5-2005

A Comparison of Academic Achievement and Value-Added Grades on the State Report Cards in Tennessee, 2001-2003.

Kyle Allen Evans
East Tennessee State University

A Comparison of Academic Achievement and Value-Added Grades on the State Report Cards in Tennessee, 2001-2003.

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 in Education

---

by
Kyle A. Evans
May 2005

Dr. Hal Knight, Chair
Dr. Terrance Tollefson
Dr. Edward Dwyer
Dr. Russ West

Keywords: standardized testing, state report cards, value-added assessment

ABSTRACT

A Comparison of Academic Achievement and Value-Added Grades on the State Report Cards in Tennessee, 2001-2003

by

Kyle A. Evans

This study uses the state report cards published by the Tennessee Department of Education to compare the academic achievement grades to the value-added grades to determine if there is a relationship between the two grading systems. The data used for this study are from the 2001, 2002, and the 2003 state report cards published for each school using the five subject areas of reading, language, math, science, and social studies.

One thousand sixty schools in the state of Tennessee were for this study. The socioeconomic status (SES) of the schools was used as a covariate to determine if the socioeconomic status of the school has an effect on the relationship between achievement grades and value-added grades on the state report card. Schools were grouped into one of three categories. Schools with 0 – 33% of their students eligible for free/reduced meals were categorized as an upper SES school. Schools with 34-66% of their students eligible for free/reduced meals were categorized as a middle SES school, and schools with 67-100% of their students eligible for free/reduced meals were categorized as a lower SES school.

The data used to determine the grades on the report card are based upon the results of the state mandated achievement tests given in the state of Tennessee. The cumulative three year averages of the normal curve equivalent scores (NCEs) are used to determine the achievement grades while the cumulative three-year value-added percentages are used to determine the value-added grades on the state report cards.

There was a statistically significant relationship between academic achievement grades and value-added grades in math, language, and social studies on the 2001, 2002, and 2003 state report cards. In reading, the 2002 state report card did not show a significant relationship between the grades while the 2001 and 2003 report cards did indicate a significant relationship. In science, the 2001 and 2003 report cards did not indicate a significant relationship between achievement and value-added grades while the 2002 report card did indicate a significant relationship between the grades.

CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The educators who operate America's public schools, along with policy makers and the general public, have been increasingly preoccupied with students' scores on standardized achievement tests (Popham, 2001). Usually created by commercial test publishers, standardized tests are designed to give a common measure of students' performance (Bagin & Rudner, 1994). Standardized testing achieves standardization by using precise instructions for administration, standard formats for tests and recording responses, and typically include machine scoring of multiple choice questions (Sanders & Horn, 1995). Just as medical tests help diagnose and treat patients physically, rigorous and meaningful educational assessments could help ensure the academic health of students (Gandal & McGiffert, 2003).

There are several reasons for the great appeal of standardized testing to policy makers as an agent of reform. First, standardized tests and assessments are relatively inexpensive compared to other reform measures. Second, it is easier to mandate testing and assessment requirements at the state or district level than it is to take action to change what happens in the classroom (Linn, 2000).

In recent decades, there has been an increase in standardized testing and the use of the scores from the tests as a way to hold schools accountable. According to the American Educational Research Association, spending on K-12 tests among the 50 states nearly doubled from $165 million in 1996 to $330 million in 2000 (McAdams, 2002). The use of test scores has become the primary source of data used to evaluate schools and their accomplishments (Amrein & Berliner, 2002). It is not uncommon for a school within a district to be publicly labeled as "exemplary" or "low performing," solely based upon test scores with rewards and sanctions that follow such designations (Perreault, 2000).

There are two sides to the debate regarding standardized testing. On the one side are those who see standardized testing as the only legitimate way of determining how schools and students measure up, while on the other side are those who say no single test can adequately gauge the performance of an individual student or school (Owens, 2002). Those who support high stakes testing make the claim that teachers need to be held accountable and the test results can be used to improve student learning and design better professional development for teachers (Amrein & Berliner, 2002). Opponents of high stakes testing argue that it encourages schools to "teach to the test," thereby improving results without improving learning (Greene, Winters, & Forster, 2003).

In the State of Tennessee, students in grades three through eight take achievement tests as part of the Tennessee Comprehensive Assessment Program (TCAP). "The TCAP Achievement Test has fresh, non-redundant test items and is customized yearly to measure academic basic skills in reading, language arts, mathematics, science, and social studies" (Tennessee Department of Education, 2004, p.1).

Each year the Tennessee Department of Education publishes a report card for the state and for each public school system and school in the state. The report card provides a summary of performance on academic and non-academic measures using letter grades (Pruett, 2002). The academic information for grades three through eight is based upon cumulative three-year averages in two areas for each of the five subject areas on the state achievement tests administered in grades three through eight. The first area is academic achievement, which is based upon the Normal Curve Equivalent (NCE) average for each school and school district. The second area is based upon the average value-added growth for each subject area at the school and district level. Instead of comparing students to each other or to an established level of proficiency, value-added assessment compares students using their score on the previous years'

tests in order to determine if they are advancing academically (Hellend, 2001). A database containing the merged records of all students in Tennessee who have taken the achievement tests during the past three years is maintained (Baker, Xu, & Detch, 1995). By following the progress of each individual student, this process filters out most of the socioeconomic factors that people have worried about during previous efforts to use student achievement data in assessment and evaluation (Sanders, 1998).

Because Tennessee uses this dual reporting system on its state report card to evaluate its schools, this study will examine the grades in academic achievement and value-added for each school in the state containing grades four through eight using the five subject areas that are assessed on the achievement tests. The grades from the 2001, 2002, and 2003 report cards will be used for this study in order to determine if there is a relationship between the achievement grades and value-added grades for each of the five subject areas on the achievement tests.

Background of the Problem

The state report card in Tennessee for schools that contain grades three through eight is based upon two very different rating scales. Schools are given grades in five subject areas of academic achievement which is based upon the students' average achievement level and for value-added which is based upon the students' amount of growth from previous years.

If a school has high grades in academic achievement, does it have less potential for growth causing the school to have lower grades in value-added? Also, does a school with lower grades in academic achievement have more potential for growth creating higher grades in value-added on the state report card? In many urban schools, Sanders has noticed a pattern in which students with the lowest past performance make the greatest gains, but those who start with higher scores make little headway. A graph of

such gains against past performance creates a downward sloping line from left to right. He calls these “shed patterns,” referring to the sloped roof of a tool shed (Archer, 1999).

According to statewide aggregate data, the evidence suggests at the highest levels of achievement that students exhibited somewhat less academic growth from year to year than their lower achieving peers (Sanders, 1998). Lower achieving students are the first to benefit as teacher effectiveness improves. With many exceptions, higher achieving students do not have the opportunity to demonstrate academic growth at the same rate as lower achieving students (Pipho, 1998). On statistical grounds alone, it is easier to elevate the performance of low scoring students than it is to elevate the scores of any other group of students (Popham, 2001).

Purpose of the Study

The purpose of this study is to compare academic achievement and value-added grades on the state report card for each of the five subject areas of reading, math, language arts, science, and social studies in order to determine if there is a significant relationship between the two rating scales. The data for this study were obtained from the Tennessee Department of Education web-site using the grades from the 2001, 2002, and 2003 state report cards published for each school in the state. All schools containing grades four through eight and have complete information in all five subject areas for all three years were included in this study.

Schools that have 0-33% of its students eligible for free/reduced meals were coded as an upper socioeconomic (SES) school. Schools that have 34-66% of its students eligible for free/reduced meals were coded as a middle socioeconomic (SES) school and schools that have 67-100% of its students eligible for free/reduced meals were coded as a lower socioeconomic (SES) school. “It is well documented that economically disadvantaged and minority students score significantly lower on standardized tests” (Orfield & Wald, 2000, p38). According to several reports,

especially the Coleman Report, the level of student achievement is strongly related with family and community variables (Bracey, 2004). As a result, schools in Tennessee that serve a high percentage of lower socioeconomic students would have more difficulty obtaining a good grade in academic achievement on the state report card as opposed to a school that serves a lower percentage of economically disadvantaged students.

Research Questions

The following questions and the related hypotheses were used to guide this study:

Question 1

Is there a significant relationship between academic achievement and value-added grades in reading on the 2001, 2002, and 2003 state report cards?

H1: There is no statistically significant relationship between academic achievement and value-added grades in reading on the 2001, 2002, and 2003 state report cards.

Question 2

Is there a significant relationship between academic achievement and value-added grades in math on the 2001, 2002, and 2003 state report cards?

H2: There is no statistically significant relationship between academic achievement and value-added grades in math on the 2001, 2002, and 2003 state report cards.

Question 3

Is there a significant relationship between academic achievement and value-added grades in language on the 2001, 2002, and 2003 state report cards?

H3: There is no statistically significant relationship between academic achievement and value-added grades in language on the 2001, 2002, and 2003 state report cards.

Question 4

Is there a significant relationship between academic achievement and value-added grades in science on the 2001, 2002, and 2003 state report cards?

H4: There is no statistically significant relationship between academic achievement and value-added grades in science on the 2001, 2002, and 2003 state report cards.

Question 5

Is there a significant relationship between academic achievement and value-added grades in social studies on the 2001, 2002, and 2003 state report cards?

H5: There is no statistically significant relationship between academic achievement and value-added grades in social studies on the 2001, 2002, and 2003 state report cards.

Question 6

Is there a significant relationship between academic achievement and value-added grades in lower SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards?

H6: There is no statistically significant relationship between academic achievement and value-added grades in lower SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards

Question 7

Is there a significant relationship between academic achievement and value-added grades in upper SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards?

H7: There is no statistically significant relationship between academic achievement and value-added grades in upper SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards.

Question 8

Is there a significant relationship between academic achievement and value-added grades in middle SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards?

H8: There is no statistically significant relationship between academic achievement and value-added grades in middle SES schools in any of the five subject areas on the 2001, 2002, and 2003 state report cards.

Significance of the Study

In 2003, 70% of the elementary schools in the state received failing grades in language arts on the state report card and 33% received failing grades in reading based on value-added scores (Olson, 2004). With so much emphasis placed upon evaluating the quality of the schools, it is important to make sure the process of evaluating schools is a fair and equitable process. For example, do schools that have a large percentage of higher achieving students have less potential for value-added growth as opposed to a school with a large percentage of lower achieving students creating an unfair rating scale. Tennessee is the state most strongly identified throughout the nation with value-added assessment (Evergreen Freedom Foundation, 2002). "The Tennessee value-added assessment system (TVAAS) was developed on the premise that society has a right to expect schools to provide students with the opportunity for academic growth regardless of the level at which students enter the educational venue" (Sanders & Horn, 1995, p12).

For many people, one of the primary considerations when moving to a neighborhood is the quality of the schools. According to a study by Amrein and Berliner, test scores have been shown to affect housing prices, resulting in a difference of about $9,000 between homes in grade "A" or grade "B" neighborhoods (2002, p3). It is extremely important for people seeking a quality school within their neighborhood to

get accurate information regarding the surrounding schools so any method of rating and evaluating schools should be fair and precise.

Definition of Terms

Academic Achievement

This score on the state report card is calculated using the three year cumulative normal curve equivalent (NCE) score from the state mandated achievement test for each of the five subject areas. The minimum score or expectation (average) is that the average score for a school or school system will be at the national average. Higher or lower performance is rated accordingly (Tennessee Department of Education, 2002).

High-Stakes Test

High-stakes testing means that the consequences for good (high) or poor (low) performance on a test are substantial. In other words, some very important decisions, such as promotion or retention, entrance into an educational institution, teacher salary, or a school district's ranking depend on a single test score. Tests that have no specific decision tied to them can become high-stakes to teachers and school administrators when they must face public pressure after scores are made public.

Normal Curve Equivalent (NCE)

Normal Curve Equivalent (NCE's) are represented on a scale ranging from 1 to 99. NCE's have some of the characteristics of percentile ranks but have the advantage of being based on an equal interval scale. That is, the difference between any two successive scores on the scale has the same meaning throughout the scale. Theoretically, because of this characteristic, a student's NCE on different tests can be compared. NCE's obtained by different groups of students on the same set of tests can also be compared by averaging the scores for the groups (CTB/McGraw-Hill, 2002).

Norm Target Gain

The norm target gain is the amount of scale score growth needed to show a years growth and is used to calculate the value-added score for each subject area. The amount of scale score growth varies for each subject area at each grade level because each subtest is scaled separately.

Scale Score

The scale score describes growth on a continuum of achievement that typically occurs as a student progresses through school. The use of a single scale for each content area across all tests and all grade levels enables direct comparisons over time. Plotting the mean (or average) scale scores for each grade, within a school or school district, for successive test administrations, as well as the mean scale scores for the norm group allows relative academic growth to be tracked to help educational planning. However, because the test content areas are scaled separately, tests in one content area cannot be compared with tests in another. For example, a scale score of 468 on a reading test would not have the same meaning as a scale score of 468 on a mathematics test (CTB McGraw-Hill, 2002).

Socioeconomic Status (SES)

The socioeconomic status for a school is determined by the percentage of its students that are eligible to receive free/reduced meals. For the purpose of this study, schools will be grouped into one of three categories of socioeconomic status. The categories are upper, middle, or lower based upon the percentage of students eligible for free/reduced meals.

Upper SES Schools

Schools that have from 0 -33% of its student population receiving free/reduced meals were coded as an upper SES school.

Middle SES Schools

Schools that have from 34-66% of its student population receiving free/reduced meals were coded as a middle SES school.

Lower SES Schools

Schools that have from 67-100% of its student population receiving free/reduced meals were coded as a lower SES school.

TerraNova

The TerraNova test is a nationally normed standardized test that includes subtests in the five subject areas of reading, math, language arts, science and social studies. Standardized testing procedures (with exact directions, time limits, and scoring criteria) ensure that testing conditions are the same for all students.

Value-added Assessment

"Value-added assessment is any method of analyzing student test data to ascertain students' growth in learning by comparing students' current level of learning to their own past learning" (Evergreen Freedom Foundation, 2002, p3).

Delimitations and Limitations

The schools included in this study are located in the State of Tennessee serving students in grades four through eight. The information for this study was obtained from the state Department of Education web-site using data from the 2001, 2002, and the 2003 state report cards. In order for schools to have information on this web-site, the school had to have complete information and the school must have been in existence for at least three years.

Schools that do not have complete information on the 2001, 2002, and the 2003 state report cards were not included in this study. Schools that do not contain grades four through eight in their student population are not included in this study.

Overview of the Study

This study is organized into five chapters:

Chapter 1 contains the introduction, statement of the problem, purpose of the study, research questions, definitions, and overview of the study.

Chapter 2 presents a review of selected related literature.

Chapter 3 describes the methods and procedures by which the study was conducted.

Chapter 4 contains the statistical treatment of the data.

