HESTER, LELA S., Ed.D. Coming to Grips with the Achievement Divide and the Distribution of Effective Teachers. (2006)
Directed by Drs. U. Rick Reitzug and John Willse. 112 pp.

This study explored the significance of deploying effective teachers to schools most heavily impacted by poverty as a strategy for reducing the achievement divide. The degree to which teacher assignments affect students' performance on Algebra I End-of-Course and Eighth Grade Math End-of-Grade tests was examined. Estimates of the effect of a series of effective or ineffective teachers on the students' scores were generated. Achievement scores of all students who participated in Algebra I and eighth grade math testing in Guilford County Schools, Greensboro, North Carolina in 2005 were matched with records in the value added databases maintained by SAS Institute. A variety of descriptive analyses were conducted to demonstrate the relationship between the cumulative effects of teacher quality and student achievement as measured by students' performance on Eighth Grade Math End-of-Grade and Algebra I tests. Even after adjusting for the entering achievement of the students in fourth grade, the impact of the previous fifth, sixth and seventh grade teachers, was quite significant on how eighth grade students performed on the Algebra I End-of-Course and the End-of-Grade tests. Further, the study investigated the relationship between teacher effectiveness scores and teacher years of experience. The study confirmed that teachers with more years of experience tended to be more effective than non-experienced teachers. The poorer schools were also more likely to have a higher percentage of less experienced teachers.

In addition, the distribution of teachers based on their teacher effectiveness estimates was examined across the Guilford County public school system. Generally, the highest percentage of effective teachers were assigned to schools that were least impacted by poverty. The results of the study should serve as a necessary catalyst for policy makers and personnel of Guilford County Schools and other districts across the nation to make decisions regarding the equitable deployment of effective teachers as a viable means of reducing the achievement gap.

# COMING TO GRIPS WITH THE ACHIEVEMENT DIVIDE AND THE DISTRIBUTION OF EFFECTIVE TEACHERS

by

Lela S. Hester

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Education

Greensboro
2006

Approved by

Dr. U. Rick Reitzug
Committee Co-Chair

Dr. John Willse
Committee Co-Chair

To my extraordinary loved ones who have supported me with well wishes, support, encouragement, and patience: my husband, Jeff, who is the love of my life and greatest champion; my children, Letasha, Ashlee, Lauren, and Jeffrey and grandchildren, Jonathan and Baileigh, who bring me tremendous joy and motivate me to work relentlessly for a better tomorrow; my sisters, Jessie and Gloria, who always have a listening ear and creative ways to lift my spirits; my brothers, Tommy, Clint, Don, and Ronnie, who stand behind me no matter what; my nieces, Jackie and Tewauna, who care for me like sisters; my sister-in-law, Rita, who is a tremendous example of faithfulness and courage; my father-in-law, Ervin Hester, Sr., who is a living legend and a true inspiration; my fellow cohort members, colleagues and dear friend Le, who laughed, lamented and labored with me, especially Doris and Lynne; my other dear family members and friends, who provided love and prayers; my parents, Sam and Lela, and brothers, Sam and George, who have all gone to meet my Heavenly Father and watch over me both day and night, and most important, my Lord and Savior *who sticketh closer than any brother*.

# APPROVAL PAGE

This dissertation has been approved by the following committee of the Faculty of The Graduate School at the University of North Carolina at Greensboro.

Committee Co-Chairs Dr. U. Rick Reitzug

Dr. John Willse

Committee Members Dr. C. P. Gause

Dr. Carl Lashley

July 13, 2006
Date of Acceptance by Committee

July 13, 2006
Date of Final Oral Examination

# ACKNOWLEDGMENTS

I extend heartfelt thanks to my co-chair, Dr. John Willse, for his invaluable feedback, attention to details and belief in my abilities to conduct a quantitative study. I offer sincere gratitude to Drs. Reitzug, Lashley, Gause, and Cooper for providing ongoing feedback and encouragement. A special thank you is expressed to Dr. June Rivers and Mr. Steve Enck of SAS Institute. Finally, thank you is extended to Guilford County Schools and Ms. Shirley Morrison and Ms. Sherri Thomas for their invaluable assistance.

This is the value of the teacher, who looks at a face and says there's something behind that and I want to reach that person, I want to influence that person, I want to encourage that person, I want to enrich, I want to call out that person who is behind that face, behind that color, behind that language, behind that tradition, behind that culture. I believe you can do it. I know what was done for me.

Maya Angelou

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

Table

Page

Page

## LIST OF FIGURES

# CHAPTER I

# INTRODUCTION

For years, educators and non-educators have suggested possible reasons, and in too many instances, excuses for low student achievement and the achievement gap. Among the leading causes cited for low achievement and the achievement gap are: genetics, family background, socio-economic status, low wealth schools/districts, non-charismatic or influential school leaders, low teacher-student expectations, and racism, just to name a few. As an African-American educator, I am incensed by the notion that genetics alone is the cause of low achievement and the achievement gap. I am equally disturbed about the prevailing definition of intelligence as a stagnant or fixed quantity, which is known to many as an Intelligence Quotient or IQ score (Herrnstein & Murray, 1994). If one believes this theory, then there is very little education can do to increase intelligence. I am aware that heredity may predispose us to certain conditions, but I am convinced that intelligence is fluid and impacted significantly by effective effort on the part of the student and the teacher (Howard, 1995).

