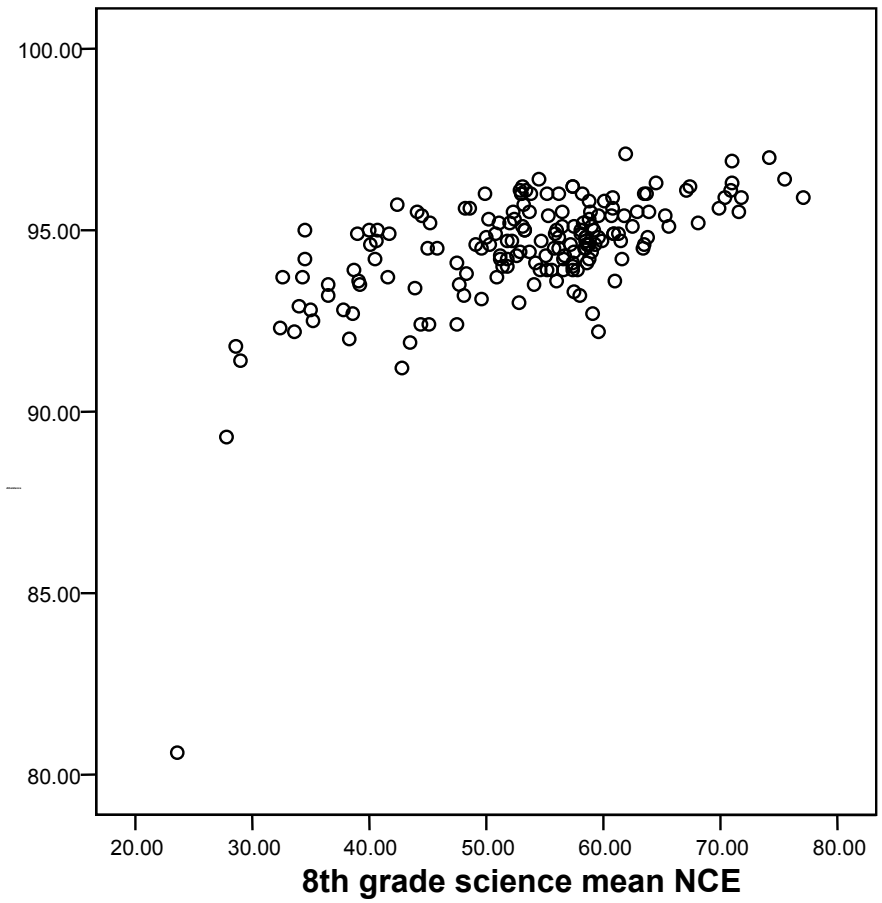

Figure 4. Scatter Plot Showing the Relationship Between Attendance and 8th Grade Mean Science NCE Scores from 2008 TCAP Report Data.

Each dot in the scatter plot represents a Tennessee school configured 6-8. The horizontal axis indicates the mean 2007-2008 8th grade science NCE for each school. The vertical axis indicates the attendance reported for each school. If the dots were perfectly aligned, the *r* value would equal one. The *r* value for this relationship is .60.

Ho: 1_5 There is no relationship between percentage of students qualifying for free and reduced-priced meals and mean science TCAP scores of 8th graders in Tennessee.

A Pearson correlation coefficient was computed to determine whether there was a correlation between mean science TCAP scores of 8th graders and percentage of students

qualifying for free and reduced-priced meals in Tennessee. A p value of less than .008 (.05/6 = .008) was required for significance and the Bonferroni approach controlled for Type I error. The correlation between the percentage of students qualifying free and reduced-priced meals and mean science TCAP scores of 8th graders in Tennessee was significant, $r(179) = -.84$, $p<.008$. The finding of a negative correlation resulted in the rejection of the null hypothesis. Figure 5 provides a descriptive picture of this relationship.