Chapter 5 includes the summary, findings, conclusions, and recommendations of the study.

CHAPTER 2

REVIEW OF LITERATURE

Introduction

The wide spread use of standardized testing can be traced back to the time prior to the United States involvement in World War I.

> "During this time the army had to field a fighting force of immense size so they developed a test known as the Army Alpha. The Alpha represented the first truly large scale use of multiple choice test items, and its items were subjected to all sorts of statistical analyses. It determined where each new test taker stood in relation to a collection of previous Alpha test takers, known as the norm group" (Popham, 2001, p41).

Since that time, tests of various sorts have determined which immigrants could enter the United States, who could serve in the armed forces, who was gifted, and who may be in need of special education services in public schools (Amrein & Berliner, 2003).

The Effects of Standardized Testing in Public Schools

The increased use of standardized testing in public schools began in the late 1980s and increased throughout the decade of the 1990s. This increase was primarily due to the release of a report in 1983 during the Reagan administration titled "A Nation at Risk" (Amrein & Berliner, 2002). "This report was part of a centered campaign and was based on exaggerated and misleading evidence that stirred widespread concerns about schools and consequently demands for more testing"(Kohn, 2000,p3). Despite its lack of scholarly credibility, the report produced massive effects. The National Commission on Education recommended that states institute more rigorous standards and accountability measures to hold schools accountable for meeting those standards (Amrein & Berliner, 2002).

According to a 2001 Quality Counts study conducted by *Education Week*, all states test students however, not all the tests are correlated to the state standards. In 27 states, there is some type of system to hold schools accountable for results by rating performance or identifying low performing schools, 11 states identify low performing schools solely on the basis of test scores, 15 states test every student in reading and math in at least every grade from three through eight (McAdams, 2002).

On January 8, 2002, President Bush signed into law the *No Child Left Behind* Act (NCLB). This new law represented the President's education reform plan and contained the most sweeping changes to the Elementary and Secondary Education Act since it was enacted in 1965 (Tennessee Department of Education, 2003). As a part of the requirements of the NCLB, schools in every state must assess students annually in reading and math in grades three through eight and again before they graduate high school (Neil, 2003).

> "This requirement must be met by the 2005 – 2006 school year and science assessments in key grades will follow in the 2007 – 2008 school year. Due to these requirements, at least 36 states will have to develop more than 200 new tests within the next few years to be in compliance with the federal law" (Gandal & McGiffert, 2003, p39).

The law does not specifically mandate standardized tests, so a few states plan to use local assessments, including classroom-based information, rather than state exams. However, standardized achievement test scores are the near-exclusive means for determining adequate yearly progress, and in most cases standardized test score results are believed to be essential to monitor individual student progress and school accountability (Bleim & Shepard, 1995).

According to the NCLB, students must reach the proficient level on state assessments by the 2012 school year. In order to accomplish this goal, states will need to move an additional 4% to 6% of their students into the proficient category every year.

Using the National Assessment of Educational Progress (NAEP) tests as a measure, only three of 33 states made even 1% gains in reading per year from 1992 to 1998 (Neil, 2003).

Most of the standardized tests used in schools are multiple choice, and a multiple choice test simply does not measure the same cognitive skills as are measured by similar problems in free-response form (Kohn, 2000). Take for instance the Graduate Record Exam (GRE), a timed, multiple choice test that most of the nation's graduate programs require for admission or financial aid. University graduate departments continue to depend on the GRE to predict a candidate's chances of success even though a student's score on the exam has no relationship whatsoever to his or her performance in graduate school (Sacks, 2000). Until there is agreement on what basics make up the ideal learner, it will be difficult to consider the best assessment approaches in order to know if schools are succeeding (Graves, 2002).

> "Americans are fascinated with mental measurement to a degree that is rare in other countries. The examinations for college or university admission in other industrial countries are typically essay tests, in which students demonstrate knowledge of various subjects they have learned in the classroom. These tests are not unlike what American educators call performance assessments. Compared to other countries, Americans appear to be far more obsessed with the notion that intelligence is both inborn and represented as a single numerical score" (Sacks, 2000, p7).

The current measurement of students' achievement focuses everyone's attention on student achievement. Superintendents, principals, and teachers now spend more time trying to link the structure and work of the organization to student learning (McAdams, 2002). Teachers feel strong pressure, especially from district administrators and the media, to improve their test scores (Hermon & Golan, 1998). Teachers believe they spend an inordinate amount of time on drills leading to the memorization of facts

rather than spending time on problem solving and the development of critical thinking skills (Amrein & Berliner, 2002).

Using high-stakes tests to drive instruction will not improve schools. It is well known that test preparation, alignment of the curriculum with the test, as well as rewards and sanctions for students and other school personnel, will almost always result in gains on whatever instrument is used by the state to assess its schools (Amrein & Berliner, 2002). Scores on standardized tests can increase by narrowing the curriculum and focusing on only the subjects that are tested at the expense of other subjects such as art, music, and physical education that are not tested. The National Research Council cautions "An assessment should provide representative coverage of the content and process of the domain being tested, so that a score is a valid measure of the student's knowledge of the broader domain, not just the particular sample of items on the test" (Amrein & Berliner, 2002, p16). In an interview with *Family Education Magazine,* Monty Neil, director of Fairtest said, "What people forget is that a test score is just an estimate, and the error rate on most tests is pretty sizeable" (Rayburn, 2003, p266). Educational Testing Service leaders strongly warn against using test scores alone to make high stakes judgments about students (Orfield & Wald, 2000).

> "Current tests, which rely heavily on computer scoring, will continue to fail to measure what American schools should prize in their students, in order to maintain the number one position in the world. Such elements as initiative and the ability to formulate questions, relate and integrate sources, and engage in good, long thinking gets lost in the rush to measure quickly and cheaply" (Graves, 2002, p32-33).
>
> "A National Academy of Science/National Research Council report on school learning makes clear, schooling that closely resembles training, as in preparation for testing, cannot accomplish the task the nation has set for itself, namely the

development of adaptive and educated citizens for this new millennium" (Amrein & Berliner, 2002, p15).

"A variety of factors, such as teaching that is narrowly focused on the specifics of the assessments rather than the content standards the assessments are intended to measure, may undermine the validity of desired generalizations" (Kuppermintz, 2002, p13).

"Many things students learn simply cannot be tested with a paper and pencil test. In a high quality education, students conduct science experiments, solve real-world math problems, write research papers, read and analyze novels and stories, deliver oral presentations, evaluate and synthesize information from a variety of fields, and apply their learning to new situations. Standardized paper and pencil tests are poor tools for evaluating these important kinds of learning and the main purpose of these tests is to sort large numbers of students in as efficient manner as possible" (Neil, 2003, p44).

Graves, Kuppermintz, Neil, Amrein, and Berliner make strong arguments regarding the negative impacts of standardized testing on public education.

Only a few states such as Kentucky, Vermont, and California have taken steps to eliminate traditional multiple-choice tests but in almost all cases even those moves proved to be short-lived (Sacks, 2000). "Low cost, ease and consistency of scoring, and a mature industry of testing companies offering a comprehensive menu of services for administering, processing, scoring, analyzing, and reporting test results ensure the privileged status of multiple choice tests" (Kuppermintz, 2002, p12).

"Concepts such as intrinsic motivation and intellectual exploration are difficult for some minds to grasp, where test scores, like sales figures or votes, can be calculated and tracked and used to determine success or failure" (Kohn, 2000, p3).

"Student learning and development of academic proficiencies is a highly complex process, shaped and influenced by a multitude of factors such as, personal

characteristics (both cognitive and non-cognitive), physical and mental maturation, home environment, cultural sensitivities, institutional and informal community resources, and of course, the formal process of schooling" (Kuppermintz, 2002, p3).

"Standardized achievement tests should not be used to evaluate the quality of students' schooling because the quest for wide score spread tends to eliminate items covering important content that teachers have emphasized and students have mastered. These items are removed when the test is revised in order to increase the test reliability and create a better score spread" (Popham, 2001, p48).

Kuppermintz, Popham, and Kohn explain several reasons why standardized tests are not good tools to use for evaluating the quality of a child's education. In addition, test designers must constantly refresh the test questions, but the new items are never precisely comparable to the old ones. That is why test designers publish the margins of error expressed as "reliability coefficients" between 0 and 1 (Sacks, 2000).

In his book, *Testing is Not Teaching,* Graves states "standardized tests have a built in failure rate of roughly 18%" (Graves, 2002, p21). The National Research Council was charged by Congress to do a study on the use of standardized test scores for so called high stakes purposes and found that while testing can yield valuable information about a student's achievement, the nature and limitations of that information is widely misunderstood and warned that high stakes decisions of any kind should not be made on the basis of a single test score (Amrein & Berliner, 2002). For this reason, the standards for the measurement profession warn against using the results of any single test as the basis for making major decisions at the classroom, school, and district level (Neil, 2003). Robert Swartz, President of ACHIEVE, states "Common sense suggests that states should not rely solely on the results of one-shot assessments" (Orfield & Wald, 2000, p2).

Standardized testing has a negative impact on education, and in some cases efforts to raise scores have distorted curricula to such an extent that reading and math have been emphasized and science and social studies excluded (Bleim & Shepard, 1995). Preoccupied with raising test scores in order to satisfy parents, principals, and state legislators, schools have often neglected reforms that would promote deeper and more active ways of thinking and learning than standardized tests typically measure (Sacks, 2000). The Heisenberg's Uncertainty Principle states "the more important any quantitative social indicator becomes in social decision making, the more likely it will be to distort and corrupt the social process it is intended to monitor" (Amrien & Berliner, 2002, p5). This principle warns that attaching serious personal and educational consequences to performance on tests for schools, administrators, teachers, and students may have a negative impact on education.

A meaningful amount of what is measured by today's high-stakes tests is directly attributable not to what students learn in school but what they bring to school in the form of their families' socioeconomic status or the academic aptitude they develop (Popham, 2001). "Call it the Volvo Effect, a good guess could be made about a child's standardized test scores by simply looking at how many degrees his or her parents have and what kind of car they drive" (Sacks, 2000, p3).

The amount of poverty in communities where schools are located, along with other variables not impacted by what happens in the classroom, accounts for the majority of the difference in test scores from one area to the next (Kohn, 2000). Schools serving low-income communities start with students who are less academically ready, have greater social needs, and receive less academic support at home. An overemphasis on testing will undermine the ability of schools to ensure a high quality academic and social experience for all their students which is what happens when you combine the limitations of standardized tests with unreasonable pressures upon schools without resources to solve long-term societal inequities (Neil, 2003).

For decades, critics have complained that many standardized tests are unfair because the questions require a set of knowledge and skills more likely to be possessed by children from a privileged background (Kohn, 2000). Affluent children who attend higher scoring schools will continue to receive an education that prepares them for college, high paying careers, and significant social and political influence.

> "Sadly, when score gaps between poor and affluent schools close some people will be fooled into thinking that education is being equalized. In reality, lower income students will rarely be taught the many attributes of success in college and life that cannot be measured by tests" (Neil, 2003, p44).

Teachers from higher socioeconomic schools report an increase over the last three years in instructional time devoted to higher level thinking skills. Teachers in lower socioeconomic schools do not report such an increase and low socioeconomic schools report giving substantially more instructional attention to test content through planning and delivery of instructional programs than do higher socioeconomic schools (Hermon & Golan, 1998).

"One fairly certain way of telling whether a high-stakes test is a winner or loser is to determine if unexciting drill activities can actually raise students' test scores. If so, the test is almost certainly inappropriate measuring only low level outcomes" (Popham, 2001, p21). Standardized tests include only a few questions on any particular topic and they provide too little information to produce accurate, comprehensive, or detailed results or analysis (Neil, 2003). "We've got very interesting studies where teachers do 35 or 38 weeks of what they think is best for kids, and then they will give them three weeks of test cramming just before the test, and the kids do just as well as kids who have 40 weeks of test driven curriculum" Harvey Daniels (as cited in Kohn, 2000, p52).

High-stakes testing assumes that rewards and consequences attached to rigorous tests will motivate the unmotivated to learn (Orfield & Wald, 2000). Eighteen states currently use exams to grant or withhold diplomas. In the 18 states, looking at

SAT, advanced placement tests, and National Assessment of Educational Progress (NAEP), a strong case can be made that high-stakes testing policies have resulted in no measurable improvement in student learning (Amrein & Berliner, 2003).

The worst tests are timed so a premium is placed on speed as opposed to thoughtfulness and even thoroughness and a test that is norm referenced can only tell that one student is more or less proficient than another but does not tell how proficient either of them is with respect to the subject matter. Tests that are given every year assume that all students learn at the same pace, and standardized tests should never be given to young children because it is difficult for them to communicate a depth of understanding (Kohn, 2000).

If the nation is to leave no child behind, schools must comprehensively address poverty and its consequences and support higher-quality educational practices. More than standardized test scores should be used to determine whether schools are improving and students are learning. Continuing the course of high-stakes testing will only deepen the crisis in schools serving the most vulnerable children creating terrible consequences on their communities in the future (Neil, 2003).

Value-Added Assessment

When the *Education Improvement Act* was passed in the State of Tennessee in 1992, it was initiated by a state supreme-court order to make Tennessee's school financing system more equitable. Legislators, under pressure from businesses, adopted a strong accountability model, now known as TVAAS, requiring evidence of satisfactory year-to-year improvements in student achievement down to the classroom level (Kuppermintz, 2002).

Even though value-added assessment is used in other areas, Tennessee has the most comprehensive value-added system in the country and is the only state so far to put in place a statewide process of gathering information needed to determine the

effects of teaching on students' academic growth (Carey, 2004). The Amenberg Foundation is supporting value-added analysis in Florida and in the Washington area where school officials are measuring each school against its past performance (Matthews, 2000). Other states with school districts using value-added analysis of schools and/or teachers include Colorado, Minnesota, North Carolina, and Wisconsin (Carey, 2004).

Value-added assessment represents a variety of technologies from many different academic areas that build upon the statistical advantages of mixed model theory and methodology (Sanders, 1998). It was developed by William Sanders, who at the time was a professor at the University of Tennessee. While doing statistical analysis for agricultural research scientists, Sanders states, "those people were constantly trying for better ways to take performance data and to better partition genetic influences from environmental influences, such that they could improve breeding efficiencies of plants and animals" (Archer, 1999, p27). Sanders focuses not on one set of test results but on how the scores change over time, and he contends that by looking at a student's test score gain or loss from the previous year, the role played by the classroom teacher can be determined (Matthews, 2000)

In the early 1980s, Sanders began to explore the feasibility of combining techniques and measures of student achievement to evaluate school influences on student data (Baker, Xu, & Detch, 1995). Sander's first attempt to prove his work in the early 1980s went nowhere when Lamar Alexander, who was the Governor of Tennessee at that time, was seeking to award merit pay for teachers (Archer, 1999). In 1984, Sanders used three years of test data from the Knox County School System to test his value-added concept. The study yielded estimates of teacher effectiveness that were relatively consistent from year to year, and school administrators confirmed the data were consistent with their own impressions of which teachers were most effective (Baker et al., 1995).