If one believes that intelligence is indeed a fixed quantity then one also believes that there is little that can be done to rectify the achievement divide. On the contrary, we have known for years what matters most in successfully educating America's children: the teacher (Haycock, 1998). Our parents of yesterday and parents of today know that there are tremendous differences between teachers that teach at the same grade level, the same subject, and in the same building. As a building-level administrator, I recall communicating to parents that, in all fairness, I could not accept teacher requests, but I encouraged them to describe the unique characteristics of their child, their preferred learning style, and the optimum environment they felt was needed for continued growth. Many of the parents held true to my criteria but described the effective teachers and their classrooms with immense detail, yet never put forth a name. They had an opinion, as I, that some teachers were better than others, but neither of us truly knew how effective a teacher was at helping her students learn and make progress from the start of the year to the end. In other words, we had a data problem. There was nothing available that could measure what a student knew when s/he arrived the first day of school and when s/he departed the last day of school.

The time to address high achievement for all students and eliminate the achievement divide is upon us. The national education goals call for closing the achievement gap among socio-economically deprived and minority students. On the topic of the achievement gap, Secretary of Education Margaret Spellings stated that, “For the first time ever, we are looking ourselves in the mirror and holding ourselves accountable for educating every child. That means all children, no matter their race or income level or zip code” (U. S. Department of Education, 2005, para. 1). This nation’s new educational goals are tied directly to closing the achievement gap for low-income and minority students. The *No Child Left Behind Act* (NCLB) of 2001 contains provisions designed to increase student achievement and close the achievement gap.

NCLB calls for equity among student populations and seeks to provide quality educational programs to all disadvantaged children. NCLB is designed to promote high standards. It acknowledges that low achievement often results from the use of inferior programs where instruction is delivered by inadequately trained or uncertified teachers. In addition, the Act requires that schools funded with Title I dollars align their efforts with state standards, hire highly qualified staff, develop research-based initiatives, and achieve measurable results within a specified timeframe. If schools fail to meet academic standards or Adequate Yearly Progress (AYP) for two consecutive years, the schools are deemed as in need of improvement or in “School Improvement.” As a result, parents must be given the choice of sending their child to another public or charter school in the district that is not in School Improvement.

If parents act on their choice of attending another school, transportation must be provided by the district. Schools that fail to meet standards for three consecutive years must offer school choice and tutoring at not cost to eligible students who receive free and reduced priced meals which is a school measure of socio-economic status. Title I schools deemed as failing are required to develop corrective plans of action. If these plans do not bring about desired results, more radical measures may be applied including a complete overhaul of schools’ staff, curricular changes or even a takeover by the state (U. S. Department of Education, 2005).

The *No Child Left Behind Act* (NCLB) is a strong indictment against educationally shortchanging any of America's children and represents zero tolerance for an achievement gap among groups of youngsters. However, the Adequate Yearly Progress or all or nothing standard will not adequately predict district, school or teacher effectiveness.

Currently, AYP is based on how students perform in a given year on standardized tests that are administered at the end of each school year. States have developed incremental achievement goals in anticipation of meeting the goal of 100% of students on grade level by 2013-2014. For example, North Carolina's mathematics goal for the 2005-2006 school term is for 81% of its students in a school or district to demonstrate proficiency or perform at or above grade level on a criterion-referenced test administered in grades three through eight. While this is an honorable goal, the use of simple raw averages to determine proficiency is inappropriate due to the lack of consideration of factors that fall outside of the control of the school or district such as socio-economics. As a result, a school that serves a large number of children who are deprived economically and more than a year below grade level at the start of school, and helps its students make more than a year's worth of gains by the end of the year is designated as an ineffective or failing school if the school's overall achievement falls short of the state's standard. Conversely, an affluent school may meet the state proficiency standard and deemed as effective, yet the students in the school may not experience a year's worth of growth. Although the NCLB goals are well-intended, the goals may not be realized if we do not find efficient and reliable ways to provide children who have high academic needs with not just highly qualified teachers but teachers that can add value to students' learning year after year.

## Problem Statement: Assessing Teacher Effectiveness

In 1966, the most convincing research regarding the state of education was published in the Coleman Report (Coleman et al., 1966). The researchers conducted a social science project involving 600,000 children in 4,000 schools. Most of the data, aggregated across schools and neighborhoods, were compiled by the U. S. Bureau of Census. The researchers found a strong relationship between academic success and the qualities of students such as socio-economic status and race. The report found a much weaker link between academic achievement and the qualities of the schools. School effects were not isolated by the researchers, not because they were not there, but because the researchers did not have the methods or the data to link school effects to academic success (Carey, 2004). As a result, many people chose to interpret the results of the report to suggest that teachers did not matter a great deal.