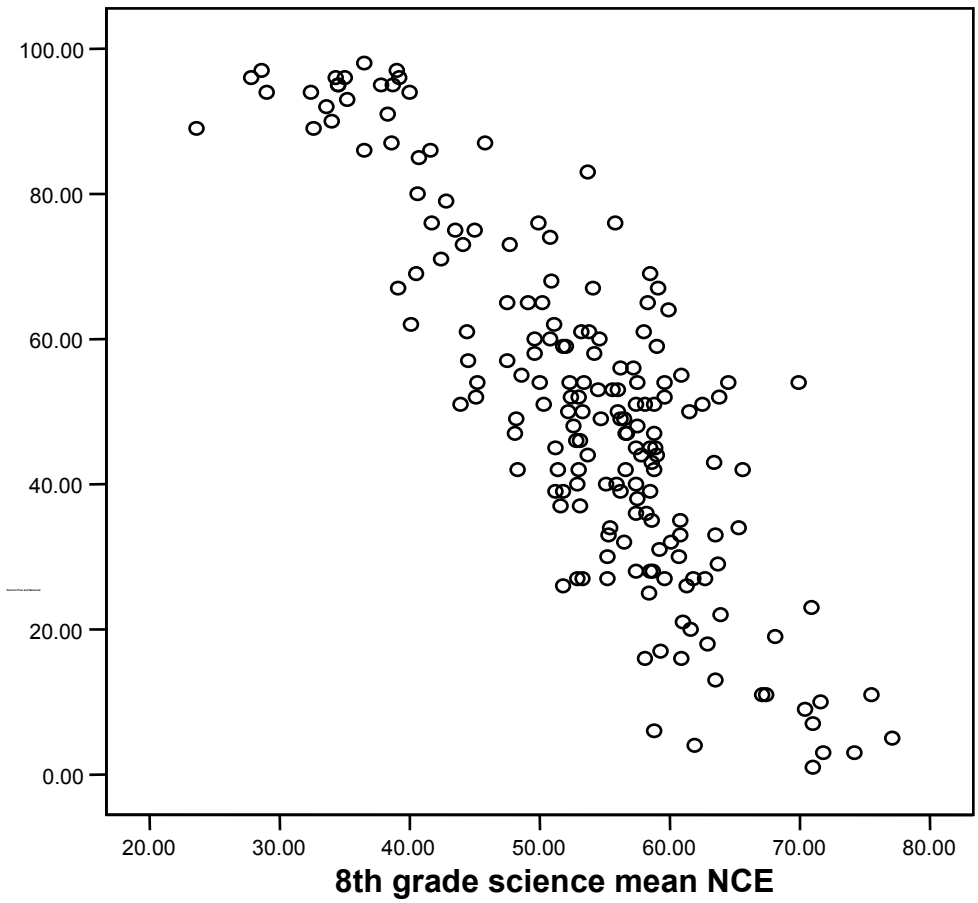

Figure 5. Scatter Plot Showing the Relationship Between Percentage of Students Qualifying for Free/Reduced Meals and 8th Grade Mean Science NCE Scores from 2008 TCAP Report Data.

Each dot in the scatter plot represents a Tennessee school configured 6-8. The horizontal axis indicates the mean 2007-2008 8th grade science NCE for each school. The vertical axis indicates the mean 2007-2008 8th grade reading NCE for each school. If the dots were perfectly aligned, the *r* value would equal one. The *r* value of -.84 represents a strong relationship.

Research Question 2

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state mean of $8,345?

The most efficient way to answer this question was to conduct an independent-samples *t* test to evaluate whether or not schools with above mean per-pupil expenditures scored better on the science TCAP test than schools with below mean per-pupil expenditures.

Ho: 2 There is no difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state mean of $8,345.

An independent- samples *t* test was conducted to evaluate whether 8th grade mean science scores and the state per-pupil expenditure were different between schools with per-pupil expenditures above the state mean and schools with per-pupil expenditures below the state mean. The 8th grade mean science scores was the test variable and the grouping variable was above or below the state per-pupil expenditure of $8,435. The test was significant, t (179) = -9.84, $p<.001$. Therefore, null hypothesis Ho: 2 was rejected. Schools with per-pupil expenditures above the state mean (M = 44.27, SD = 9.77) scored lower than those with higher per-pupil expenditures (M = 57.18, SD = 7.25). The 95% confidence interval for the differences in means was -15.50 to -9.97. The η^2 index was .35, which indicated a large effect size. Schools with above state mean per-pupil expenditures tended to score lower than schools with below state mean per-pupil expenditures.

Research Question 3

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state mean of $8,345?

The most efficient way to answer this question was to conduct an independent-samples *t* test to evaluate whether or not schools with above mean per-pupil expenditures scored better on the reading TCAP than schools with below mean per-pupil expenditures.

Ho: 3 There is no difference between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state mean of $8,345.

An independent-samples t test was conducted to evaluate whether 8th grade mean reading scores and the state per-pupil expenditure were different between schools with per-pupil expenditures above the state mean and schools with per-pupil expenditures below the state mean. The 8th grade mean reading scores was the test variable and the grouping variable was above or below the state's per-pupil expenditure of $8,435. The test was significant, t (179) = -9.40, $p<.001$. Therefore, null hypothesis Ho: 3 was rejected. Schools with per-pupil expenditures above the state mean ($M = 53.00$, $SD = 7.04$) scored lower than those with higher per-pupil expenditures ($M = 62.10$, $SD = 5.44$). The 95% confidence interval for the difference in means was -11.00 to -6.96. The η^2 index was .33, which indicated a large effect size. Schools with above state mean per-pupil expenditures tended to score lower than schools with below state mean per-pupil expenditures.