In 1996, Sanders used data from two Tennessee school districts and divided their teachers into five groups ranging from the least effective to the most effective based upon their achievement test score results (Sanders, 1998). He found students who had been taught by three of the least effective teachers in a row scored below the 50th percentile on math by the end of the third year. By contrast, those who had three highly effective teachers scored at the 80th percentile or above in math by the end of the third year (Archer, 1999). The differences separating teachers cannot be attributed to differences in students, because value-added systems isolate the teacher's impact by controlling prior student achievement and other factors. Value-added data show that, even in the same schools or districts, even with students whose prior achievement was similar, some teachers get great gains while others allow achievement to lag (Carey, 2004). Sanders argument is an effective teacher can produce improvement in any student, low income or affluent (Matthews, 2000).

Value-added analysis is so named because it seeks to answer one question. "How much value has a school added to a student's learning?" (Doran, 2003, p57). Value-added assessment is statistically robust but the validity of its results depends upon certain preconditions. Those conditions include the requirement of annual testing in consecutive grade levels, items used in each annual test must be fresh and non-redundant, the scoring must be tied to an underlying linear scale, and the scores are reported in a common scale (Evergreen Freedom Foundation, 2002).

Other accountability plans rely on simple statistics, such as group averages, and are status-based accountability systems which can provide misleading and invalid results for the following reasons.

> First, test scores are subject to external variables such as the economic status of a student's parents. Second, it punishes schools serving disadvantaged populations. Third, a test score for an eighth grade student is invalid for evaluating eighth grade instruction because it reflects the cumulative impact of

schooling over all previous years. Fourth, the cut score categories such as basic and proficient are gross measures like measuring the height of a child with a yardstick, but only acknowledging growth when he or she has exceeded thirty-six inches of growth (Doran, 2003, p56-57).

After determining the overall gain in student achievement, TVAAS produces a measurement of teacher effectiveness by comparing the actual growth in student learning to the expected growth (Sanders & Horn, 1995).

> "The expected growth level is created by starting with the normal amount of academic progress that a typical student is expected to make in a given subject and grade, and then using statistical controls to adjust that anticipated progress up or down, based upon the previous achievement history of each student. If a teacher has a student that has previously struggled to make academic progress over a number of years-because of motivation, aptitude, family life, or whatever the reason, then the amount of growth that a teacher is expected to help that student achieve is adjusted down accordingly" (Carey, 2004, p5).

This has the effect of screening out whatever non-teaching factors affect student learning and isolating the individual teacher's contribution (Bratton, 1998). The TVAAS statistical model aggregates student growth increases using a design that accommodates for missing data (Evergreen Freedom Foundation, 2002).

> In order for a student to be included in the individual teacher's value-added report, the student must be in attendance a minimum of 150 days during the school year. For teachers who have taught a given student less than a full year, only those who have been the teacher's responsibility for more than 75 days are counted. Students who receive special education services are not included in the teacher's individual value-added information (Tennessee Code Annotated 49-1-606, 1996).

The value-added system provides evidence to suggest that the single largest factor affecting the academic growth of students is the difference in teacher effectiveness (Carey, 2004).

> "Using a massive database of student test scores from thousands of schools and hundreds of thousands of students in Texas, researchers analyzed the math performance of individual students over time, calculating the effect of individual teachers on how much students learn. The conclusion: teacher effectiveness varied dramatically and had a major impact on student performance, so much so that having a high quality teacher throughout elementary school can substantially offset, even eliminate, the disadvantage of a low socioeconomic background" (Rivkin, Hanushek & Kain, 2002, p5).

"Of all the factors we study, class size, ethnicity, location, and poverty, they all pale to triviality in the face of teacher effectiveness" (Long & Cass, 2001, p1). Sanders also states,

> "I believe that school districts, schools, and individual teachers should never be held for solving all of society's problems, but I believe equally strongly that the educational community is responsible for taking each kid as they found that kid and allowing each student, each year, to make academic progress from where he or she is" (Archer, 1999, p3).

Much of the frustration does not center on the value-added method itself but on a step Sanders uses to determine if the scale used to weigh the easiness or difficulty of individual test items is equivalent from year to year so that test results are comparable from one year to the next. That is important, because the value-added system averages scores over three years and if the scale is inconsistent, Sanders makes an adjustment (Olson, 2004). Sanders states "his contract with the Tennessee Department of Education requires him to make such adjustments when the test scales are not consistent from year to year" (Olson, 2004, p9).

Sanders compares value-added to measuring a child's height at age two, three, and four. They may not grow the same amount each year. One child may show more growth during one year as opposed to another (Sanders, 1998). Systems, schools and teachers who do best under value-added assessment are those who provide academic growth opportunities for students at every academic level. TVAAS data have shown that some schools have successfully addressed the needs of all students as evidenced by their ability to consistently show normal, and sometimes exceptional, academic progress for students of all academic abilities (Sanders, 1998). According to Sanders, over the last decade about 40% of districts have made reasonable progress, 50% have stayed the same and the rest have lost ground (Matthews, 2000).

People contend standardized tests are imperfect measures of student learning making the tests inappropriate measures of teacher quality. It is true tests aren't perfect and any test covers only a sample of the knowledge and skills a student has acquired (Carey, 2004). Sanders states, "No responsible person claims that any form of assessment can appraise the totality of a student's school experience or even the entirety of the learning that is part of that experience" (Sanders & Horn, 1995, p2). A strong assessment system is essential element for value-added data. "Since the information is derived from test scores, the value is seriously deflated if the tests lack rigor, are overly reductive or don't exhibited steady progress from year to year" (Carey, 2004, p33). Ben Brown, the State Executive Director for Evaluation and Assessment in Tennessee, is quoted as saying "the real dynamic in education is the gain or growth of children and to me that is the only socially, politically and morally correct way to determine accountability for educators" (Matthews, 2000, p3).

State Report Cards

Banks send out statements, doctors perform check-ups, and schools issue report cards. These checks and balances let us know what is happening with our money, our health, and our educational system. Report cards are considered a central feature of state accountability systems by making the assumption a state will improve education by providing the public with better information, spurring low performers into action and inspiring parents to become more involved (Olson, 1999).

An effective report card assumes the system that gathers and produces the data contained in the report card is comprehensive, valid, and reliable (Ananda & Rabbinowitz, 2001; Linn, 2000). States must keep report cards short, simple, and focused on results if parents and other citizens see them as useful, and they must provide comprehensive information if educators are to view them as legitimate (Olson, 1999). Furthermore, the indicators used and reported in the report cards (poverty rates, student demographic data) are sometimes at odds with what indicators the public wants reported (Edwards, 1999). One problem with report cards is that much of the information on them may have little bearing on what schools can do to improve achievement. Report cards are less likely to include data about the school climate, course taking patterns, levels of parent involvement, and the population of teachers with a college major in the subjects they teach even though research has linked these factors with improvements in test scores (Olson, 1999).

> "Research by the Public Agenda Foundation for *Education Week* found that parents in small focus groups wanted information about the quality of life in the school, school leadership, different program offerings, parent and student satisfaction rates, and the levels of parent involvement, among other concerns. Only about one-third of participants in the focus group believed test scores should be used as the main measurement. One half of taxpayers, 25% of educators, and 36% of parents had the same opinion. Parents and other

> taxpayers raised a number of drawbacks to standardized testing, including not trusting the results, concerns that not all children test well, and fears that teachers spend too much time teaching to the test" (Olson, 1999, p6).

The annual school report cards generate much debate among parents, educators, administrators, and students. Although experts believe each school is different, parents demand these report cards while administrators complain they are unfair. The most common complaint from principals is that these report cards do not take into account many other factors. Giving a school credit for advantages children bring to school from home isn't telling the audience much about the quality of the school, and neither is blaming the school for problems the children bring as a result of the impact of lower family incomes (Cohen, 1999).

Because schools can be structured quite differently from one another in their size, the grades they teach, and in the programs they offer, report cards are not a useful tool for ranking schools. It would be incorrect to determine that one school is better than another based on slight differences between data points on a report card. Some critics contend that report cards are often a waste of time and money because they sit on a shelf and gather dust (Olson, 1999).

> "Across the fifty states, school report cards vary tremendously with no two states reporting exactly the same information in the same format. Some report cards are just several pages of statistics with no explanatory text while others are a dozen or more pages with sample test questions and detailed descriptions of what constitutes exemplary performance" (Olson, 1999, p1).

Many states also assign an overall rating to a school's performance based largely on test scores by labeling them with such terms as "acceptable," "unacceptable," or "exemplary". A growing number of states are resorting to this rating system in an effort to get poor performing public schools to improve (Steinberg, 1998). A majority of educators thought using either labels or letter grades to describe how well a school is

doing was a bad idea indicating their general distrust of test scores and other statistical measures for judging schools (Olson, 1999).

The State of Tennessee uses a combination of letter grades and terminology on its report card for school systems and individual schools. On the report card the minimum expectation is considered a "C" or "average" with higher or lower performance rated accordingly. Meeting those standards identified as maximum goals to be attained is considered an "A" or "exemplary" and other ratings are based on a prescribed scale. Value-added gains -- The minimum standard or expectation is that all students will gain a year's average growth (compared to the national norm) for a year's instruction in each subject area. This is expressed as a 100% gain. Higher or lower performance is rated accordingly. Elementary achievement -- The minimum standard or expectation is that average score for a school or school system will be at the national average. Higher or lower performance is rated accordingly (Tennessee Department of Education, 2002).

Summary

One of the driving forces behind the ever increasing use of standardized testing is the *No Child Left Behind Act*. This federal legislation requires all states to develop accountability plans measuring the effectiveness of each public school, and the use of standardized test score information has become the primary source of these data in most cases (Doran, 2003).

Many experts argue that standardized tests are poor tools to measure student learning because the primary intent of these standardized tests are to spread out scores and not to rate instructional effectiveness (Popham, 2001). Standardized tests that use multiple choice answers measure only low level skills and are incapable of measuring the skills students need to be successful in life (Kohn, 2000). Many of the tests are not correlated to state standards and teachers feel compelled to teach to the test due to the

high-stakes surrounding the test results. Also, studies have shown that standardized tests penalize students from disadvantaged backgrounds.

Other experts contend that standardized testing forces accountability upon students, teachers, and school districts and helps to determine gaps in the education of students. Ideally, testing will show which school districts are failing and research will be able to find out the best methods to teach children and those methods can be implemented.

Value-added assessment is a process that measures the influence that systems, schools, and teachers have on the rate of academic growth for populations of students. To accomplish this, value-added uses statistical mixed-model methodology and student scale scores from the norm-referenced component of the achievement test. Value-added was developed on the premise that society has a right to expect that schools will provide students with the opportunity for academic growth regardless of the level at which students enter the educational venue. In other words, all students can and should learn regardless of their ability level (Sanders & Horn, 1994).

All states and school systems are now required to publish report cards with information regarding their progress. The format of these report cards varies among the states with some containing only the mandated information with little explanation while others are very extensive. The purpose of these report cards is to inform parents and the public about the progress of their local schools.

CHAPTER 3

METHODS OF RESEARCH

Introduction

Chapter 3 consists of an explanation of the population and the sample that will be used for this study. It also includes information and tables detailing important concepts that are related to this study. A discussion regarding the validity and the data analysis procedures used for this study is included in this chapter.

Population

The population for this study consists of schools in the state of Tennessee that contain at least one grade of grades four through eight. Data for this study were obtained from the Tennessee Department of Education web-site which contains the report card data for all the public schools in the state.

In order to be considered for this study, schools must have complete data on the 2001, 2002, and the 2003 state report cards. The student enrollment for each school included in the study was obtained from the 2002 state report cards. The total number of students enrolled in 2002 for the schools included in this study was 534,562. The grade span of the schools was coded in order to determine the number of schools with that particular grade span that are included in the study. Any schools not serving at least one grade of grades four through eight or with missing or incomplete data on the state report card were not considered for this study.

The state of Tennessee publishes a report card for each school district and for each school in the district. In order to access the report card information from the state web-page, use the URL www.state.tn.us/education/ and click on the report card symbol. An alphabetical list of all the school districts in the state is listed. Once a particular district is chosen, an alphabetical list of all the schools in that district is available in order to review each individual school's report card.

Each school containing grades four through eight is given a grade in academic achievement and value-added gain for each subject area of reading, math, language arts, science, and social studies on the state report card. The academic achievement grade for each school is calculated using the cumulative three-year average NCE score for each of the five subject areas tested on the state mandated achievement tests.

The value-added grade on the state report card is based upon a complicated formula used to calculate academic growth in each subject area using scale scores from the achievement test. The value-added score is reported as a percentage based upon a predetermined amount of scale score growth. Each grade level and each subject area has a norm target gain of scale score growth needed to show what is equivalent to one year of growth. For example, because 25 is the expected amount of scale score growth needed in fourth grade math to demonstrate one year of growth, a school that averages a scale score growth of 25 in fourth grade math on the achievement test would have a value-added percentage of 100% in math at that grade level. If a school had an average scale score growth of 30 in fourth grade math on the achievement test, its value-added percentage would be 120% for math in fourth grade (Pruett, 2002).

Because value-added percentages measure the amount of growth, one year of data is needed as a baseline in order to calculate the value-added percentage. The state of Tennessee mandates the achievement tests be administered in grades three through eight. For this reason, value-added scores are calculated in schools that contain grades four through eight. The norm target gains of scale score growth for each grade level and each subject area are listed in Table 1 (CTB/McGraw Hill, 2002).

Table 1

Norm Target Scale Score Growth for Each Subject and Grade

Subject	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Math	25	20	18	14	16
Reading	12	13	10	9	8
Language	15	15	4	8	8
Social Studies	12	13	10	11	6
Science	19	16	10	14	10

The state uses a letter grading system on its report card in academic achievement and value-added growth for each subject area. Because the Tennessee Department of Education considers the 50th NCE average and the 100% value-added average as the minimum standard for each school and each school district in the state, a school that has a cumulative three-year NCE average of 50 in any subject area would earn a letter grade of "C" for that subject area on the report card. A school with a cumulative three-year value-added average of 100% in a subject area would receive a letter grade of "C" for that particular subject area on the state report card. A breakdown of the grading scale used by the state of Tennessee to calculate the letter grades on the state report cards is listed in Table 2 (Pruett, 2002).