After the Coleman Report, there were quite a few studies of student academic success in relation to school variables, and socio-economic and motivational factors. Hanushek (1986) provided a critical review of the aforementioned factors in the production function literature in education. One study conducted by Robert Strauss and Elizabeth Sawyer (1986) provided a statistical analysis of the average student performance on standardized tests and the extent to which students fail these tests. The study also, unlike other studies during that time, used the quality of teachers as measured by standardized test scores as a determinant of performance. The most critical finding was that a 1% increase in teacher quality as measured by standardized test scores was accompanied by a 5% decline in the level of failure or rate of failure of students on high school competency tests. Although the findings were significant, using test scores alone as a measure of student performance does not take into account the students' readiness levels upon entering the class. Therefore, the findings did not demonstrate a convincing link between student achievement and teacher quality, but the study was a much needed springboard for further research in this area.

Researchers and educators have debated which school variables impact student achievement for many years. Some researchers have suggested that "schools bring little influence to bear upon a child's achievement that is dependent of his background and general social context" (Coleman et al., 1966). Factors like teacher qualifications (Ferguson, 1991), class size (Glass, Cayhen, Smith, & Filby, 1982) and others have all been attributed to student achievement.

As the nation searches for the formula for student success and moves toward a stronger focus on accountability and closing the divide in achievement, attention is drawn to the person who spends the most amount of time with students during the course of a day and is thought to be the most influential school-related factor in student achievement, the teacher (Darling-Hammond, 1997).

A growing body of research suggests that the teacher is the most important variable in a student's schooling. Teacher effect studies using the Tennessee Value-Added Assessment System (TVAAS), and a similar system in Dallas, Texas, found that teacher effectiveness is a more powerful determinant of student learning than the effects of class size or grouping practices (Jordan, Mendro, & Weerasinghe, 1997; Sanders & Rivers, 1996; Wright, Horn, & Sanders, 1997). TVAAS uses a statistical method to determine the effectiveness of school systems, schools and teachers (Sanders & Horn, 1998). The method includes a multivariate, longitudinal analysis of student achievement data. Each student's test data, in scaled scores, are collected over time and matched to the student's school system, school, and teachers. The achievement data are used to track learning patterns over time and define teacher effectiveness as the measure of influence of a teacher on indicators of learning (Sanders & Horn, 1998). Effective teachers are defined as those that lead students to achieve normal academic gain over a three-year period. The TVAAS reports include information on student gains for each subject and grade for the three most recent years as well as the three-year average gains. The cumulative average gain is the primary indicator of success (Sanders & Horn, 1998). In other words, effective teachers add at least a year's worth of growth annually on student learning. Therefore, it is extremely important that the most fragile students who lag behind in their achievement are assigned teachers who are deemed as "effective" by objective standards. As Sanders (1998) points out, students whose initial achievement levels are comparable experience different academic outcomes based on their assigned sequence of teachers. Teacher effects appear to be critical and not compensatory. In addition, these studies uncovered evidence of strong bias in assignment of students to teachers based on the effectiveness levels of the teachers (Jordan et al., 1997). The studies further indicate that African American students were nearly twice as likely to be assigned to the most ineffective teachers and half as likely to be assigned to the most effective teachers (Sanders & Rivers, 1996). In order to equitably distribute effective teachers, educators must adopt a system that minimizes subjectivity and can provide information about how teachers influence their students' learning during an academic school year.

Increased accountability is an educational theme in the 21st century at both the state and national levels. Although the No Child Left Behind Act signed into law by President Bush in 2002 places a great deal of emphasis on "hiring highly qualified teachers" (U. S. Department of Education, 2005), there is little agreement regarding how to measure teacher effectiveness. Teachers are any child's best hope at achieving at high levels. This is especially true for poor children and children of color. Research indicates that poor children and children of color arrive at kindergarten behind their counterparts (Barton, 2005). The divide is largely due to socio-economics and lack of prior knowledge, and as the students progress, the deficits accumulate.

Although the achievement gap between African American and White students declined from 1970 to 1980 by 50%, the gap began to increase in 1988 (Haycock, 2001). A review of the findings from the National Center for Educational Statistics (2001) showed that by the end of high school, African Americans had acquired skills in reading and mathematics that were the same as those of their eighth grade White counterparts. Perhaps even more devastating, the findings revealed that African American students were half as likely to complete four years of college as Whites.