Research Question 4

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

The most efficient way to answer this question was to conduct an independent-samples *t* test to evaluate whether or not schools reporting attendance above 93% scored better on science TCAP than schools reporting attendance below 93%.

Ho: 4 There is no difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentages above or below the state goal of 93%.

An independent-samples *t* test was conducted to evaluate whether 8th grade mean science scores were different between schools with attendance percentages above the state goal of 93% and schools with attendance reported below the state goal of 93%. The 8th grade mean science NCE scores were the test variable and the grouping variable was attendance. The test was significant, $t(175) = 7.68$, $p<.001$. Therefore, null hypothesis Ho: 4 was rejected. Schools with attendance above the state goal ($M = 55.04$, $SD = 8.62$) scored higher than schools with attendance below the state goal ($M = 38.73$, $SD = 9.72$). The 95% confidence interval for the difference in means was 12.11 to 21.15. The η^2 index was .25, which indicated a large effect size. Schools with attendance above the state goal tended to score higher than schools with attendance below the state goal.

Research Question 5

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

The most efficient way to answer this question was to conduct an independent-samples *t* test to evaluate whether or not schools reporting attendance above 93% scored better on reading TCAP than schools reporting attendance below 93%.

Ho: 5 There is no difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentages above or below the state goal of 93%. An independent-samples *t* test was conducted to evaluate whether 8th grade mean reading scores were different between schools with attendance reported above the state goal of 93% and schools with attendance reported below the state goal of 93%. The 8th grade mean reading scores were the test variable and the grouping variable was attendance. The test was significant, t (175) = 7.09, $p<.001$. Therefore, null hypothesis Ho: 5 was rejected. Schools with attendance above the state goal ($M = 60.55$, $SD = 6.43$) scored higher than schools with attendance below the state goal ($M = 49.41$, $SD = 6.79$). The 95% confidence interval for the difference in means was 8.04 to 14.54. The η^2 index was .22, which indicated a large effect size. Schools with attendance above the state goal tended to score higher than schools with attendance below the state goal.

CHAPTER 5

SUMMARY OF FINDINGS, CONCLUSIONS, AND RECOMMENDATIONS

The purpose of this study was to determine whether Tennessee Comprehensive Assessment Program (TCAP) Achievement reading scores, school district per-pupil expenditures, school size, percentage of students qualifying for free and reduced-priced meals, and attendance were related to science TCAP test scores of 8th graders from the 2007-2008 school year. Pearson correlations were conducted between the 8th mean grade science TCAP scores and 8th grade mean reading scores, per-pupil expenditure, school size, attendance, and the percentage of students qualifying for free and reduced-priced meals. Independent-samples t tests were conducted to evaluate whether 8th grade mean science TCAP scores and 8th grade mean reading scores differed depending on whether the school per-pupil expenditure was above or below the state average of $8,345. Independent-samples *t* test were also conducted to evaluate whether 8th grade mean science TCAP scores and 8th grade mean reading scores differed depending on whether the school attendance percentage was above or below the state goal of 93%. Chapter 5 contains a summary of the study, a summary of the findings, conclusions, recommendations for practice, and recommendations for further research.

Summary of Findings

This quantitative study examined the mean science scores of 8th grade students enrolled in middle schools configured 6th through 8th grade in Tennessee that participated in the spring 2008 TCAP test for 8th grade mean reading scores. Those scores were examined in relation to per-pupil expenditure, school size, attendance, and the percentage of students qualifying for free and

reduced-priced meals. The analysis focused on 5 research questions and 9 null hypotheses. The sample included data on 181 Tennessee middle schools that were configured grades 6th through 8th. The analysis of attendance only included 177 schools. All other analysis in this study used 181 schools. All schools were listed in the 2008 Tennessee Schools Report Card website. The data were analyzed using the Statistical Package for the Social Sciences (SPSS).

Research Question 1

> Is there a relationship between mean science TCAP scores of 8th graders in Tennessee and the following variables: a) mean reading TCAP scores, b) per-pupil expenditure, c) school size, d) attendance, and e) percentage of students qualifying for free and reduced-priced meals?

The Pearson correlation coefficient of .92 indicated that there was a strong positive correlation between 8th grade mean reading TCAP scores and 8th grade mean science TCAP scores. Thus, the 1st null hypothesis was rejected.

The Pearson correlation value of .92 indicated the strongest relationship found in this study. Students needed to be able to read and comprehend information to be successful on standardized tests. Reading comprehension was not only tested on the reading portions of standardized tests. Flick and Leaderman (2002) reported that with the rise in standardized testing, there was growing concern that reading comprehension was being tested in other subjects, such as science and math. Fleener and Bucher (2003) also warned that being able to read, however, did not mean that the student was able to comprehend the material being read.