Table 2

Grading Scale Used for the State Report Cards in Tennessee

Grade	Achievement 3-8 Score Range	Value-Added Percentage 4-8 Score Range
A (Exemplary)	60 - 99	115.0 or above
B (Above Average)	55 - 59	105.0 - 114.9
C (Average)	50 - 54	95.0 - 104.9
D (Below Average)	45 - 49	85.0 - 94.9
F (Deficient)	44 - 1	84.9 - 0.0

Other data taken from the state report cards and recorded for the purpose of this study include each school's name, enrollment, grade span, and socioeconomic status. Each school included in the study was coded as a lower SES school, middle SES school, or a upper SES school. For the purpose of this study, any school with 0-33% of its students eligible for free/reduced meals was coded as an upper SES school, schools with 34-66% of its students eligible for free/reduced meals was coded as a middle SES school while schools with 67-100% of its students eligible for free/reduced meals was coded as a lower SES school. The data for the free/reduced meals information were obtained from the 2002 state report card for each individual school.

Research Design

This is a quantitative study in order to compare the relationship between academic achievement grades and value-added grades on the state report card for schools across the state of Tennessee using three years of report card data. The letter grades on the report cards were considered ordinal data. For each subject and year,

the data were organized into crosstabulated tables to examine the relationship between the achievement and value-added grade.

Validity

In order to increase the validity of this study, schools from across the entire state were considered for this study in order to limit any regional or geographical influences on the results. Three years of report card data were compared independently of each other by subject area in order to account for any major discrepancies that may occur by using one year of report card data. Consideration will be given to the socioeconomic factors that may influence the results by using each school's SES status as a covariate after making an initial comparison. Schools will be grouped into one of three categories of SES schools based upon the percentage of students receiving free/reduced meals.

The items being compared, academic achievement grades and value-added grades, are based upon two different scales used in measuring achievement test results. The grading scale used on the report card to calculate the letter grades for each category was established by the Tennessee Department of Education and could have some influence on the results.

Data Analysis

Frequency tables with counts and percentages for each variable were used to describe the data. In addition, new variables will be created to measure the discrepancy between achievement and value-added grades for each subject by year. These new variables represent the number of letter grades the value-added grade is above or below the achievement grade.

The range of these new variables is -4 to +4. The sign of the variables, positive or negative, indicates whether the value-added grade is higher (positive sign) or lower (negative sign) than the achievement grade. For example, a school that has the same

value-added grade as the achievement grade would receive a score of zero on the new variable. A school with a value-added grade that is one letter grade higher than the achievement grade would receive a score of +1 on the new variable (e.g., a school with a value-added grade of B and an achievement grade of C). In other words, the absolute value of these new variables count the number of letter grades separating the value-added and achievement grades, while the sign of the score indicates whether the value-added grade was higher or lower than the achievement grade.

To evaluate the relationship between achievement and value-added grades, the data were organized into crosstabulated tables, first to examine the relationship between the variables for the total sample, then into three partial tables, one for each level of socioeconomic status. In all crosstablulated tables, the achievement grade will serve as the independent variable, while the value-added grade is the dependent variable.

The Kendall's tau b measure of association was used to measure the strength of the relationship between the two variables of academic achievement grades and value-added grades. The .05 significance level was used to reject or retain the null hypotheses.

CHAPTER 4

DATA ANALYSIS

Introduction

Chapter 4 contains the statistical treatment of the data used for this study in order to answer the research questions from Chapter 1. Several tables are used to organize and explain the data used for this study along with the results of the statistical analysis.

The schools used for this study consists of 1,060 schools across the state of Tennessee. The percentage of students receiving free/reduced meals was used to classify schools into one of three categories of socioeconomic status (SES). Schools that have 0 – 33% of their students eligible for free/reduced meals were coded as an upper SES school; schools with 34-66% of their students eligible for free/reduced meals were coded as a middle SES school; and schools with 67-100% of their students eligible for free/reduced meals were coded as a lower SES school. The SES status of the school was used as a control variable in order to determine its influence on the achievement grades and value-added grades for each subject on the state report card for each of the three years of data. There were 323 lower SES schools (30.5%), 520 middle SES schools (49%), and 217 upper SES schools (20.5%) across the state of Tennessee included in this study.

The mean student enrollment of the schools in this sample was 504.3 with a standard deviation of 229.9. The median student enrollment was 473. The report card information for the schools is available on the Tennessee Department of Education web-site (www.state.tn.us/education/).

A breakdown of the grade span of the schools included in this study along with the number and percentage of schools serving those grade levels is provided in Table 3.

Table 3

Grade Span of Schools

Grade Span	# and	% of Schools	Grade Span	# and	% of Schools
K-9	2	0.2	5	1	0.1
K-8	181	17.1	5-12	1	0.1
K-7	4	0.4	4-8	2	0.2
K-6	127	12.0	4-6	6	0.6
K-5	372	35.1	4-5	2	0.2
K-4	91	8.6	3-8	1	0.1
K-12	26	2.5	3-6	4	0.4
7-9	9	0.8	3-5	12	1.1
7-8	20	1.9	3-4	3	0.3
7-12	3	0.3	1-8	2	0.2
6-8	132	12.5	1-6	2	0.2
6-7	1	0.1	1-5	1	0.1
6-12	3	0.3	5-6	9	0.8
5-8	43	4.1			
Total	1,060	100%			

Analysis of Reading Achievement and Value-Added Grades

The first research question was: Is there a significant relationship between academic achievement grades and value added grades in reading on the 2001, 2002, and 2003 state report cards? Also, this study explores the impact of the socioeconomic status of the schools on the achievement and value-added grades on the state report

cards by categorizing each school into one of three SES categories which are upper, middle, and lower.

A frequency table was computed containing the frequency and percentage of each letter grade for achievement and value-added for each subject area on the schools' state report cards. The grades were obtained from the 2001, 2002, and the 2003 state report cards for each of the schools. The frequency tables for the subject area of reading are provided in Table 4.

Table 4

Reading Achievement and Value-Added Grades

Year/Grade	Reading Achievement		Reading Value-Added	
2001	N	Percent	N	Percent
A	126	11.9	273	25.8
B	182	17.2	219	20.7
C	388	36.6	198	18.7
D	203	19.2	185	17.5
F	161	15.2	185	17.5
Total	1,060	100.0	1,060	100.0
2002				
A	128	12.1	241	22.7
B	210	19.8	218	20.6
C	386	36.4	252	23.8
D	186	17.5	167	15.8
F	150	14.2	182	17.2
Total	1,060	100.0	1,060	100.0

(Table 4 continued)

Year/Grade	Reading Achievement		Reading Value-Added	
2003				
A	130	12.3	203	19.2
B	223	21.0	192	18.1
C	400	37.7	232	21.9
D	156	14.7	169	15.9
F	151	14.2	264	24.9
Total	1,060	100.0	1,060	100.0

As discussed in Chapter 3, new variables were created to measure the discrepancy between value-added and achievement grades. A discrepancy score ranging from +4 to -4 was computed for each school for each subject area on the report card for each year (2001, 2002, and 2003). Positive discrepancy scores indicate the value-added grades were higher than the achievement grades and negative scores indicate the value-added grades were lower than the achievement grades on that subject for that year. The absolute value of the discrepancy score is the number of letter grades between the value-added and achievement grade. For example, if a school had the same letter grade for both value-added and achievement grades, the discrepancy score would be zero. A school with a value-added grade of A and an achievement grade of C would have a discrepancy score of +2 (the value-added grade is two letter grades higher than the achievement grade). Likewise, a school with a value-added grade of F and an achievement grade of A would score a -4 on the discrepancy variable (the value-added grade is four letter grades lower than the achievement grade). The discrepancy scores for reading are provided in Table 5.

Table 5

Discrepancy Scores for Reading (2001, 2002, 2003)

Discrepancy Score	Year 2001		2002		2003	
	N	Percent	N	Percent	N	Percent
-4.00	15	1.4	19	1.8	15	1.4
-3.00	41	3.9	85	8.0	59	5.6
-2.00	120	11.3	148	14.0	130	12.3
-1.00	166	15.7	170	16.0	179	16.9
0.00	249	23.5	275	25.9	233	22.0
+1.00	191	18.0	160	15.1	176	16.6
+2.00	167	15.8	124	11.7	152	14.3
+3.00	79	7.5	51	4.8	79	7.5
+4.00	32	3.0	28	2.6	37	3.5
Total	1,060	100.0	1,060	100.0	1,060	100.0

Note. +Positive scores indicate the value-added grade was higher than the achievement grade. -Negative scores indicate the value-added grade was lower than the achievement grade.

On the 2001 state report card in the subject area of reading, 23.5% (249) of the schools had the same grade on reading achievement and reading value-added while 15.7% (166) of the schools had a reading value-added grade that was one letter grade lower than the reading achievement grade. Approximately one third of the schools 32.3% (342) had reading value-added grades from one to four letter grades lower than their reading achievement grade on the 2001 state report card while 44.2% (469) of the

schools had reading value-added grades from one to four letter grades higher than their reading achievement grade on the 2001 state report card.

On the 2002 state report card, slightly over one fifth of the schools 22% (233) had the same score on reading achievement and reading value-added while 36.1% (383) of the schools had a value-added grade from one to four letter grades lower than their reading achievement grade on the 2002 state report card while 41.9% (444) of the schools had a reading value-added grade from one to four letter grades higher than the reading achievement grade on the 2002 state report card.

Approximately one fourth of the schools 25.9% (275) had the same grade in reading achievement and reading value-added on the 2003 state report card while 39.8% (422) of the schools had a reading value-added grade from one to four letter grades lower than their reading achievement grade. Slightly over one third of the schools 34.2% (363) had reading value-added grades from one to four letter grades higher than their reading achievement grade on the 2003 state report card.

The mean of the discrepancy scores were calculated in order to determine the average difference between the reading value-added grade and the reading achievement grade at each level of SES. The mean discrepancy scores for the subject area of reading on the state report card for the sample of schools used for this study are provided in Table 6.

Table 6

Mean Discrepancy Scores for Reading

Year	Upper SES	Middle SES	Lower SES
2001	-0.6	0.1	1.1
2002	-0.9	-0.0	1.2
2003	-1.0	-0.3	0.8
Total	217	520	323

The upper SES schools included in this study had a negative average discrepancy score in reading for each year of the state report card included in this study. The negative scores indicate the reading value-added grades were on average lower than the reading achievement grades on the state report card. The average difference in 2001 was -0.6 which is slightly over one half of a letter grade difference. The 2002 average was -0.9 while the 2003 average was -1.0 indicating the value-added grade was on the average approximately one letter grade lower than the reading achievement grade on the 2002 and the 2003 state report cards.

The middle SES schools included in this study had an average discrepancy score of 0.1 in 2001, 0.0 in 2002, and -0.3 in 2003. This information indicated the reading value-added grades and achievement grades were on the average basically the same grade in 2001 and 2002 while the reading value-added grade was on the average approximately one third of a grade lower than the reading achievement grade on the 2003 state report cards for the middle SES schools included in this study.

The lower SES schools included in this study all had positive scores indicating the reading value-added grade was on the average higher than the reading achievement grade on the state report card. In 2001 the mean discrepancy score was 1.1 and the 2002 mean discrepancy score was 1.2 indicating the grades in reading value-added were on the average slightly over one letter grade higher than the reading achievement grade for the lower SES schools included in this study. The 2003 discrepancy score was 0.8 indicating the reading value-added grade was on the average slightly under one letter grade higher than the reading achievement grade for the lower SES schools included in this study.

A cross tabulation table is used to compare the reading achievement grade to the reading value-added grade for the 2001, 2002, and 2003 report card data from the sample of 1,060 schools included in this study. A two-way contingency table analysis

using crosstabs evaluates whether a statistical relationship exists between the two variables (Green, Salkind, & Akey, 2000).

The Kendall's Tau-b will be used to measure the strength of the relationship between school achievement and value-added grades for the total sample and for each level of SES. Kendall's Tau-b is used when both variables are measured at the ordinal level. Kendall's Tau-b has a potential range from -1 to +1. The closer the Kendall's Tau-b is to $\pm$1, the stronger the relationship while a value of zero indicates no relationship between variables (Green et al., 2000). The results for reading achievement grades and reading value-added grades for 2001, 2002, and 2003 are provided in Table 7.

Table 7

Kendall's Tau-B for the Relationship between Reading Achievement and Reading Value-Added Grades

SES	2001	2002	2003
Total Analysis	.093*	-.004	.069*
Upper	.127*	.054	.130*
Middle	-.096*	-.137*	-.100*
Lower	.091	-.004	.053

* Significant at the .05 level

For the total analysis of 2001 reading grades, the Kendall's Tau-b (.093) indicated a weak positive relationship between achievement and value-added grades which was statistically significant at the .05 level ($p<.0005$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.127) indicated a weak positive relationship which was statistically significant ($p = .027$) while the Kendall's Tau-b for the middle SES schools (-.096) indicated a weak negative

relationship which was statistically significant (p = .013). The lower SES schools (.091) exhibited a weak positive relationship which was not statistically significant (p = .053).

For the total analysis of 2002 reading grades, the Kendall's Tau-b (-.004) indicated virtually no relationship between reading achievement grades and reading value-added grades which was not statistically significant at the.05 level (p = .872). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.054) indicated a weak positive relationship which was not statistically significant (p = .343). The Kendall's Tau-b for the middle SES schools (-.137) indicated a weak negative relationship which was statistically significant at the .05 level (p < .0005) while the Kendall's Tau-b for the lower SES schools (-.004) indicated no significant relationship (p = .939).

For the total analysis of 2003 reading grades, the Kendall's Tau-b (.069) indicated a weak positive relationship between achievement and value-added grades which was statistically significant at the .05 level (p = .008). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.130) indicated a weak positive relationship which was statistically significant (p = .018) while the Kendall's Tau-b for the middle SES schools (-.100) indicated a weak negative relationship which was statistically significant (p = .010). The Kendall's Tau-b for the lower SES schools (.053) indicated a weak positive relationship which was not statistically significant (p = .258).

Based upon the information from this data, there is a statistically significant relationship between reading achievement grades on the 2001 and the 2003 state report cards so the null hypothesis for the first research question was rejected for the 2001 and 2003 state report cards but was retained for the 2002 state report card.

After factoring in the SES status of the schools, the upper SES schools had a statistically significant relationship between reading achievement and reading value-added grades on the 2001 and 2003 state report card while there was not a statistically

significant relationship between reading achievement and value-added grades for the upper SES schools on the 2002 state report card. Therefore, the null hypothesis that there is no statistically significant relationship between academic achievement and value added grades in reading in the upper SES schools was rejected for the 2001 and 2003 state report card and was retained for the 2002 state report card.

The null hypothesis stating there is no statistically significant relationship between reading achievement grades and reading value-added grades for the middle SES schools was rejected for all three years (2001, 2002, and 2003) because there was a statistically significant relationship.

The null hypothesis stating there is no statistically significant relationship between reading achievement grades and reading value-added grades for the lower SES schools was retained for all three years (2001, 2002, and 2003) because there was not a statistically significant relationship.