The achievement gap between the two groups as measured by standardized measures begins before children enter kindergarten and continues into adulthood (Jencks & Phillips, 1998a). The gap narrowed between 1972 and 1988 by about 50% (Haycock, 1997). African Americans showed the most gains among children who began school in 1968 through 1972 and from 1976 through 1980. The reading gap between African Americans and Whites was 30 points; on a 100 point scale; however, the gap narrowed to 18 points in 1988 and increased to 30 again in 1992 (Haycock, 1997). SAT averages increased for all students between 1991 and 2001; however, a significant gap existed between African American and White students (Boehner, 2001). Analyses of data from the National Assessment of Educational Progress (2000) indicated that the achievement gap between African American and White students begins in elementary school and continues throughout high school. The achievement gap is evident in grades, test scores, course selection and graduation rates (Comer, 2001). According to Comer (2001), by the time an African American student completes the fourth grade, he is two years behind his White counterparts in mathematics and reading achievement. Likewise, when the student enters eighth grade, he is three years behind, and by grade 12, four years behind.

Although achievement data in both mathematics and reading for North Carolina students in grades 3 through 8 as shown in Table 1 indicate that the gap has declined over the years, a significant gap remains (North Carolina Department of Education, 2005). In 1992-93, the state posted a 33.3 percentage point gap in students at grade level in reading and mathematics between white and black students; 30 percentage points in 2000-2001; 27.8 percentage points in 2001-2002; 21.9 percentage points in 2002-2003; 21.5 percentage points in 2003-2004, and in 2004-2005, a 21.8 percentage point divide in achievement between the two races.

These results are staggering, yet we impose minimum standards rife in bureaucracy such as teacher certification programs on teachers as verification of their effectiveness (Darling-Hammond, 2001). Too often, ineffective teachers are assigned to our most needy schools or schools heavily impacted by poverty (Carey, 2004). Across the nation and this state, there are plenty of examples of how high poverty and schools of color achieve at high levels (Haycock, 1998). When achievement results are examined closely, one key factor is revealed:

**Table 1**

***1992-1993 to 2004-2005 North Carolina End of Grade Test Results: Percent of Students at or Above Grade Level in Both Reading and Mathematics in Grades 3-8 for Black and White Students***

| Year | Black (Percent of Students) | White (Percent of Students) |
|---|---|---|
| 1992-1993 | 30.1 | 63.4 |
| 2000-2001 | 52.0 | 82.0 |
| 2001-2002 | 56.6 | 84.4 |
| 2002-2003 | 66.9 | 88.8 |
| 2003-2004 | 67.7 | 89.2 |
| 2004-2005 | 67.2 | 89.0 |

when teachers teach what students need to know to perform at high levels, the students do. Why are schools, especially schools that have children with increased academic needs, not filled with effective teachers? Schools are not filled with effective teachers largely because school systems have not accessed reliable measures to verify effectiveness and have not acted on the information in the distribution of teachers (Carey, 2004).

**Significance**

In light of what we know about teacher quality, teacher effect or value-added data, and the achievement gap, districts across the nation should embrace the opportunity to find out more about teacher effectiveness and how highly effective teachers are distributed across districts. The more we know about teacher effectiveness, the better chance we will have leveraging that information to improve the educational system. We need to uncover where the effective teachers teach and deploy them differently if they are not teaching minorities and children of poverty. In addition, teacher effectiveness and its relationship to teaching experience should be explored since experience has been identified as an indicator of teacher effectiveness. Therefore, this study, patterned after the Rivers (1999) study, was designed to look at the relationship between the sequences of teacher effectiveness, years of teaching, and the math achievement of students. The distribution of teachers deemed as “highly effective” as measured by the value-added approach will also be examined.

In Guilford County Schools, Greensboro, North Carolina, SAS EVAAS or value-added or Teacher Effect data are a part of the reflective practice process for all teachers in grades 3 through 12; however, the data are not used to validate claims that teachers matter most, and ensure placement of effective teachers to the schools that need them the most. Given that the district has a number of schools impacted by poverty, schools that did not make Adequate Yearly Progress as defined by NCLB for more than two consecutive years, and an achievement gap between students of color and white students, the district needs definitive data that validate the importance of hiring and maintaining effective teachers and deploying them to schools that need them the most.

The implications of the study are far reaching and may cause the Board of Education and district leaders in Guilford County Public Schools, Greensboro, North Carolina to make some well-founded decisions regarding how to increase achievement in all schools and eliminate the achievement gap. In addition, this study will add to the limited body of knowledge regarding the promising use of value-added data as districts across the nation operationalize the *No Child Left Behind* legislation.

## Background and Overview of Methodology

The measurements for this study included two independent estimates of student achievement or mastery in mathematics (Algebra I, eighth grade math End-of-Grade and grade four math scale scores) and the SAS Educational Value Added Assessment System (EVAAS) estimates of teacher effectiveness. At the end of each school year, upon the conclusion of the Algebra I and eighth grade math courses, students take the Algebra I End-of-Course (EOC) and End-of-Grade (EOG) tests respectively. Taking the Algebra I End-of-Course test and scoring at the proficient level is a North Carolina graduation requirement for all students and a NCLB/AYP indicator. In North Carolina, End-of-Grade tests are administered in reading and math in grades three through eight; however, only the mathematics scale scores are included in this study. In addition, students in grades three, five and eight must obtain proficient scores on the eighth grade math and reading EOG tests to get promoted to the next grade level.