The Pearson correlation coefficient of -.62 indicated there was a strong negative correlation between per-pupil expenditure and 8th grade mean science TCAP scores. Thus, the 2nd null hypothesis was rejected.

The closer the value was to either -1 or 1, the stronger the relationship. The Pearson correlation value of -.62 indicated a strong negative relationship between per-pupil expenditure and 8th grade mean science TCAP scores. This might offer the impression that spending more money on students would negatively affect achievement. However, Webb (2005) conducted a similar study identifying school characteristics most closely associated with ACT scores. His findings yielded a negative relationship between per-pupil expenditure and ACT scores. However, when he delved further into his analysis and eliminated the inner city schools from his data, his initial findings changed from a Pearson's correlation value for per-pupil expenditure with ACT scores from negative .28 to positive .41 (Webb, 2005). This example demonstrated how easily variables could be misconstrued.

The Pearson correlation coefficient of .23 indicated that there was a positive correlation between school size and 8th grade mean science TCAP scores. Thus, the 3rd null hypothesis was rejected.

The closer the value was to either -1 or 1, the stronger the relationship. The Pearson correlation value of .23 indicated a positive relationship between school size and 8th grade mean science TCAP scores, but was the weakest relationship found among all the variables in this study. As previously mentioned, economy of scale was becoming of more concern in the financing of school districts. Increasing school sizes might make them more economically efficient, but when schools became too large, they tended to become economically inefficient (Alspaugh, 1994).

The Pearson correlation coefficient of .60 indicated that there was a strong positive correlation between attendance and 8th grade mean science TCAP scores. Thus, the fourth null hypothesis was rejected.

Pearson correlation values ranged from -1 to 1. The closer the value was to either -1 or 1, the stronger the relationship. The Pearson correlation value of .60 indicated a significant positive relationship between attendance and 8th grade mean science TCAP scores. Cassell (2007) conducted a similar study identifying relationships between student attendance and test scores on the Virginia Standards of Learning Test. Cassell also found a significant positive correlation between attendance and test scores. These findings were not surprising because it was known and widely accepted that students attending school on a regular basis was crucial to their academic success.

The Pearson correlation coefficient of -.84 indicated that there was a strong negative correlation between the percentage of students qualifying for free and reduced meals and 8th grade mean science TCAP scores.

Pearson correlation values ranged from -1 to 1. The closer the value was to either -1 or 1, the stronger the relationship. The Pearson correlation value of -.84 indicated a strong negative correlation between the percentage of students qualifying for free and reduced meals and 8th grade mean science TCAP scores. Aside from the positive correlation between 8th grade mean reading TCAP scores and 8th grade mean science TCAP scores, the negative correlation of percentage of students qualifying for free and reduced meals possessed the second strongest correlation to 8th grade mean science TCAP scores. Caldwell and Ginthier (1996) reported that because of increased economic, legal, and psychological problems, students with low socioeconomic status constituted the largest population of individuals at risk of not graduating from high school. The factors associated with socioeconomic characteristics of students could not be controlled by schools, but Dunn and Griggs (1988) urged educators not to give up on these students. Educators could alter the path their students choose to take.

Research Question 2

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

An independent-samples *t* test was conducted to evaluate whether 8th grade mean science scores and the state per-pupil expenditure were different between schools with per-pupil expenditures above the state average and schools with per-pupil expenditures below the state average. The test was significant, t (179) = -9.83, $p < .001$. Thus, the null hypothesis was rejected.

Research Question 3

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with per-pupil expenditures above or below the state average of $8,345?

An independent-samples *t* test was conducted to evaluate whether 8th grade mean reading scores and the state per-pupil expenditure were different between schools with per-pupil expenditures above the state average and schools with per-pupil expenditures below the state average. The test was significant, t (179) = -9.39, $p < .001$. Thus, the null hypothesis was rejected.

Research Question 4

Is there a difference between mean science TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

An independent-samples *t* test was conducted to evaluate whether 8th grade mean science scores were different between schools with attendance percentages above the state goal of 93% and schools with attendance reported below the state goal of 93%. The test was significant, t (175) = 7.68, $p < .001$. Thus, the null hypothesis was rejected.

Research Question 5

Is there a difference between mean reading TCAP scores of 8th graders in Tennessee and schools with attendance percentage above or below the state goal of 93%?

An independent-samples *t* test was conducted to evaluate whether 8th grade mean reading scores were different between schools with attendance reported above the state goal of 93% and schools with attendance reported below the state goal of 93%. The test was significant, t (175) = 7.09, $p < .001$. Thus, the null hypothesis was rejected.