Analysis of Math Achievement and Value-Added Grades

The second research question was: Is there a statistically significant relationship between math achievement and value-added grades on the 2001, 2002, and 2003 state report cards? Also, the SES status of the schools (upper, middle, and lower) will be considered in order to determine if it has an influence on the achievement and value-added scores in math on the state report cards.

A frequency table was computed containing the frequency and percentage of each letter grade for math achievement and math value-added. The frequency tables for the subject area of math are provided in Table 8.

Table 8

Math Achievement and Value-Added Grades

Year/Grade	Math Achievement		Math Value-Added	
2001	N	Percent	N	Percent
A	205	19.3	304	28.7
B	269	25.4	203	19.2
C	307	29.0	214	20.2
D	154	14.5	173	16.3
F	125	11.8	166	15.7
Total	1,060	100.0	1,060	100.0
2002	N	Percent	N	Percent
A	236	22.3	348	32.8
B	294	27.7	227	21.4
C	296	27.9	217	20.5
D	122	11.5	141	13.3
F	112	10.6	127	12.0
Total	1,060	100.0	1,060	100.0
2003				
A	237	22.4	307	29.0
B	332	31.3	223	21.0
C	273	25.8	203	19.2
D	115	10.8	169	15.9
F	103	9.7	158	14.9
Total	1,060	100.0	1,060	100.0

Discrepancy scores ranging from +4 to -4 were computed for each school for the subject area of math on the state report card for each year (2001, 2002, and 2003). Positive discrepancy scores indicate the math value-added grades were higher than the math achievement grades and negative discrepancy scores indicate the math value-added grades were lower than the math achievement grades for that year. The absolute value of the discrepancy score is the number of letter grades between the math value-added grade and the math achievement grade. The discrepancy scores for math are provided in Table 9.

Table 9

Discrepancy Scores for Math (2001, 2002, 2003)

Discrepancy Score	Year 2001		2002		2003	
	N	Percent	N	Percent	N	Percent
-4.00	6	0.6	11	1.0	19	1.8
-3.00	40	3.8	43	4.1	58	5.5
-2.00	126	11.9	97	9.2	136	12.8
-1.00	178	16.8	198	18.7	197	18.6
0.00	324	30.6	293	27.6	298	28.1
+1.00	219	20.7	238	22.5	185	17.5
+2.00	123	11.6	118	11.1	116	10.9
+3.00	36	3.4	49	4.6	35	3.3
+4.00	8	0.8	13	1.2	16	1.5
Total	1,060	100.0	1,060	100.0	1,060	100.0

Note. +Positive scores indicate the value-added grade was higher than the achievement grade. -Negative scores indicate the value-added grade was lower than the achievement grade.

Approximately 30.6% (324) of the schools had the same grade on math achievement and math value-added on the 2001 state report card. One third of the schools 33% (350) had a math value-added grade from one to four letter grades lower than their math achievement grade on the 2001 state report card while 36.4% (386) of the schools had a value-added grade from one to four letter grades higher than their math achievement grade on the 2001 state report card.

Approximately 27.6% (293) of the schools had the same grade in math achievement and math value-added on the 2002 state report card. One third of the schools 32.9% (349) had a math value-added grade from one to four letter grades lower than their math achievement grade on the 2002 state report card while 39.4% (418) of the schools had a math value-added grade from one to four letter grades higher than their math achievement grade on the 2002 state report card.

On the 2003 state report card 28.1% (298) of the schools had the same grade on math achievement and math value-added while 38.7% (410) of the schools had a math value-added grade from one to four letter grades lower than their math achievement grade on the 2003 state report card. Approximately one third of the schools 33.2% (352) had a math value-added grade from one to four letter grades higher than their math achievement grade on the 2003 state report card.

The means of the discrepancy scores were calculated in order to determine the average difference between the math value-added grade and the math achievement grade at each level of SES. The mean discrepancy scores for the subject area of math on the state report card for the sample of schools used for this study are provided in Table 10.

Table 10

Mean Discrepancy Scores for Math

Year	Upper SES	Middle SES	Lower SES
2001	-0.5	-0.1	0.7
2002	-0.6	-0.1	0.9
2003	-0.7	-0.4	0.7
Total	217	520	323

The mean discrepancy score for the upper SES schools included in this study in math indicated the math achievement grade was on the average one half of a letter grade higher than the math value-added grade on the 2001 state report card while the math achievement grade was on the average slightly over one half of a letter grade higher than the math value-added grade on the 2002 state report card for the upper SES schools included in this study. The average discrepancy score indicated the math achievement grade was on the average slightly two thirds of a letter grade higher than the math value-added grade for the upper SES schools included in this study on the 2003 state report card.

The middle SES schools included in this study had negative discrepancy score of -0.1 for 2001 and 2002 indicating the math achievement grade was on the average approximately one tenth higher than the math value-added grade on the 2001 and 2002 state report card for the middle SES schools included in this study. The mean discrepancy score of -0.4 for math on the 2003 state report card indicated the grade in math achievement was on the average almost one half of a letter grade higher than the math value-added grade.

The lower SES schools included in this study had positive mean discrepancy scores indicating the math value added grade was on the average higher than the math achievement grade. The mean discrepancy score of 0.7 on the 2001 and the 2003 state report cards for math indicated the math value-added grade was on the average slightly over two-thirds of a grade higher than the math achievement grade on the 2001 and the 2003 state report cards. The mean discrepancy score of 0.9 indicated the math value-added grade was on the average approximately one letter grade higher than the math achievement grade on the 2002 state report card.

The Kendall's Tau-b was used to measure the strength of the relationship between math achievement and value-added grades for the total sample and for each level of SES. The results for math achievement grades and math value-added grades for 2001, 2002, and 2003 are provided in Table 11.

Table 11

Kendall's Tau-B for the Relationship between Math Achievement and Math Value-Added Grades

SES	Year		
	2001	2002	2003
Total Analysis	.325*	.252*	.205*
Upper	.339*	.359*	.284*
Middle	.220*	.206*	.153*
Lower	.328*	.250*	.209*

*Significant at the .05 level

For the total analysis of 2001 math grades, the Kendall's Tau-b (.325) indicated a positive relationship between math achievement grades and math value-added grades which was statistically significant at the .05 level ($p < .0005$). After factoring in

the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.339) indicated a positive relationship which was statistically significant (p<.0005) while the Kendall's Tau-b for the middle SES schools (.220) indicated a positive relationship which was statistically significant (p<.0005). The lower SES schools (.328) exhibited a positive relationship which was statistically significant (p<.0005).

For the total analysis of 2002 math grades, the Kendall's Tau-b (.252) indicated a positive relationship between math achievement grades and math value-added grades which was statistically significant at the .05 level (p<.0005). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.359) indicated a positive relationship which was statistically significant (p<.0005) while the Kendall's Tau-b for the middle SES schools (.206) indicated a positive relationship which was statistically significant p<.0005). The lower SES schools (.250) exhibited a positive relationship which was statistically significant (p<.0005).

For the total analysis of 2003 math grades, the Kendall's Tau-b (.205) indicated a positive relationship between math achievement grades and math value-added grades which was statistically significant at the .05 level (p<.0005). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.284) indicated a positive relationship which was statistically significant (p<.0005) while the Kendall's Tau-b for the middle SES schools (.153) indicated a positive relationship which was statistically significant (p<.0005). The lower SES schools (.209) exhibited a positive relationship which was statistically significant (p<.0005).

Based upon the information from these data, there is a statistically significant relationship between math achievement grades on the 2001, 2002, and 2003 state report cards so the null hypothesis for the second research question was rejected for the 2001, 2002, and 2003 state report cards.

After factoring in the SES status of the schools, the upper SES schools had a statistically significant relationship between math achievement and math value-added

grades on the 2001, 2002, and 2003 state report cards. Therefore, the null hypothesis there is no statistically significant relationship between academic achievement and value added grades in reading in the upper SES schools was rejected for the 2001, 2002, and 2003 state report cards.

The null hypothesis stating there is no statistically significant relationship between math achievement grades and math value-added grades for the middle SES schools was rejected for all three years (2001, 2002, and 2003) because there was a statistically significant relationship, and the null hypothesis stating there is no statistically significant relationship between math achievement grades and math value-added grades for the lower SES schools was rejected for all three years (2001, 2002, and 2003).

Analysis of Language Achievement and Value-Added Grades

The third research question was: Is there a statistically significant relationship between language achievement and value-added grades on the 2001, 2002, and 2003 state report cards. Also, the SES status of the schools (upper, middle, and lower) will be considered in order to determine if it has an influence on the achievement and value-added grades in language on the state report cards.

A frequency table was computed containing the frequency and percentage of each letter grade for language achievement and language value-added. The frequency tables for the subject area of language are provided in Table 12.

Table 12

Language Achievement and Value-Added Grades (2001, 2002, 2003)

Grade/Year	Language Achievement		Language Value-added	
2001	N	Percent	N	Percent
A	182	17.2	251	23.7
B	268	25.3	125	11.8
C	333	31.4	147	13.9
D	172	16.2	177	16.7
F	105	9.9	360	34.0
Total	1,060	100.0	1,060	100.0
2002				
A	198	18.7	228	21.5
B	288	27.2	95	9.0
C	336	31.7	148	14.0
D	132	12.5	162	15.3
F	106	10.0	427	40.3
Total	1,060	100.0	1,060	100.0
2003				
A	207	19.5	169	15.9
B	299	28.2	105	9.9
C	330	31.1	120	11.3
D	122	11.5	157	14.8
F	102	9.6	509	48.0
Total	1,060	100.0	1,060	100.0

Discrepancy scores ranging from +4 to -4 were computed for each school for the subject area of language on the state report cards for each year (2001, 2002, and 2003). Positive discrepancy scores indicate the language value-added grades were higher than the language achievement grades and negative discrepancy scores indicate the language value-added grades were lower than the language achievement grades for that year. The absolute value of the discrepancy score is the number of letter grades between the language value-added grade and the language achievement grade. The discrepancy scores for language are provided in Table 13.

Table 13

Discrepancy Scores for Language (2001, 2002, 2003)

Discrepancy	Year					
Score	2001		2002		2003	
	N	Percent	N	Percent	N	Percent
-4.00	44	4.2	74	7.0	89	8.4
-3.00	116	10.9	134	12.6	167	15.8
-2.00	191	18.0	190	17.9	209	19.7
-1.00	173	16.3	186	17.5	181	17.1
0.00	206	19.4	207	19.5	205	19.3
+1.00	165	15.6	127	12.0	114	10.8
+2.00	112	10.6	103	9.7	68	6.4
+3.00	42	4.0	30	2.8	21	2.0
+4.00	11	1.0	9	0.8	6	0.6
Total	1,060	100.0	1,060	100.0	1,060	100.0

Note. +Positive scores indicate the value-added grade was higher than the achievement grade. -Negative scores indicate the value-added grade was lower than the achievement grade.

Approximately 19.4% (206) of the schools had the same grade on language achievement and language value-added on the 2001 state report cards. Almost one half of the schools 49.4% (524) had a language value-added grade from one to four letter grades lower than their language achievement grade on the 2001 state report cards while slightly less than one third of the schools 31.1% (330) had a language value-added grade from one to four letter grades higher than their language achievement grade on the 2001 state report cards.

On the 2002 state report cards 19.5% (207) of the schools had the same grade on language achievement and language value-added and over one half of the schools 55.1% (584) had a language value-added grade from one to four letter grades lower than language achievement on the 2002 state report cards. One fourth of the schools 25.3% (269) had a language value-added grade from one to four letter grades higher than language achievement on the 2002 state report cards.

On the 2003 state report cards 19.3% (205) of the schools had the same grade on language achievement and language value-added while 60.9% (646) of the schools had a language value-added grade from one to four letter grades lower than their language achievement grade on the 2003 state report cards. Approximately one fifth of the schools 19.7% (209) had a language value-added grade from one to four letter grades higher than their language achievement grade on the 2003 state report cards.

The mean of the discrepancy scores were calculated in order to determine the average difference between the language value-added grade and the language achievement grade at each level of SES. The mean discrepancy scores for the subject area of language on the state report card for the sample of schools used for this study are provided in Table 14.

Table 14

Mean Discrepancy Scores for Language

Year	Upper SES	Middle SES	Lower SES
2001	-1.4	-0.6	0.2
2002	-1.7	-0.8	-0.1
2003	-2.1	-1.1	-0.2
Total	217	520	323

The mean discrepancy score in language for the upper SES schools included in this study are all negative indicating the language achievement grade is on the average higher than the language value-added grade. The mean discrepancy score for language on the 2001 report card was -1.4 indicating the language achievement grade was on the average approximately one and one-half grade higher than the language value-added grade. The mean discrepancy score of -1.7 on the 2002 report indicated the language achievement grade was on the average one and two-thirds of a letter grade higher than the language value-added grade while the discrepancy score of -2.1 on the 2003 state report cards indicated the language achievement grade was on the average slightly over two letter grades higher than the language value-added grade.

The mean discrepancy score in language for the middle SES schools included in this study are all negative indicating the language achievement grade is higher than the language value-added grade. The mean discrepancy score of -0.6 on the 2001 state report cards and -0.8 on the 2002 state report card indicated the language achievement grade was on the average from three fifths to four fifths of a letter grade higher than the language value-added grade on the 2001 and 2002 state report cards for the middle SES schools in this study. The mean discrepancy score of -1.1 on the 2003 state report

cards indicated the language achievement grade was on the average slightly over one letter grade higher than the language value-added grade for the middle SES schools in this study.

The mean discrepancy score in language for the lower SES schools included in this study are 0.2 for the 2001 report cards, -0.1 for the 2002 report card, and -0.2 for the 2003 state report cards. The language achievement grades were slightly higher than language value-added grades for the middle SES schools on the 2002 and the 2003 state report cards while the language value-added grade was on the average slightly higher than the language achievement grade on the 2001 state report cards for the middle SES schools in this study.

The Kendall's Tau-b was used to measure the strength of the relationship between language achievement and value-added grades for the total sample and for each level of SES. The results for language achievement grades and language value-added grades for 2001, 2002, and 2003 are provided in Table 15.

Table 15

Kendall's Tau-B for the Relationship between Language Achievement and Language Value-Added Grades

SES	Year		
	2001	2002	2003
Total Analysis	.121*	.091*	.113*
Upper	.032	-.029	.030
Middle	.007	-.031	.019
Lower	.096*	.122*	.167*

*Significant at the .05 level

For the total analysis of 2001 language grades, the Kendall's Tau-b (.121) indicated a weak positive relationship between language achievement and language value-added

grades which was statistically significant at the .05 level ($p < .0005$). After factoring in the SES status of the school's, the Kendall's Tau-b for the upper SES schools (.032) indicated a weak relationship which was not statistically significant ($p = .594$) and the middle SES schools (.007) indicated a weak relationship that was not statistically significant ($p = .844$). The lower SES schools (.096) exhibited a weak positive relationship that was statistically significant ($p = .037$).