Guilford County Schools has participated in value-added assessments for the past five years and worked directly with Drs. William Sanders and June Rivers, pioneers of the Tennessee Value Added Assessment System (TVAAS). The TVAAS was developed to provide impartial estimates of the influences that school systems, schools, and teachers have on the academic gains of students in a number of subjects (Sanders, Saxton, & Horn, 1997). The database of student achievement data was created and a statistical methodology was applied to the database using a software package designed to handle years of longitudinal data. Student scale scores from norm-referenced or criterion-referenced tests are used in the multivariate, longitudinal analysis (Sanders, 2000). The mixed-model process is designed to provide the best linear unbiased estimates and predictors of the influence of school systems, schools, and teachers upon student learning.

Because of the model's key multivariate, longitudinal features, estimates of influence of the school district, schools, and teachers on student learning have been shown to be not correlated with socio-economic factors and prior achievement levels of students (Sanders & Horn, 1998). This phenomenon allays concerns that a number of other variables must be included in the analysis to guarantee fair and impartial assessments. According to Sanders and Topping (1999):

> The TVAAS model does not use gains as the dependent variable, but fits the entire observational vector. The district, school, or teacher effects are simultaneously estimated considering the variance-covariance structure of the data. Using an appropriate set of estimable functions, these effects are back mapped into mean gains for reporting purposes. Thus gains from different baselines are not directly compared, and the starting point of the students is irrelevant so long as the testing regime provides sufficient scale elongation to allow measurement of progress for the lowest and highest achieving students. (p. 3)

The statistical methodology that is the foundation of TVAAS, which is now synonymous to the SAS EVAAS model, is the Henderson mixed model (Sanders et al., 1997). The equations included in the Henderson model and the modifications to them, enable the use of all test data for each student regardless of how sparse or incomplete the data may be. Incomplete data may result from students changing schools, moving to other districts or states, or missing tests. The approach used by SAS EVAAS is a major advantage over traditional statistical approaches that use fewer years of data in order to have complete information or use more years of data with fewer students. Incorporating either traditional approach can result in biased estimates. The SAS EVAAS approach reduces these problems.

Two advantages of the SAS EVAAS statistical model are processes known as the “shrinkage estimates” and the “layered” model (Sanders, 2000). Although, there is some protection against false estimates because of the number of student records at the school and district levels, that same protection is not available at the individual teacher or classroom level because of the small number of student records. As a result, the risk of individual teachers receiving false negative reports because of the small numbers is extremely high. To combat this occurrence and increase the validity of the results, SAS EVAAS incorporates the statistical process known as “shrinkage estimates.” With this approach all teachers are assumed to be the average of their school system until the weight of the data pulls their specific estimates away from their school systems’ mean. Therefore, if teachers have small quantities of student records, it is virtually impossible to distinguish among individuals; the teachers’ estimates will not be considerably different from their systems’ mean (Sanders, 2000; Sanders et al., 1997).

The student data necessary for these analyses were acquired from Guilford County Schools and prepared for analyses by the SAS Institute, Cary, North Carolina. This preparation required the merging of year 2001 and 2005 data from the district with achievement and teacher effectiveness for the same students. A general linear model analysis, PROC GLM was used by SAS Institute to ascertain any interactions or effects. The procedure used the method of least-squares and analyzed data within the framework of General linear models. In contrast to means, the LS MEANS statement performs multiple comparisons on interactions and main effects. The GLM procedure can be used for such as analyses as: simple and multiple regression, analysis of variance and covariance, response-surface models, weighted and polynomial regression, partial correlation, multivariate analysis of variance and repeated measures analysis of variance. Through estimability, the GLM procedure can provide tests of hypotheses for the effects of a linear model regardless of the number of missing cells or the extent of the confounding. (SAS Institute, 2006).

Value-added or teacher effect data on teachers who teach mathematics and reading in grades three through eight, Algebra I & II, Geometry, and other high school End-of-Course tests were merged in the value-added database maintained by SAS Institute. The database provides the basis for the district's estimation of district, school, and teacher effects in the same subjects. District and school results are used to make data-guided decisions regarding improvement efforts. Teacher effect data are used to assist teachers in understanding how their students learn and grow as a result of their instruction, regardless of their starting point; help teachers understand the effectiveness of their instruction for various levels of student achievement; assist principals in placing students and assigning subjects to teachers for the most effective growth, and provide support for other measures of teacher effectiveness. The first teacher effect reports were released to teachers in 2003-2004; however, the data are not a part of the teachers' formal evaluation process.