Conclusions

Conclusion #1

The results of this study indicated a strong positive relationship (.92) between 8th grade mean science scores and 8th grade mean reading scores. These results might not be particularly surprising because, as Flick and Lederman (2002) acknowledged, with the rise in standardized testing, there was growing concern that reading comprehension was being tested in other subjects, such as science. Jongsma (2004) noted that teachers had to understand the importance of excellent comprehension instruction in the primary grades and continue to refine them through the intermediate and middle grades. Ediger (2005) also noted that all teachers regardless of content area need to be instructors of reading. If schools wish to increase their TCAP scores in science, they need to increase the reading comprehension of their students.

Conclusion #2

The results of this study indicated a strong positive relationship (.60) between 8th grade mean science scores and attendance. Attendance was traditionally viewed as possessing a strong correlation to academic performance in the classroom (Atkinson, 2005; Moore, 2005) but little

research was conducted on the topic until the implementation of No Child Left Behind. This study supports the assumption that attendance is directly related to student achievement and therefore, contributes to a growing body of research resulting from the enactment of the No Child Left Behind legislation.

Conclusion #3

The results of this study indicated a strong negative relationship (-.84) between 8th grade mean science scores and the percentage of students qualifying for free and reduced lunch. Unless students from poorer socioeconomic status can access quality instruction and remediation and attend school regularly, it is probable that large numbers will continue to receive low scores on the TCAP test, thus, exacerbating the social divide between the rich and the poor.

Recommendations to Improve Practice

1. Administrators should provide all teachers, regardless of content area, with staff development opportunities centered on teaching reading comprehension. Fleener and Bucher (2003) emphasized that being able to read did not mean that the student was able to comprehend the material being read. Schools that wished to improve their TCAP scores should concentrate on reading comprehension. The format of the TCAP requires students be able to read and comprehend questions of various lengths.

2. Educators must recognize the importance of teacher efficiency in approving student attendance. Educators should develop and implement programs that encourage, support, and improve sustained student attendance. Promoting active student engagement, for example, has motivated students to attend school and remain there until graduation (National Research Council, 2003). Klem and Connell (2004)

suggested that by creating a personalized student environment levels of engagement would increase along with student attendance and test scores. Student motivation plays a key role in several facets of education. Administrators should provide staff development opportunities to enrich teacher knowledge of effective instructional strategies that actively engage and interest students.

3. Students who are motivated are more likely to attend class on a regular basis (Tretter & Jones, 2003). Ediger (2001c) found that the teacher plays a key role in student motivation. Teachers should experiment with both intrinsic and extrinsic motivation to decide which approach or combination of intrinsic and extrinsic works best with their particular group of students.

Recommendations for Further Research

1. Additional research can be conducted between the remaining two subjects (math and social studies) tested by TCAP for relationships they may possess to reading scores.

2. Additional research can be conducted that seeks other factors that contribute to increased test scores.

3. Studies can be conducted to identify relationships that may exist between various grade configurations.

4. The practices of schools with high mean scores on TCAP exams that are serving poor and minority students should be studied and the strategies and practices should be disseminated to other Tennessee schools.

5. This study addressed the relationship between reading scores and science scores. However, Neill (2003) reported that research shows that high-stakes testing does not improve long term educational outcomes. Additional research could be conducted comparing college entrance exams scores such as the ACT to school, district, and state test accountability measures. Perhaps the emphasis on high stakes achievement tests in public schools is misplaced.

6. The relationship between reading scores and science scores in this study raises significant questions about whether standardized tests used to assess students and schools truly measure what they are assumed to measure. A good deal of the literature discussed in Chapter 2 raises questions about how valid theses tests are at measuring what they claim to measure. My findings surface a significant possibility that Tennessee's science TCAP achievement tests are a better measure of a students' ability to read and comprehend than their knowledge of science. Additional research could be conducted on the validity of standardized tests to see if they accurately measure the skills they are designed to measure.

REFERENCES

Aldridge, J. (2003). Rethinking the No Child Left Behind Act of 2001. *Childhood Education*, *80*(1), 45-50.

Alspaugh, J. W. (1994). The relationship between school size, student teacher ratio and school efficiency. *Education*, *114*, 593-601.

Amrein, L., & Berliner, D. (2003, February). The effects of high-stakes testing on student motivation and learning. *Educational Leadership*, *60*(5), 32-37.

Anderson, G. L. (2002). Critique of the test for school leaders. *Educational Leadership*, 59(8), 67-70.

Atkinson, A. J. (Ed.). (2005). *Improving school attendance: A resource guide for Virginia schools.* Richmond, VA.: Virginia Department of Education.