For the total analysis of 2002 language grades, the Kendall's Tau-b (.091) indicated a weak positive relationship between language achievement and language value-added grades that was statistically significant at the .05 level ($p <.0005$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (-.029) indicated a weak negative relationship that was not statistically significant ($p =.620$) and the middle SES schools (-.031) indicated a weak negative relationship that was not statistically significant ($p=.405$). The lower SES schools (.122) indicated a weak positive relationship that was statistically significant ($p = .008$).

For the total analysis of 2003 language grades, the Kendall's Tau-b (.113) indicated a weak positive relationship that was statistically significant at the.05 level ($p < .0005$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.030) indicated a weak positive relationship that was not statistically significant ($p = .590$) and the middle SES schools (.019) indicated a weak positive relationship that was not statistically significant ($p = .606$). The lower SES schools (.167) exhibited a weak positive relationship that was statistically significant ($p < .0005$).

Based upon the information from these data, there is a statistically significant relationship between language achievement grades on the 2001, 2002, and 2003 state report cards so the null hypothesis for the third research question was rejected.

After factoring in the SES status of the schools, the upper SES schools did not have a statistically significant relationship between language achievement and language value-added grades on the 2001, 2002, and 2003 state report cards. Therefore, the null

hypothesis there is no statistically significant relationship between academic achievement and value added grades in language in the upper SES schools was retained for the 2001, 2002, and 2003 state report cards.

The null hypothesis stating there is no statistically significant relationship between language achievement grades and language value-added grades for the middle SES schools was retained for all three years (2001, 2002, and 2003) because there was not a statistically significant relationship.

The null hypothesis stating there is no statistically significant relationship between language achievement grades and language value-added grades for the lower SES schools was rejected for all three years (2001, 2002, and 2003) because there was a statistically significant relationship.

Analysis of Science Achievement and Value-Added Grades

The fourth research question was: Is there a statistically significant relationship between science achievement and value-added grades on the 2001, 2002, and 2003 state report cards? Also, the SES status of the schools (upper, middle and lower) were considered in order to determine if it has an influence on the achievement and value-added grades in science on the state report cards.

A frequency table was computed containing the frequency and percentage of each letter grade for science achievement and science value-added. The frequency tables for the subject area of science are provided in Table 16.

Table 16

Science Achievement and Value-Added Grades (2001, 2002, 2003)

Grade/Year	Science Achievement		Science Value-Added	
2001	N	Percent	N	Percent
A	97	9.2	405	38.2
B	177	16.7	255	24.1
C	400	37.7	215	20.3
D	208	19.6	101	9.5
F	178	16.8	84	7.9
Total	1,060	100.0	1,060	100.0
2002				
A	97	9.2	417	39.3
B	169	15.9	223	21.0
C	391	36.9	162	15.3
D	215	20.3	107	10.1
F	188	17.7	151	14.2
Total	1,060	100.0	1,060	100.0
2003				
A	92	8.7	281	26.5
B	177	16.7	229	21.6
C	394	37.2	255	24.1
D	205	19.3	152	14.3
F	192	18.1	143	13.5
Total	1,060	100.0	1,060	100.0

Discrepancy scores ranging from +4 to -4 were computed for each school for the subject area of science on the state report cards for each year (2001, 2002, and 2003). Positive discrepancy scores indicate the science value-added grades were higher than the science achievement grades and negative discrepancy scores indicate the science value-added grades were lower than the science achievement grades for that year. The absolute value of the discrepancy score is the number of letter grades between the science value-added grade and the science achievement grade. The discrepancy scores for science are provided in Table 17.

Table 17

Discrepancy Scores for Science (2001, 2002, 2003)

Discrepancy	Year					
Score	2001		2002		2003	
	N	Percent	N	Percent	N	Percent
-4.00	12	1.1	5	0.5	22	2.1
-3.00	31	2.9	16	1.5	42	4.0
-2.00	52	4.9	61	5.8	92	8.7
-1.00	93	8.8	102	9.6	129	12.2
0.00	214	20.2	271	25.6	213	20.1
+1.00	246	23.2	238	22.5	221	20.8
+2.00	225	21.2	215	20.3	189	17.8
+3.00	117	11.0	108	10.2	99	9.3
+4.00	70	6.6	44	4.2	53	5.0
Total	1,060	100.0	1,060	100.0	1,060	100.0

Note. +Positive scores indicate the value-added grade was higher than the achievement grade. -Negative scores indicate the value-added grade was lower than the achievement grade.

One fifth of the schools 20.2% (214) had the same score on science achievement and science value-added on the 2001 state report card while 17.7% (188) of the schools had a science value-added grade from one to four letter grades lower than their science achievement grade on the 2001 state report cards. Approximately 62.1% (658) of the schools had a science value-added grade from one to four letter grades higher than their science achievement grade on the 2001 state report cards.

One fourth of the schools 25.6% (271) had the same grade on science achievement and science value-added on the 2002 state report cards while 17.3% (184) of the schools had a science value-added score from one to four letter grades lower than their science achievement score on the 2002 state report cards. Over one half of the schools 57.1% (605) had a science value-added grade from one to four letter grades higher than their science achievement grade on the 2002 state report cards.

One fifth of the schools 20.1% (213) had the same grade on science achievement and science value-added on the 2003 state report cards while over one fourth of the schools 26.9% (285) had a science value-added grade from one to four letter grades lower than their science achievement grade on the 2003 state report cards. Over one half of the schools 53% (562) had a science value-added grade from one to four letter grades higher than their science achievement grade on the 2003 state report cards.

The mean of the discrepancy scores were calculated in order to determine the average difference between the science value-added grade and the science achievement grade at each level of SES. The mean discrepancy scores for the subject area of science on the state report cards for the schools used for this study are provided in Table 18.

Table 18

Mean Discrepancy Scores for Science

Year	Upper SES	Middle SES	Lower SES
2001	-0.4	0.8	2.0
2002	-0.1	0.8	1.5
2003	-1.0	0.5	1.7
Total	217	520	323

The mean discrepancy scores in science for the upper SES schools in this study are all negative indicating the science achievement grades are on the average higher than the science value-added grades. The mean discrepancy score was -0.4 on the 2001 state report cards indicating the science achievement grade was on the average almost one-half of a letter grade higher than the science value-added grade. The mean score was -0.1 on the 2002 state report cards indicated the science achievement grade was on the average only one-tenth of a grade higher than the science value-added grade. The mean discrepancy score of -1.0 on the 2003 state report cards indicated the science achievement grade was on the average one letter grade higher than the science value-added grade for the upper SES schools included in this study.

The mean discrepancy scores in science for the middle SES schools in this study are all positive scores indicating the science value added grade was on the average higher than the science achievement grade. The discrepancy score in science was 0.8 on the 2001 and the 2002 report cards indicating the science value-added grade was on the average approximately four fifths of a letter grade higher than the science achievement grade while the mean discrepancy score of 0.5 on the 2003 state report cards indicated the science value-added grade was on the average one half of a letter

grade higher than the science achievement grade for the middle SES schools in this study.

The mean discrepancy scores in science for the lower SES schools in this study are all positive indicating the science value-added grade was higher than the science achievement grade. The discrepancy score was 2.0 on the 2001 state report cards indicating the science value-added grade was on the average two letter grades higher than the science achievement grade for the lower SES schools. The discrepancy score was 1.5 on the 2002 state report cards in science indicating the science value-added grade was on the average one and one half of a letter grade higher than the science achievement grade for the lower SES schools in this study. The discrepancy score of 1.7 in science on the 2003 state report cards indicated the science value added grade was on the average one and two thirds of a letter grade higher than the science achievement grade for the lower SES schools in this study.

The Kendall's Tau-b will be used to measure the strength of the relationship between science achievement and value-added grades for the total sample and for each level of SES. The results for science achievement grades and science value-added grades for 2001, 2002, and 2003 are provided in Table 19.

Table 19

Kendall's Tau-B for the Relationship between Science Achievement and Science Value-Added Grades

SES	Year		
	2001	2002	2003
Total Sample	-.008	.222*	-.051
Upper	-.084	.164*	-.112
Middle	-.029	.141*	-.048
Lower	.090	.282*	.076

*Significant at the .05 level

For the total analysis of 2001 science grades, the Kendall's Tau-b (-.008) indicated a weak negative relationship between achievement and value-added grades which was not statistically significant at the .05 level ($p = .760$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (-.084) indicated a weak negative relationship which was not statistically significant ($p = .163$) and the middle SES schools (-.029) exhibited a weak negative relationship which was not statistically significant ($p = .468$). The lower SES schools (.090) exhibited a weak positive relationship which was not statistically significant ($p = .055$).

For the total analysis of 2002 science grades, the Kendall's Tau-b (.222) indicated a weak positive relationship which was statistically significant at the .05 level ($p < .0005$). After factoring in the SES status of the schools, the upper SES schools (.164) indicated a weak positive relationship that was statistically significant ($p = .009$) and the middle SES schools (.141) indicated a weak positive relationship that was statistically significant ($p < .0005$). The lower SES schools (.282) exhibited a positive relationship which was statistically significant ($p < .0005$).

For the total analysis of 2003 science grades, Kendall's Tau-b (-.051) indicated a weak negative relationship between achievement and value-added grades which was not statistically significant at the .05 level ($p = .051$). After factoring in the SES status of the schools, Kendall's Tau-b for the upper SES schools (-.112) indicated a weak negative relationship which was not statistically significant ($p=.067$) and the middle SES schools (-.048) exhibited a weak negative relationship that was not statistically significant ($p = .190$). The lower SES schools (.076) exhibited a weak positive relationship which was not statistically significant ($p = .095$).

Based upon the information from these data, there is a statistically significant relationship between science achievement and value-added grades on the 2002 state report cards so the null hypothesis for the fourth research question was rejected for the 2002 state report cards but was retained for the 2001 and 2003 state report cards.

After factoring in the SES status of the schools, the upper SES schools did not have a statistically significant relationship between science achievement and science value-added grades on the 2001 and 2003 state reports. Therefore, the null hypothesis there is no statistically significant relationship between academic achievement and value added grades in science in the upper SES schools was retained for the 2001 and 2003 state report cards but was rejected for the 2002 state report cards.

The null hypothesis stating there is no statistically significant relationship between science achievement grades and science value-added grades for the middle SES schools was retained for the 2001 and 2003 because there was not a statistically significant relationship but was rejected for the 2002 state report cards.

The null hypothesis stating there is no statistically significant relationship between science achievement grades and science value-added grades for the lower SES schools was retained for the 2001 and 2003 state report cards because there was not a statistically significant relationship but the null hypothesis was rejected for the science achievement and value-added grades for the 2002 state report card because there was a statistically significant relationship.

Analysis of Social Studies Achievement and Value-Added Grades

The fifth research question was: Is there a statistically significant relationship between social studies achievement and value-added grades on the 2001, 2002, and 2003 state report cards? Also, the SES status of the schools (upper, middle, and lower) was considered in order to determine if it has an influence on the achievement and value-added grades in social studies on the state report cards.

A frequency table was computed containing the frequency and percentage of each letter grade for social studies achievement and social studies value-added. The frequency tables for the subject area of social studies are provided in Table 20.

Table 20

Social Studies Achievement and Value-Added Grades (2001, 2002, 2003)

Grade/Year	Social Studies Achievement		Social Studies Value-Added	
2001	N	Percent	N	Percent
A	90	8.5	458	43.2
B	161	15.2	204	19.2
C	418	39.4	167	15.8
D	226	21.3	123	11.6
F	165	15.6	108	10.2
Total	1,060	100.0	1,060	100.0
2002				
A	96	9.1	605	57.1
B	166	15.7	147	13.9
C	420	39.6	107	10.1
D	212	20.0	83	7.8
F	166	15.7	118	11.1
Total	1,060	100.0	1,060	100.0
2003				
A	91	8.6	283	26.7
B	178	16.8	188	17.7
C	415	39.2	205	19.3
D	205	19.3	180	17.0
F	171	16.1	204	19.2
Total	1,060	100.0	1,060	100.0

Discrepancy scores ranging from +4 to -4 were computed for each school for the subject area of social studies on the state report card for each year (2001, 2002, and 2003). Positive discrepancy scores indicate the social studies value-added grades were higher than the social studies achievement grades and negative discrepancy scores indicate the social studies value-added grades were lower than the social studies achievement grades for that year. The absolute value of the discrepancy score is the number of letter grades between the social studies value-added grade and the social studies achievement grade. The discrepancy scores for social studies are provided in Table 21.

Table 21

Discrepancy Scores for Social Studies (2001, 2002, 2003)

Discrepancy	Year					
Score	2001		2002		2003	
	N	Percent	N	Percent	N	Percent
-4.00	3	0.3	1	0.1	14	1.3
-3.00	26	2.5	7	0.7	34	3.2
-2.00	54	5.1	19	1.8	127	12.0
-1.00	98	9.2	67	6.3	148	14.0
0.00	216	20.4	247	23.3	256	24.2
+1.00	239	22.4	245	23.1	198	18.7
+2.00	265	25.0	332	31.3	171	16.1
+3.00	113	10.7	123	11.6	75	7.1
+4.00	46	4.3	19	1.8	37	3.5
Total	1,060	100.0	1,060	100.0	1,060	100.0

Note. +Positive scores indicate the value-added grade was higher than the achievement grade. -Negative scores indicate the value-added grade was lower than the achievement grade.

One fifth of the schools, 20.4% (216) had the same grade on social studies achievement and social studies value-added on the 2001 state report cards while 17.1% (181) of the schools had a social studies value-added grade from one to four letter grades lower than their social studies achievement grade on the 2001 state report cards. Well over one half of the schools 62.4% (662) had a social studies value-added grade from one to four letter grades higher than their social studies achievement grade on the 2001 state report cards.

Approximately one fourth of the schools 23.3% (247) had the same grade on social studies achievement and social studies value-added on the 2002 state report cards while 8.9% (94) of the schools had a social studies value-added grade from one to four letter grades lower than their social studies achievement grade on the 2002 state report cards. Approximately two thirds of the schools 67.8% (719) had a social studies value-added grade from one to four letter grades higher than their social studies achievement grade on the 2002 state report cards.

Almost one fourth of the schools 24.2% (256) had the same grade in social studies achievement and social studies value-added on the 2003 state report cards while 30.5% (323) of the schools had a social studies value-added grade from one to four letter grades lower than their social studies achievement grade on the 2003 state report cards. Approximately 45.3% (481) of the schools had a social studies value-added grade from one to four letter grades higher than their social studies achievement grade on the 2003 state report cards.