## Statistical Assumptions for Applications of the General Linear Model

The least-squares approach provides estimates of the linear parameters that are unbiased and have minimum variance among linear estimators. Further, under the assumption that the errors have a normal distribution, the least squares estimates are the maximum likelihood estimates and their distribution is known. This assumption is necessary because normality is required in order for the significance levels (“p values”) and confidence intervals to be valid (SAS Institute, 2006).

## Organization of the Study

This dissertation is organized into five chapters. The first chapter includes an introduction, problem statement, significance of the study, background information and overview of the methodology, and definition of terms. Chapter II consists of a review of literature on effectiveness studies and teacher quality. The research design or methodology and data analysis plan are described in Chapter III. The analysis and discussion of the results and recommendations are presented in Chapter IV and Chapter V respectively.

## Definition of Terms

1. *Value-added*. Value Added is a statistical way to analyze test data to determine the influence of teachers, schools and districts on student learning.
2. *The Tennessee Value-Added Assessment System (TVAAS).* The process by which the effects or influence of school, school systems, and teachers on the academic growth of students in grades three through eight in science, math, social studies, language arts, and reading are estimated in Tennessee. TVAAS uses a mixed-model methodology to produce a multivariate response analysis which allows the inclusion of all available student achievement data regardless of the degree of missing information (Sanders et al., 1997).

3. *SAS Educational Value-Added Assessment System (EVAAS)*. The process that is synonymous to TVAAS and executed by SAS Institute, Inc., Cary, North Carolina.
4. *Effectiveness*. When students experience at least a year's worth of growth from the beginning of the school year to the end of the same year.
5. *Teacher Effect Data*. Teacher effect scores refer to effectiveness of teachers.
6. *Cumulative Teacher Effect*. a) The accumulations of measurable effects of teachers on students' learning years after being taught by that teacher, and b) The accumulations of measurable effects of the sequence of teachers on students' learning (Sanders and Rivers, 1996).
7. *Value-Added Data*. Value-added scores refer to effectiveness of schools and districts.
8. *Tertiles or Student Ranks*. Distribution of effectiveness scores into thirds with the lowest degree of effectiveness in the first tertile (lowest 25%), middle 50%, and the greatest degree of effectiveness in the third tertile (top 25%).
9. *End-of-Grade Tests*. Criterion referenced assessments in reading and mathematics administered to North Carolina students in grades three

through eight at the conclusion of each school term. These results are reported in Scale Scores.

10. *End-of-Course Tests*. Criterion referenced assessments administered to North Carolina students who take Algebra I, Biology, English 9, Physics, U. S. History, and Geometry. These results are reported in Scale Scores.
11. *General Linear Model Procedure or PROC GLM*. PROC GLM analyzes data within the framework of general linear models using the method of least squares. It handles models relating one or several continuous dependent variables to one or several independent variables. The procedure can be used for regressions, analyses of variance and covariance, multivariate analyses of variance and partial correlation. (SAS Institute, 2006).
12. *Estimable Functions*. The use of estimable functions will allow the analysis of various combinations of teacher sequences.
13. *LS MEANS Statement*. Least-squares means (LS-means) are calculated for each effect listed in the LSMEANS statement. In contrast to MEANS the statement, the LSMEANS statement performs multiple comparisons on interactions and main effects. (SAS Institute, 2006).

14. *Type I Tests*. Type I sums of squares (SS) also called sequential sum of squares are the incremental improvement in error sums of squares as each effect is added to the model (SAS Institute, 2006).
15. *Type III Tests*. Type III sums of squares (SS) are sometimes referred to as partial sums of squares. (SAS Institute, 2006).
16. *Degrees of Freedom (DF)*. The number of independent pieces of information remaining after estimating one or more parameters (Howell, 2004).
17. *Free/Reduced Price Lunch (FRPL)*. Awarded to children through a federally funded program at school who qualify due to their parent's financial status. Receipt of FRPL is often used as a poverty index indicator. In this study, low poverty will most often be referred to schools that fall within the 0-59% FRPL category and high poverty, between 60-99%.
18. *Odds Ratio*. The ratio of two odds.
19. *Cochran-Mantel-Haenszel Statistic*. Computation of chi square to test the independence of two variables (Howell, 2004).

# CHAPTER II

# REVIEW OF THE LITERATURE

## Why Adopt a System that Measures a Teacher's Effectiveness in Terms of How Much His Students Learn from the Beginning of the Year to the End?