Avalon Project of Yale Law School (2005). *Northwest ordinance: July 1787.* Retrieved September 1, 2007, from http://www.yale.edu/lawweb/avalon/nworder.htm

Azzam, A. M. (2007). The intervention called NCLB. *Educational Leadership*, *65*(2), 92-93.

Berliner, D. C., & Biddle, B. J. (1995). *The manufactured crisis: Myths, fraud, and the attack on America's public schools.* Cambridge, MA: Perseus Books.

Bracey, G. W. (1993). Money matters. *Phi Delta Kappan*, *74*, 573-575.

Bracey, G. W. (1996). Money improves test scores- even state-level SATs. *Phi Delta Kappan*, *78*, 91-96.

Bracey, G. W. (1997). Money matters: No it doesn't, yes it does. *Phi Delta Kappan*, *78*, 162-167.

Bracey, G. W. (2003). Research: We're number one. *Phi Delta Kappan*, *85,* 87-92.

Bracey, G. W. (2007). The first time 'everything changed': The 17th Bracey Report on the condition of public education. *Phi Delta Kappan*, *89*, 119-140.

Butler, R. J., & Monk, D. H. (1985). The cost of public schooling in New York State: the role of scale and efficiency in 1978-79. *The Journal of Human Resources, 20*, 361-381.

Caldwell, G. P., & Ginthier, D. W. (1996). Differences in learning styles of low socioeconomic status for low and high achievers. *Education*, *117*, 141-149.

Cassell, J. D. (2007). *Relationships between student attendance and test scores on the Virginia Standards of Learning tests*. Unpublished doctoral dissertation, East Tennessee State University.

Chen, Y., & Martin, M. A. (2000). Using performance assessment and portfolio assessment together in the elementary classroom. *Reading Improvement*, *37*(1), 32.

Corville-Smith, J., Ryan, B. A., Adams, G. R., & Dalicandro, T. (1998). Distinguishing absentee students from regular attendors: The combined influence of personal, family, and school factors. *Journal of Youth and Adolescence*, *27*, 629-635.

Cruickshank, D. R., & Haefele, D. (2001, January). Good teachers, plural. *Educational Leadership*, *58*(5), 26-30.

Donegan, M., & Trepanier-Street, M. (1998, February). Teacher and parent views on standardized testing: A cross-cultural comparison of the uses and influencing factors. *Journal of Research in Childhood Education*, *13*, 85-93.

Dunn, R., & Griggs, S. A. (1988). High school dropouts: Do they learn differently from those who remain in school. *The Principal*, *34*, 1-8.

Eberts, R. W. (2007). Teachers unions and student performance: Help or hindrance? *The Future of Children*, *17*, 175-196.

Ediger, M. (2001a). *Assessment of school principal achievement* (Report No. TM032521). Published by. (ERIC Document Reproduction Service No. ED452235)

Ediger, M. (2001b). *Assessing approaches in upping student test scores* (Report No. TM032832). Published by. (ERIC Document Reproduction Service No. ED453293)

Ediger, M. (2001c). *Phonics activities in the reading curriculum* (Report No. CS014512). Published by. (ERIC Document Reproduction Service No. ED458546)

Ediger, M. (2003). *Testing and predictions of pupil success* (Report No. TM034832). Published by. (ERIC Document Reproduction Service No. ED475356)

Ediger, M. (2005). Assessing reading in the science curriculum. *College Student Journal*, *39*(1), 26-27.

Eisner, E. (2001). What does it mean to say a school is doing well? *Phi Delta Kappan*, *82*, 367-372.

Flick, L. B., & Lederman, N. C. (2002). The value of teaching reading in the context of science and mathematics. *School Science and Mathematics*, *102*, 105-107.

Finder, M. (2004). *Educating America: How Ralph W. Tyler taught America to teach.* Westport, CT: Praeger.

Fleener, C. E., & Bucher, K. T. (2003). Linking reading, science and fiction books. *Childhood Education*, *80*(2), 76-78.

Flippo, R., & Riccards, M. (2000, September). Initial teacher certification testing in Massachusetts: A case of the tail wagging the dog. *Phi Delta Kappan*, *82*, 34-37.

Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education.* New York: McGraw Hill.

Gandal, M., & McGiffert, L. (2003, February). The power of testing. *Educational Leadership*, *60*(5), 39-42.

Gardner, H. (1993). *Multiple intelligences: The theory in practice.* New York: Basic Books.

Gibson, R. (2001). *Outfoxing the destruction of wisdom* (Report No. TM033318). Published by. (ERIC Document Reproduction Service No. ED457220)

Gough, P. B. (2002). The editor's page: Outside the big-school box. *Phi Delta Kappan*, *83*, 342.