The means of the discrepancy scores was calculated in order to determine the average difference between the social studies value-added grade and the social studies achievement grade at each level of SES. The mean discrepancy scores for the subject

area of social studies on the state report card for the sample of schools used for this study are provided in Table 22.

Table 22

Mean Discrepancy Scores for Social Studies

Year	Upper SES	Middle SES	Lower SES
2001	-0.0	0.9	1.7
2002	0.6	1.2	1.4
2003	0.1	0.2	1.2
Total	217	520	323

The mean discrepancy scores for the upper SES schools in this study were 0.0 on the 2001 state report card and 0.1 on the 2003 state report cards indicating the grades in social studies achievement and value-added were on the average basically the same on the 2001 report cards and the social studies value-added grade was on the average only one tenth of a letter grade higher than the social studies achievement grade on the 2003 state report cards. The mean discrepancy score for the upper SES schools on the 2002 state report cards was 0.6 indicating the social studies value-added grade was on the average onehalf of a letter grade higher than the social studies achievement grade for the upper SES schools included in this study.

The mean discrepancy scores for the middle SES schools were all positive indicating the social studies value-added grade was on the average higher than the social studies achievement grade. The mean discrepancy score on the 2001 state report cards in social studies was 0.9 indicating the social studies value-added grade was on the average almost one letter grade higher than the social studies achievement grade. The mean discrepancy score on the 2002 state report cards was 1.2 indicating the social studies value-added grade was on the average slightly over one letter grade

higher than the social studies achievement grade. The mean discrepancy score for science in 2003 was 0.2 indicating the grades in social studies value-added and social studies achievement grades are on the average approximately the same.

The mean discrepancy scores for the lower SES schools included in this study were positive for all three years. The mean discrepancy score in social studies was 1.7 in 2001 indicating the social studies value-added grade was on the average more than one letter grade higher than the social studies achievement grade. The mean discrepancy score on the 2002 state report card was 1.4 indicating the letter grade in science value-added was on the average approximately one and one half of a letter grade higher than the social studies achievement grade. The mean discrepancy score for the 2003 state report card was 1.2 indicating the social studies value-added grade was on the average slightly over one letter grade higher than the social studies achievement grade for the lower SES schools included in this study.

The Kendall's Tau-b was used to measure the strength of the relationship between social studies achievement and value-added grades for the total sample and for each level of SES. The results for social studies achievement grades and social studies value-added grades for 2001, 2002, and 2003 are provided in Table 23.

Table 23

Kendall's Tau-B for the Relationship between Social Studies Achievement and Social Studies Value-Added Grades

SES	Year		
	2001	2002	2003
Total Analysis	.165*	.416*	.098*
Upper	.107	.356*	.148*
Middle	.039	.159*	.002
Lower	.251*	.456*	.096*

*Significant at the .05 level

For the total analysis of 2001 social studies grades, Kendall's Tau-b (.165) indicated a weak positive relationship between achievement and value-added grades which was statistically significant at the .05 level ($p < .0005$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.107) indicated a weak positive relationship which was not statistically significant ($p = .075$) while the middle SES schools (.039) exhibited a weak positive relationship which was not statistically significant ($p = .344$). The lower SES schools (.251) exhibited a positive relationship which was statistically significant ($p< .0005$).

For the analysis of 2002 social studies grades, the Kendall's Tau-b (.416) indicated a positive relationship which was statistically significant at the .05 level ($p < .0005$). After factoring in the SES status of the schools, the Kendall's Tau-b for the upper SES schools (.356) indicated a positive relationship which was statistically significant ($p<.0005$) while the Kendall's Tau-b for the middle SES schools (.159) exhibited a positive relationship which was statistically significant ($p<.0005$). The Kendall's Tau-b for the lower SES schools (.456) indicate a positive relationship which was statistically significant ($p<.0005$).

For the analysis of 2003 social studies grades, the Kendall's Tau-b (.098) indicated a positive relationship which was statistically significant at the .05 level ($p < .0005$). After factoring in the SES status of the schools, the upper SES schools (.148) indicated a weak positive relationship which was statistically significant ($p = .013$) while the middle SES schools (.002) indicated a weak positive relationship which was not statistically significant ($p= .954$). The lower SES schools (.096) exhibited a weak positive relationship which was statistically significant ($p = .045$).

Based upon the information from these data, there is a statistically significant relationship between social studies achievement and value-added grades on the 2001, 2002, and 2003 state report cards so the null hypothesis for the fifth research question was rejected.

After factoring in the SES status of the schools, the upper SES schools did not have a statistically significant relationship between social studies achievement and social studies value-added grades on the 2001 state reports. Therefore, the null hypothesis there is no statistically significant relationship between academic achievement and value added grades in science in the upper SES schools was retained for the 2001 state report cards but was rejected for the 2002 and 2003 state report cards.

The null hypothesis stating there is no statistically significant relationship between social studies achievement grades and social studies value-added grades for the middle SES schools was retained for the 2001 and 2003 because there was not a statistically significant relationship but was rejected for the 2002 state report cards.

The null hypothesis stating there is no statistically significant relationship between social studies achievement grades and social studies value-added grades for the lower SES schools was retained for the 2001 and 2003 state report cards because there was not a statistically significant relationship but the null hypothesis was rejected for the social studies achievement and value-added grades for the 2002 state report cards because there was a statistically significant relationship.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Introduction

The purpose of this study was to determine if there was a relationship between academic achievement and value-added grades on the state report cards published for schools in the state of Tennessee. Three years of state report card data (2001, 2002, 2003) were used to compare the five subject areas of reading, language, math, science, and social studies which receive letter grades on the state report card in academic achievement and value-added. The state uses a pre-determined rating scale to determine the grades in achievement and value-added using data from the state mandated achievement tests administered in grades three through eight. The grades in both academic achievement and value-added are based upon three year cumulative averages and a data base of test score results is maintained by the State Department of Education in order to calculate the three-year averages.

Summary of Results

Reading

The relationship between reading value-added grades and reading achievement grades for schools included in this study was statistically significant on the 2001 and the 2003 state report cards. The relationship between reading value-added grades and achievement grades was not statistically significant on the 2002 state report cards.

Approximately 25.8% of the schools included in this study had an “A” in reading value-added on the 2001 state report cards, but the percentage of schools receiving an “A” declined to 22.7% on the 2002 state report cards and to 19.2% on the 2003 state report cards. Approximately 24.9% of the schools included in this study had a grade of “F” in reading value-added on the 2003 state report cards which was an increase from 17.2% on the 2002 state report cards.

Approximately 44% of the schools included in this study had reading value added grades from one to four letter grades higher than their reading achievement grade on the 2001 state report cards while 32.3% of the schools in this study had reading achievement grades that were from one to four letter grades higher than their reading value-added grade on the 2001 state report cards. Approximately 36.1% of the schools in this study had a reading achievement grade higher than their reading value-added grade on the 2002 state report cards while 41.9% of the schools had a reading value-added grade that was from one to four letter grades higher than their reading achievement grade. A total of 39.8% of the schools in this study had a reading achievement grade higher than their value-added grade on the 2003 state report cards while 34.2% of the schools in this study had a reading value-added grade from one to four letter grades higher than their reading achievement grade on the 2003 state report cards

After factoring in the SES status of the schools on the 2001 state report cards, the upper SES schools had a positive statistically significant relationship and middle SES schools had a negative statistically significant relationship between their reading value-added and reading achievement grades while the lower SES schools did not have a statistically significant relationship between reading value-added and achievement. On the 2002 state report cards, the upper SES schools and the lower SES schools did not have a statistically significant relationship between their reading value-added and achievement grades, but the middle SES schools exhibited a negative statistically significant relationship between the reading value-added and reading achievement grades on the 2002 state report cards. On the 2003 state report cards, the upper SES schools had a positive statistically significant relationship and the middle SES schools had a negative statistically significant relationship between reading value-added and achievement while the lower SES schools did not have a statistically significant

relationship between reading value-added and achievement on the 2003 state report cards.

On the average for the three years of report card data, the upper SES schools had reading achievement grades that were higher than their reading value-added grades while the lower SES schools had on the average higher reading value-added grades than reading achievement grades based upon the mean discrepancy scores for reading.

Math

The relationship between math value-added and math achievement grades for schools included in this study was statistically significant on the 2001, 2002, and the 2003 state report cards.

Approximately 47.9% of the schools included in this study had an "A" in math value added on the 2001 state report cards and the percentage increased to 54.2% on the 2002 state report cards before declining to 50% on the 2003 state report cards.

On the 2001 state report cards, 33% of the schools included in this study had math value-added grades that were from one to four letter grades lower than their math achievement grades while 36.4% of the schools had math value-added grades that were from one to four letter grades higher than their math achievement grade. On the 2002 state report cards, 32.9% of the schools in this study had math value-added grades that were from one to four letter grades lower than their math achievement grades while 39.4% of the schools had math value-added grades from one to four letter grades higher than their math achievement grades. On the 2003 state report cards, 38.7% of the schools in this study had math value-added grades from one to four letter grades lower than their math achievement grades while 33.2% of the schools had math value-added grades from one to four letter grades higher than their math achievement grades on the 2003 state report cards

After factoring in the SES status of the schools, there was a positive statistically significant relationship between value-added grades and math achievement grades for the upper, middle, and lower SES schools included in this study for all three years (2001, 2002, and 2003) of report card data.

On the average for the three years of report card data, the upper and middle SES schools had math achievement grades higher than their math value-added grades while the lower SES schools had higher math value-added grades compared to math achievement grades based upon the mean discrepancy scores for math.

Language

The relationship between language value-added and language achievement grades for schools included in this study was statistically significant on the 2001, 2002, and the 2003 state report cards.

Approximately 34% of the schools included in this study had a grade of "F" in language value-added on the 2001 state report cards. This percentage increased to 40.3% on the 2002 state report cards and to 48% on the 2003 state report cards.

On the 2001 state report card, 49.4% of the schools in this study had language value-added grades from one to four letter grades lower than their language achievement grades while 31.1% of the schools had language value-added grades from one to four letter grades higher than their language achievement grades. On the 2002 state report cards, 55.1% of the schools in this study had language value-added grades from one to four letter grades lower than their language achievement grades while 25.3% of the schools had language value-added grades from one to four letter grades higher than their language achievement grades. On the 2003 state report cards, 60.9% of the schools in this study had language value-added grades from one to four letter grades lower than their language achievement grades while 19.7% had language value-

added grades from one to four letter grades higher than their language achievement grades.

After factoring in the SES status of the schools, there was a statistically significant positive relationship between language value-added grades and language achievement grades at the lower SES schools on the 2001, 2002, and the 2003 state report cards. There was not a statistically significant relationship between language value-added grades and language achievement grades for the middle and upper SES schools on the 2001, 2002, and 2003 state report cards.

On the average, the upper SES schools included in this study had language achievement grades from one and one half to two letter grades higher than the language value-added grade. All three levels of SES schools had on the average language achievement grades higher than language value-added grades with the exception of the lower SES schools on the 2001 state report cards based upon the mean discrepancy scores for language.

Science

The relationship between science value-added and science achievement grades for schools included in this study was statistically significant on the 2002 state report cards. The relationship between science value-added and science achievement grades was not statistically significant on the 2001 and the 2003 state report cards.

Approximately 9.2% of the schools in this study had an “A” in science achievement on the 2001 and 2002 state report cards while 8.7% had an “A” on the 2003 state report card. Approximately 38.8% of the schools in this study had an “A” in science value-added on the 2001 state report cards while 39.3 % had an “A” in science value-added on the 2002 state report cards. Approximately 26.5% of the schools had an “A” in science value-added on the 2003 state report cards.

Approximately 17.7% of the schools included in this study had science value-added grades from one to four letter grades lower than their science achievement grades while 62.1% of the schools had science value-added grades from one to four letter grades higher than their science achievement grade on the 2001 state report cards. Approximately 17.3% of the schools in this study had science value-added grades from one to four letter grades higher than their science achievement grades while 57.1% of the schools had science value-added grades from one to four letter grades higher than their science achievement grade on the 2002 state report cards. Approximately 26.8% of the schools had science value-added grades from one to four letter grades lower than their science achievement grade while 53% of the schools had science value-added grades from one to four letter grades higher than their science achievement grade on the 2003 state report cards.

After factoring in the SES status of the schools, all three levels of SES schools in this study had a positive statistically significant relationship between science value-added grades and science achievement grades on the 2002 state report cards. There was not a statistically significant relationship between science value-added grades and science achievement grades at any level of SES school on the 2001 and the 2003 state report cards.

On the average the upper SES schools had science achievement grades higher than science value-added grades while the middle and lower SES schools had science value-added grades higher than science achievement grade based upon the mean discrepancy scores for science.

Social Studies

The relationship between social studies value-added grades and social studies achievement grades for schools included in this study was statistically significant on the 2001, 2002, and the 2003 state report cards.

Approximately 8.5% of the schools in this study had an “A” in social studies achievement while 43.2% of the schools had an “A” in social studies value-added on the 2001 state report cards. Approximately 9.1% of the schools had an “A” in social studies achievement while 57.1% of the schools had an “A” in social studies value-added on the 2002 state report cards. Approximately 8.6% of the schools had an “A” in social studies achievement while 26.7% had an “A” in social studies value-added on the 2003 state report cards. The percentage of schools with a “C” in social studies achievement on the state report cards remained constant with 39.4% in 2001, 39.6% in 2002, and 39.2% in 2003.

Approximately 17.1% of the schools in this study had social studies value-added grades from one to four letter grades lower than their social studies achievement grade while 62.4% of the schools had social studies value-added grades that were from one to four letter grades higher than their social studies achievement grade on the 2001 state report cards. Approximately 8.9% of the schools in this study had social studies value-added grades from one to four letter grades lower than their social studies achievement grades while 67.8% of the schools had social studies value-added grades from one to four letter grades higher than their social studies achievement grades on the 2002 state report cards. Approximately 30.5% of the schools in this study had social studies value-added grades from one to four letter grades lower than their social studies achievement grades while 45.3% of the schools had social studies value-added grades from one to four letter grades higher than their social studies achievement grades on the 2003 state report cards.

After factoring in the SES status of the schools, the lower SES schools exhibited a positive statistically significant relationship between social studies value-added grades and social studies achievement grades on the 2001, 2002, and 2003 state report cards. The upper SES schools exhibited a positive statistically significant relationship between socials studies value-added grades and achievement grades on the 2002 and the 2003

state report cards but not on the 2001 state report cards. The middle SES schools exhibited a positive statistically significant relationship between social studies value-added and achievement grades on the 2002 state report cards but not on the 2001 or the 2003 state report cards.

On the average the social studies value added grade is higher than the social studies achievement grade at all three levels of SES schools for all three years on the state report card with the exception of the upper SES schools on the 2001 state report cards which averages the same grade in social studies achievement and value-added based upon the mean discrepancy scores for social studies.