Although the *No Child Left Behind* (NCLB) legislation emphasizes accountability, this accountability rests at the steps of the school rather than individual teachers. According to Hershberg (2005), the legislation falls short for two reasons. First, because there is greater variability in quality of instruction within schools than between schools, data at the classroom level must be reported and evaluated. Second, comprehensive school reform can only take place when every one's careers are linked directly to learning outcomes. In addition, the NCLB legislation is underscored by the use of simple raw averages to draw conclusions regarding school effectiveness. These averages are so riddled with factors, like socio-economics, which lie outside of a school's control, that it is virtually impossible to reach any sensible conclusions regarding a school's effectiveness (Sanders, 2000; Wang, Haertel, & Walberg, 1993). Under the legislation, accountability is high and the stakes are high, yet they are based on one achievement score. One can draw many unfair conclusions by using raw averages. For example, a school that serves primarily children in poverty could have experienced tremendous growth or academic progress, yet when compared with their district's average may be deemed as a "failing" school. Likewise, a school that serves children from affluent families may be deemed a "school of excellence" because of its overall achievement average but the students in the school may not have experienced high rates of growth. Some attempts were made to address misinterpretation of data. One attempt seeks to disaggregate raw averages by socioeconomic groups. However, this approach has significant limitations in terms of drawing meaningful comparisons due to the obscurity of other confounds within the stratification schema (Sanders, 2000). For example, if averages are reported by ethnic groups, then the influence of a student's parents' educational history is not taken into consideration. The use of a different measure of teacher effectiveness is definitely warranted.

Many states have taken a long time to establish the conditions necessary to validate the opinions of parents and administrators regarding the effectiveness of teachers. Few states, until fairly recently, had a common curriculum or standards, yearly standardized assessments, and computers to house the achievement data over time (Haycock, 1998). By the 1990s, some states had adopted the necessary components to fairly measure what students know when they arrive at the beginning of the year and what they know at the end of the year. As a result, researchers were able to use databanks to track yearly progress for thousands of students matched with specific teachers. The findings were quite conclusive: teachers, both effective and ineffective do matter a great deal in the academic success of students (Sanders & Rivers, 1996).

There are basically two systems that render a fairer representation of test results than raw averages: regression and mixed models (Darlington, 1997; Sanders, 2000). In the regression approach, data are collected for each student that might predict a student's score on end-of-year achievement tests. The data may include previous test scores, IQ scores, attendance, English as a Second Language, absences from class, socio-economic status (eligibility for free/reduced priced meals), in or out of school suspensions, and whether the parents live in the same residence. The regression method combines all of these data points to predict a child's achievement test score with the average attributed to the teacher. The "school effect" is then computed using the teacher averages and the "district effect" encompasses the school averages (Darlington, 1997; Sanders, 2000). Researchers denote two major concerns of the regression model. The first concern is that the use of socioeconomics may open the door for low or different expectations (Darlington, 1997; Sanders, 2000). Sanders (2000) used himself as an example. If socioeconomics were included as a predictor of academic achievement, then because he lived in a community that had less income than any district in the county, he may have been expected to achieve less than others even his wealthier counter parts with similar abilities who attended the same high school. The second concern with regression models is that require complete data sets. If a child has missing data from one prediction variable, that child is typically left out of the analysis. Some of the most fragile students in a district miss a lot of school or frequently move. Lower scoring students disproportionately miss more school than high scoring students; therefore, the data used for the regression analysis is often a truncated sample of the district or school's student population which can result in an over estimate of student achievement (Sanders, 2000).

The mixed model approach can be used if multiple years of achievement results in scale scores from norm- or criterion-referenced tests are available and if the achievement tests are highly aligned with curricular objectives. The process incorporates methods appropriate for longitudinal analyses with lean or incomplete the data sets are for each child. This statistical mixed model seeks to eliminate the shortcomings of other "value-added" assessment approaches (Sanders et al., 1997). One great example of the use of a method that links the progress of students over time to the teachers who taught them and measures the impact of instruction on a student's academic growth is the Tennessee Value-Added Assessment (TVAAS). TVAAS was created by law in Tennessee in the early 1990s to determine the effectiveness of districts, schools and teachers using an unbiased approach (Sanders & Horn, 1998). In the system, a statistical mixed-model and methodology are used to conduct a multivariate, longitudinal analysis of student achievement results. These data include student scores on the Tennessee Comprehensive Assessment Program which includes a group of five tests in math, science, social studies, reading, and language arts in grades three through eight and two end-of-course tests in high school subjects.

The method is a “value-added” system because it is designed to measure the additional amount of learning that a district, school or teacher adds to their students during a given school year in annual tests in different subjects. In other words, the effectiveness measures are based on the academic gain of students from the beginning of the year to the end. Each student’s test data are accumulated over time and are linked to that student’s teacher(s), school(s), and school system(s) to depict learning patterns. Because individual students rather than cohorts are tracked over time, each student becomes his or her own “baseline” or control. This strategy virtually removes all the influence of characteristics such as race or socio-economic indicators (Sanders, 2000).

TVAAS compares the actual and expected growth in student learning or the normal amount of academic growth that a typical student is expected to make in a given subject and grade. The anticipated progress (variance) is statistically controlled and adjusted up or down based on the previous history of each student (Sanders & Horn, 1998). For example, if a teacher has a student that has traditionally struggled to make progress over time the amount of growth that teacher is expected to help that student achieve is adjusted down. In essence, non-teacher variables that may affect student learning (e. g., home situations) are screened out thus isolating the teacher’s influence.