Gratz, D. (2000). High standards for whom? *Phi Delta Kappan*, *81*, 681-687.

Hargreaves, A. (2001). Beyond anxiety and nostalgia: Building a social movement for educational change. *Phi Delta Kappan*, *82*, 373-377.

Harrison, J., & McAfee, H. (2002). *Beyond standardized testing: Examining, developing, and validating the interview for admission into the teacher education program* (Report No. TM034684). Published by. (ERIC Document Reproduction Service No. ED471655)

Health & Education Research Operative Services, Inc. (2003). Retrieved January 2009, from http://www.heros-inc.org/star.htm

Hirsh, E. D. (2006). *The knowledge deficit: Closing the shocking education gap for American children.* New York: Houghton Mifflin.

Holloway, J. (2001). Research link: The use and misuse of standardized tests. *Educational Leadership*, *59*(1), 77-78.

Holloway, J. (2002). A defense of the test for school leaders. *Educational Leadership*, *59*(8), 71-75.

Hughes, S., & Bailey, J. (2001, December). What students think about high-stakes testing. *Educational Leadership*, *59*(4), 74-76.

Illig, D. C. (1996, June). Reducing class size: A review of the literature and options for consideration. Retrieved February 5, 2008, from http://www.library.ca.gov/CRB/classz/clssiz.html.

Jennings, J., & Rentner, D. S. (2006). Ten big effects of the No Child Left Behind act on public schools. *Phi Delta Kappan*, *88*, 358-365.

Jongsma, K. (2004). Teaching reading in social studies, science, and math. *The Reading Teacher*, *57*, 544.

Key, S. (1996). Economics or education: The establishment of American land-grant universities. *Journal of Higher Education*, *67*, 196-201.

Khaliq, S., & Sireci, S. (2002). *An analysis of the psychometric properties of dual language test forms* (Report No. TM034399). Published by. (ERIC Document Reproduction Service No. ED468489)

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health*, *74*, 262-273.

Kohn, A. (2001). Fighting the tests: A practical guide to rescuing our schools. *Phi Delta Kappan*, *82*, 349-356.

Lopus, J. S. (1990). Do additional expenditures increase achievement in the high school economic class? *Journal of Economic Education*, *21*, 277.

Lynd, C. (2000). *The new generation of standardized testing*. Published by. (ERIC Document Reproduction Service No. ED452239)

McColskey, W., & McMunn, N. (2000, October). Strategies for dealing with high-stakes state tests. *Phi Delta Kappan*, *82*, 115-120.

McDermott, T. K., & McDermott, D. F. (2002, March). High-stakes testing for students with special needs. *Phi Delta Kappan*, *83*, 504-544.

Moore, R. (2005). Attendance: Are penalties more effective than rewards? *Journal of Developmental Education*, *29*(2), 26-32.

Morgan, J. G., Moore, R., Detch, E., & Walton, J. (2007). *Tennessee's graduation exams: Past, present, and future.* Retrieved December 2007, from http://www.comptroller.state.tn.us/orea/reports/highstakestest.pdf

National Research Council. (2003). *Engaging schools: Fostering high school students' motivation to learn.* Washington, DC: The National Academies Press.

Neill, M. (2003). The dangers of testing. *Educational Leadership*, *60*(5), 43-45.

Ohanian, S. (2001). News from the test resistance trail. *Phi Delta Kappan*, *82*, 363-366.

Pangle, L. S., & Pangle, T. L. (1993). The learning of liberty: The educational ideas of the American founders. Lawrence, KS: University Press of Kansas.

Pate-Bain, H., Fulton, D., & Boyd-Zaharias, J. (1999). *Effects of class-size reduction in the early grades (K-3) on high school performance.* Retrieved January 2009, from http://www.heros-inc.org/star-hs-p.pdf

Pellegrino, J. (2007). Should NAEP performance standards be used for setting standards for state assessments? *Phi Delta Kappan*, *88*, 539-541.

Perreault, G. (2000). The classroom impact of high-stress testing. *Education*, *120*, 705-710.

Popham, W. (1999). Why standardized tests don't measure educational quality. *Educational Leadership*, *56*(1), 8-15.

Popham, W. (2001). Uses and misuses of standardized tests. *NASSP Bulletin*, *85*(622), 24-31.

Rose, L. C., & Gallup, A. M. (2003). The 35th annual Phi Delta Kappa/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 85*, 42-52.

Rotenberg, A. (2002). *A classroom research project: The psychological effects of standardized testing on young English language learners at different language proficiency levels* (Report No. FL027596). Published by. (ERIC Document Reproduction Service No. ED472651)

Rubado, K. (2002). Empowering students through multiple intelligences. *Reclaiming Children & Youth*, *10*, 233-237.