Summary

There was a statistically significant relationship between academic achievement grades and value-added grades in math, language, and social studies for schools included in this study on the 2001, 2002, and 2003 state report cards. In reading, the 2002 state report cards did not indicate a statistically significant relationship between reading achievement grades and value added grades while the 2001 and 2002 state report cards did indicate a statistically significant relationship between reading achievement and reading value-added grades. In science, the 2001 and 2003 state report cards did not indicate a statistically significant relationship between science achievement and science value-added grades while the 2002 state report cards did indicate a statistically significant relationship between science achievement grades and science value-added grades.

A breakdown of the SES schools included in this study and the statistical significance between achievement and value-added grades by subject area and report card year is provided in Tables 24, 25, and 26.

Table 24

Breakdown of Statistical Significance for the Upper SES Schools

Statistically Significant	Not Statistically Significant
Subject/Report Card Year	Subject/Report Card Year
Reading 2001 (+), 2003 (+)	Reading 2002
Math 2001 (+), 2002 (+), 2003 (+)	Language 2001, 2002, 2003
Science 2002 (+)	Science 2001, 2003
Social Studies 2002 (+), 2003 (+)	Social Studies 2001

Note. (+) indicates a positive statistically significant relationship between value-added grades and achievement grades. (-) indicates a negative statistically significant relationship between value-added grades and achievement grades.

Table 25

Breakdown of Statistical Significance for the Middle SES Schools

Statistically Significant	Not Statistically Significant
Subject/Report Card Year	Subject /Report Card Year
Reading 2001 (-), 2002 (-), 2003 (-)	Language 2001, 2002, 2003
Math 2001 (+), 2002 (+), 2003 (+)	Science 2001, 2002
Science 2002 (+)	Social Studies 2001, 2003
Social Studies 2002 (+)	

Note. (+) indicates a positive statistically significant relationship between value-added grades and achievement grades. (-) indicates a negative statistically significant relationship between value-added grades and achievement grades.

Table 26

Breakdown of Statistical Significance for the Lower SES Schools

Statistically Significant	Not Statistically Significant
Subject/Report Card Year	Subject/Report Card Year
Math 2001 (+), 2002 (+), 2003 (+)	Reading 2001, 2002, 2003
Language 2001 (+), 2002 (+), 2003 (+)	Science 2001, 2003
Science 2002(+)	
Social Studies 2001 (+), 2002 (+), 2003 (+)	

Note. (+) indicates a positive statistically significant relationship between value-added grades and achievement grades. (-) indicates a negative statistically significant relationship between value-added grades and achievement grades.

Conclusions

The socioeconomic status of the schools had an impact on the value-added and achievement grades on the state report cards issued for schools containing grades three through eight in the state of Tennessee. The upper SES schools had on the average higher achievement grades than value-added grades in all subjects for all three years (2001, 2002, and 2003) of state report card data used for this study with the exception of social studies on the 2001 state report cards. The lower SES schools had on the average higher value-added grades than achievement grades in all subjects for all three years (2001, 2002, and 2003) of state report card data used for this study with the exception of language on the 2002 and the 2003 state report cards.

These findings indicate that teachers, parents, and other stakeholders should expect an upper SES school to have good grades in achievement with lower grades in

value-added while a lower SES school should have lower achievement grades but higher value-added grades.

All of the data used to calculate the grades for the subject areas listed on the state report cards in Tennessee were derived from the state mandated achievement tests. There is much debate regarding the reliability of standardized tests, such as those used in Tennessee, as a way to measure the amount of learning that takes place in a classroom, school, or school district. If there are concerns regarding the reliability of the tests from which the data used to calculate the grades for the state report card, then the reliability of the data on the report cards could be questionable as well.

Recommendations

The achievement grades and value-added grades on the state report cards for schools in Tennessee are based upon two very distinct rating scales. Achievement grades are based upon three year average NCE scores while value-added grades are based upon the amount of scale score growth using three year averages for each subject area at each grade level. Therefore, it is difficult to average the two scales in order to make comparisons between achievement and value-added grades. The rating scales used to calculate the letter grades for the report card was developed by the Tennessee Department of Education. Because the grades are based upon two different scales, the amount of significance between value-added grades and achievement grades could be reduced by adjusting the grading scales used to calculate the value-added and achievement grades. The actual letter grades for achievement and value-added were used for this study instead of the actual numerical averages that were used to calculate the grades.

This study evaluates the value-added grades and achievement grades for each year of the report card independently as opposed to evaluating all three years combined. The level of significance would change if all three years of report card data

were analyzed in combination. However, when the value-added grades are calculated each year, adjustments are sometimes made to previous grades making it difficult to track the actual grades over a three year time period.

The state of Tennessee may want to consider a system similar to the system used in athletics at the secondary school level. Secondary schools are divided into classifications based upon their enrollment in order to allow the schools to be more competitive in athletics. Because the socioeconomic status of the school had an impact on the achievement and value-added grades, the state may want to consider grouping schools according to the percentage of students eligible for free/reduced meals and establish a separate grading system for each group based upon the group average of achievement and value-added grades.

This study analyzes the grades at all five subject areas listed on the state report cards grouping schools into one of three levels of SES schools. It may be more feasible to more closely analyze one subject area over a longer period of time while breaking the SES groups into more than three groups.

REFERENCES

Abrams, L. M., & Madaus, G. F. (2003). The lessons of high-stakes testing. *Educational Leadership*, *61*(3), 31-35.

Amrein, A. L., & Berliner, D.C. (2002). *The impact of high-stakes tests on student academic performance: An analysis of NAEP results in states with high-stakes tests and ACT, SAT, and AP test results in state exams with high school graduation exams.* Arizona State University College of Education, Educational Policy Studies Laboratory, Retrieved May 16, 2004, from http://edpolicylab.org/.

Amrein, A. L., & Berliner, D.C. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership*, *60*(5), 39-42.

Ananda, S., & Rabinowitz, S. (2001). *Building a workable accountability system: Key decision points for policymakers & educators*. [Knowledge Brief]. San Francisco: WestEd.

Archer, J. (1999, May 5). Sanders 101. *Education Week*, pp. 26-28.

Bagin, C. B., & Rudner, L. M. (1994). *What parents should know about standardized testing in schools?*, ERIC Clearinghouse on Assessment and Evaluation: U.S. Department of Education (Contract No. RR92024001).

Baker, A. P., Xu, D., & Detch, E. (1995). *A review of the Tennessee value-added assessment system*. (Authorization Number 307236), Nashville, TN, Comptroller of theTreasury.

Basset, P. (2004, February, 4). From good to great schools. *Education Week*, pp. 44, 30.

Bleim, C. L., & Shepard, L. A. (1995). *An analysis of parent opinions and changes in opinions regarding standardized tests, teacher information, and performance assessment.* (CSE Technical Report 397), Retrieved July 14, 2004, from http://cresst96.cse.ucla.edu/Summary/397shepard.htm.

Boser, U. (2001, January 11). Pressure without support. *Education Week*, p. 68.

Bracey, G. W. (2004). Serious questions about the Tennessee value-added assessment system. *Phi Delta Kappan*, *85*, 716-717.

Bratton, S. (1998). How I respond to the critics. *The School Administrator Web Edition*, 1-2. retrieved November 3, 2003 from http://www.aasa.org/publications/sa/1998_12/bratton_side_responding.htm.

Carey, K. (2004). The real value of teachers. *The Education Trust*, *8*(1), 1-43.

Cohen, S. (1999). Using data effectively. *Niemann Reports*, *53*(1), 35.

CTB/McGraw Hill. (2002). *Guide to test interpretation*. (pp. 1-9). Monterey, CA: McGraw-Hill.

Doran, H. C. (2003). Adding value to accountability. *Educational Leadership*, *61*(3), 55-59.

Edwards, V. B. (Ed.). (1999). Quality counts 99: Rewarding results, punishing failure. (Special Issue) *Education Week*, *18*(17).

Egan, K. (2003). Testing for what? *Educational Leadership*, *61*(3), 27-30.

Evergreen Freedom Foundation. (2002). School Directors Handbook. *Evergreen Freedom Foundation* 1-24. retrieved May 9, 2004 from http://www.effwa.org/pdfs/Value-Added.pdf.

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

Graves, D. (2002). *Testing is not teaching: What should count in education*. Portsmouth, NH: Heinemann.

Greene, J. P., Winters, M. A., & Forster, G. (2003). *Testing high stakes tests: Can we believe the results of accountability tests*. (Civic Report 33), Manhattan Institute

for Policy Research, Retrieved April 14, 2004 from http://www.manhattan-institute.org/html/cr_33.htm

Green, S.B., Salkind, N.J., & Akey, T. M. (2000). *Using SPSS for windows: Analyzing and understanding data*. Upper Saddle River, NJ: Prentice Hall.

Hellend, K. (2001). Value-added assessment. *Evergreen Freedom Foundation*, *11*(5), 1-6.

Hermon, J. L., & Golan, S. (1998). *Effects of standardized testing on teachers and learning*. (CSE Technical Report 334), UCLA Graduate School of Education.

Hershberg, T., Rosenblum, I., & Simon, V. A. (2003, February 19). Adequacy, equity and accountability: Modernizing a system of public education designed for the 20th century. *Education Week*, pp. 48, 33.

Johnson, J. (2003). What does the public say about accountability. *Educational Leadership*, *61*(3), 36-40.

Kohn, A. (2000). *The case against standardized testing; Raising the scores, ruining the schools*. Portsmouth, NH: Heinemann.

Kuppermintz, H. (2002). *Teacher effect as a measure of teacher effectiveness: Construct validity considerations in TVAAS.* (CSE Technical Report 563) University of Colorado, Boulder.

Linn, R. I. (2000). Assessments and accountability. *Educational Researcher*, *29*(2), 4-16.

Long, D., & Cass, M. (2001, November 11). Analyst rocks education boat with theory that teacher, not economic status, is more important. *The Tennessean, pp.* A1, A4.

Matthews, J. (2000, March 14). Tennessee system for gauging results angers some educators but gains acceptance elsewhere. *Washington Post,* p. A7.

McAdams, D. R. (2002). Enemy of the good. *Education Next*, *2*(i2), 23.

Miller, D. W. (2001, March 2). Scholars say high-stakes test deserve a failing grade. *The Chronicle of Higher Education,* p. *A*14.

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

Olson, L. (1999). Quality Counts 99: Report cards for schools. (Special Issue). *Education Week*, *18*(17), p. 27.

Olson, L. (2004, March 3). Tennessee reconsiders value-added assessment system. *Education Week*, p. 9.

Orfield, G., & Wald, J. (2000). The high stakes testing mania hurts poor and minority students the most. *The Nation*, *270*(22), p. 38

Orlofsky, G., & Olson, L. (2001). Quality Counts 01: The state of the states. (Special Issue). *Education Week*, *20*(17), pp. 86-88.

Owens, A. M. (2002). Putting schools to the test: How standardized exams are changing education in Canada. *National Post, 4*(i18), B1, B2.

Pedulla, J. J. (2003). State mandated testing: What do teachers think?. *Educational Leadership*, *61* (3), 42-46.

Perreault, G. (2000). The classroom impact of high stress testing. *Education*, *20*(i4), 705.

Pipho, C. (1998). The value-added side of standards. *Phi Delta Kappan, 79*, 341.

Pipho, C. (1999). High-stakes accountability. *Phi Delta Kappan*, *80*, 485.

Popham, J. (2001). *The truth about testing: An educators call to action.* Alexandria, VA: Association of Supervision and Curriculum Development.

Pruett, K. (2002). *Making sense of the numbers.* Paper presented at TASL (Tennessee Academy of School Leaders) Data Analysis Workshop, Kingsport, TN.

Rayburn, L. (2003). The standardized classroom. *World and I*, *18* (i6), 266.

Rivkin, S.G., Hanushek, E. A., & Kain, J. P. (2002). Teachers, schools and academic achievement, *Dallas Texas Schools Project,* University of Texas, Amherst College, Stanford University, and National Bureau of Economic Research.

Sacks, P. (2000). *Standardized minds: The high price of America's testing culture and what we can do to change it.* Cambridge, MA: Perseus Book.

Sanders, W. L. (1998). Value-added assessment. *The School Administrator Web Edition,* 1-5. retrieved November 3, 2003, from http://www.aasa.org/publications/sa/1998_12/sanders.htm.

Sanders, W. L. & Horn, S. (1994). The Tennessee value-added assessment system (TVAAS): Mixed methodology in educational assessment. *Journal of Personnel Evaluation in Education,* 8, 299-311.

Sanders, W. L., & Horn, S. (1995). Educational assessment reassessed: The usefulness of standardized and alternative measures of students achievement as indicators for the assessment of educational outcomes. *Education Policy Analysis Archives, 3*(6), 1-15.

Shapiro, D. (2001). Will this be on the test. *Sojourners, 30* (i5), 15.

Shepard, L. A., & Bleim, C. L. (1995). *An analysis of parent opinions and changes in opinions regarding standardized tests, teacher information, and performance assessments.* University of Colorado at Boulder, Center for Research on Evaluation Standards and Student Testing (CRESST).

Steinberg, J. (1998, January 7). Shaming public schools by rating them. *New York Times*, p. 13.

Summers, A. A. (2002). Expert measures. *Education Next,* 1-5.

Tennessee Code Annotated, 49-1-606, (1996). Annual estimates of teacher effects on student progress in grades three through eight. 44.

Tennessee Department of Education (2002). How to interpret the report card. retrieved April 9, 2004 from http://evaas.sasinschool.com/tn_reportcard/help/Rpt_Card_Interp.html.

Tennessee Department of Education (2003). No Child Left Behind, A handbook for principals, Office of Federal Programs, 1-72.

Tennessee Department of Education (2004). Understanding TCAP achievement test results. Tennessee Department of Education Evaluation and Assessment Division. *CTB/McGraw Hill 2003,* 1-18.

VITA

Kyle A. Evans

Personal Data: Date of Birth: May 9, 1967
Place of Birth: Elizabethton, Tennessee
Marital Status: Married

Education: Public Schools, Hampton, Tennessee
East Tennessee State University, Johnson City, Tennessee; Elementary Education, B.S., 1989
East Tennessee State University, Johnson City, Tennessee; Supervision and Administration, M.A., 1993
East Tennessee State University, Johnson City, Tennessee; Educational Leadership, E.D., 2005

Professional Experience: Fifth Grade Teacher, Doe Elementary School; Mountain City, Tennessee, 1989-1991
Elementary Principal, Doe Elementary School; Mountain City, Tennessee, 1991-1994
Fifth Grade Teacher, Fairmount Elementary School; Bristol, Tennessee, 1994-1998
Elementary Principal, Haynesfield Elementary School; Bristol, Tennessee, 1998-Present

Honors and Awards: Fairmount Rotary Teacher of the Year, 1997