Sanchez, C. (2008). '*No Child' law picked apart as renewal fight looms*. Retrieved February 3, 2008, from http://www.npr.org/templates/story/story.php?storyId=18432881.

Scott, J. C. (2006). The mission of the university: Medieval to postmodern transformations. *Ohio State University Press*, 77, 1-21.

Shaul, M. S. (2006). *No Child Left Behind Act: States face challenges measuring academic growth*: *Testimony before the House Committee on Education and the Workforce*. Washington, DC: United States Government Accountability Office.

Shelton, K. M., & Biddle, B. J. (1998). Standards, accountability, and school reform: Perils and pitfalls. *Teachers College Record*, *100*, 164-180.

Spegelken, R. (2007). Why 'thinking like a scientist' is relevant to kids: Professor Wendy Williams is developing a curriculum for female and minority youth of low socioeconomic status that teaches kids how they can use scientific reasoning for the problems they confront in their daily lives. *Human Ecology*, *35*(1), 20-24.

Spellings, M. (2007). *Building on results: A blueprint for strengthening the No Child Left Behind Act.* Retrieved February 3, 2008, from www.ed.gov

Standardized Testing: One size fits all? (2002) *A WisKids Count Special Report* (Report No. TM034585). Published by. (ERIC Document Reproduction Service No. ED471080)

Tennessee Department of Education. (2005). *A handbook for principals.* Nashville, TN: Author.

Tennessee Department of Education. (2007a). *Achievement test.* Retrieved March 5, 2007, from http://www.state.tn.us

Tennessee Department of Education. (2007b). *Tennessee report card 2007.* Retrieved March 27, 2008, from http://www.state.tn.us

Tennessee Department of Education. (2007c). *Educators' guide to understanding TCAP achievement test results 2007.* Retrieved March 27, 2008, from http://tennessee.gov/education/assessment

Tennessee Department of Education (2007d). *History of education in Tennessee.* Knoxville, TN: University of Tennessee Press.

Tennessee education laws annotated. (2000). Charlottesville, VA: Lexis.

Thattai, D. (2001). Education. Retrieved November 10, 2008, from http://www.servintfree.net/~aidmn-ejournal/publications/2001-11/PublicEducationInTheUnitedStates.html

Thompson, S. (2001). The authentic standards movement and its evil twin. *Phi Delta Kappan*, *82*, 358-362.

Tretter, T. R., & Jones, M. G. (2003). Relationships between inquiry-based teaching and physical science standardized test scores. *School Science and Mathematics*, *103*, 345-360.

United States Department of Education. (2006). *No child left behind.* Retrieved February 16, 2008, from http://www.ed.gov/nclb/landing.jhtml

United States Department of Education. (2007). *Building on results: A blueprint for strengthening the no child left behind act.* Washington, DC: Author.

United States Department of Education. (2008). *Institute of education sciences, national center for education statistics.* Retrieved December 13, 2008, from http://nces.ed.gov/nationsreportcard/states/profile.asp

Valencia, S. W., & Buly, M. R. (2004). Behind the test scores: What struggling readers really need. *Australian Journal of Language and Literacy*, *27*, 217-231.

Wassermann, S. (2001). Quantum theory, the uncertainty alchemy of standardized testing. *Phi Delta Kappan*, *83*, 28-40.

Webb, P. B. (2005). *The association between the scores on the ACT test and Tennessee's value-added assessment in 281 Tennessee high schools*. Unpublished doctoral dissertation, East Tennessee State University.

Wyett, J. L. (1998). John Dewey & Earl Kelley: Giants in democratic education. *Education*, *151,* 73-78.

Wyman, B. F. (2000). Decentralization continued: A survey of emerging issues in site-based decision making. *Journal of Law and Education*, *29*, 255-263.

Young, K. (2006). *Standardized testing.* Retrieved February 3, 2008, from http://www.msn.edu/~youngka7/history.html

Zucker, S. (2003). *Fundamentals of standardized testing.* San Antonio, TX: Harcourt Assessment.

VITA

JULIA W. LAMONS

Personal Data:	Date of Birth: September 28, 1981 Place of Birth: Greeneville, Tennessee Marital Status: Married
Education:	Tusculum College, Greeneville, TN; B.A. in Elementary Education 2003 M.A. in K-12 Education 2005
	East Tennessee State University, Johnson City, Tennessee; Educational Leadership and Policy Analysis, Ed.D.; 2009
Professional Experience:	Teacher McDonald Elementary 2003-2005
	Mosheim Middle 2005-present
Professional Achievements:	Chair School Improvement Team Leadership Committee Assessment Coach Wal-Mart Distribution Center Teacher of the Year 2008