Although there is more than 10 years of TVAAS data that show some teachers are much more effective than others, the system is not perfect. Student test scores are only one estimate of a student’s knowledge; therefore, those scores never capture all aspects of learning and development. To address this limitation, TVAAS designers have instituted considerable safeguards to ensure accuracy in the results of the system (Carey, 2004; Sanders & Horn, 1998).

First, the effectiveness measures are based on multiple years of data and account for the different learning history of each student. In addition, the system takes into account the amount of student achievement data for each teacher. If there is not ample data to provide a reliable rating, the system will make adjustments. For example, a new teacher will not have multiple years of achievement data; therefore, the system will give the teacher the benefit of the doubt and conclude that the new teacher’s performance is the same as the system’s average.

In sum, because of the adjustments made to ensure accuracy and the fact that value-added ratings are consistent over time, TVAAS is a viable system for measuring teacher effectiveness and provides reliable information about which teachers are most effective in helping students to grow academically (Bock & Wolfe, 1996). TVAAS is now referred to as Educational Value Assessment System (EVAAS) by the SAS Institute.

## Do Good Teachers Matter?

In a study conducted in Tennessee that measured the cumulative and residual effects of teachers on student achievement, researchers estimated teacher effects for teachers who taught mathematics in grades three, four and five using a statistical mixed model approach (Sanders & Rivers, 1996). In addition, the teachers' effects were divided into five quintiles, with the least effective teachers in the first quintile and the most effective in the fifth. Student records were matched with their teachers so that the students' progress could be tracked through sequences of teachers. The findings indicated that students assigned to ineffective teachers continued to show the effects of such teachers even when these students were assigned to very effective teachers in subsequent years. Two years after the fact, the performance of fifth grade students was still impacted by the quality of their third grade teachers.

Further, when the data were aggregated by student achievement level, it was found that ineffective teachers were ineffective with all students, regardless of prior achievement levels. On average, the least effective teachers in the first quintile produced gains of about 14 percentile points in achievement results during the school year and the most effective teachers, in the fifth quintile averaged gains of 53 points among low-achieving students.

Dramatic differences were also shown for middle and high achieving students. High achieving students with the least effective teachers posted an average gain of only two points and an average of 25 points with the most effective teachers. Middle achieving students gained an average of 10 points with teachers in the first quintile and in the mid 30s with the most effective teachers in the fifth quintile.

Another interesting finding was African American and white students with the same level of academic achievement made comparable academic progress when they were assigned to teachers of comparable effectiveness. However, the study revealed that African American students in the system studied were disproportionately assigned to the least effective teachers.

Similar findings in achievement were found in a variety of studies in Texas (Jordan et al., 1997). Using some of Sanders' techniques, researchers in the Dallas school district found that the average fourth grade reading scores of students who were assigned to three highly effective teachers in a row rose from the 59th percentile in fourth grade to the 76th percentile in the sixth grade. A similar but slightly higher achieving group of students was assigned to three consecutive ineffective teachers and dropped from the 60th percentile in fourth grade to the 42nd percentile by the end of the sixth grade.

The teacher effect findings in mathematics were also compelling. A group of Dallas beginning third graders who averaged close to the 55th percentile in mathematics scored close to the 76th percentile at the end of fifth grade after being assigned to highly effective teachers for three consecutive years. In contrast, a slightly higher achieving group of third graders taught by three ineffective teachers averaged around the 57th percentile. At the end of fifth grade, the group's ranking fell to the 27th percentile. These are startling findings: not only did the youngsters with the least effective teachers regress by 30 points; they were 50 percentile points lower three years later than their counterparts who had effective teachers for three consecutive years.

In a study using the Tennessee Value-Added Assessment System (TVAAS), fourth grade student math achievement scores and TVAAS effectiveness estimates for math teachers in grades five through eight were used to predict the impact of different sequences of teacher effectiveness on the students' probability of passing the ninth grade competency test (Rivers, 1999). Data from two urban Tennessee school districts were linked longitudinally for each student. Students received a quartile ranking on their fourth grade achievement math scores. Students in the first quartile (Q1) were defined as low achieving and scores fell within the $1^{st}$ and $25^{th}$ percentile. Second quartile (Q2) students were considered below average ($26^{th}$- $50^{th}$ percentile); above average students ($51^{st}$-$75^{th}$ percentile) fell in the third quartile (Q3), and high achieving students ($76^{th}$-$99^{th}$ percentile) were in the fourth quartile (Q4). Each student's record was encoded with a success variable for each of the following assumed cut scores: 60, 65, 70, 75, and 80. If a student, for example, earned a Competency math score of 73, the success variables for 60, 65, and 70 were coded a "1" to reflect success; and the success variables 75 and 80 were coded "0" to reflect failure. Teacher effectiveness estimates were estimated in years prior to the ones when the students in the study were assigned to them. Ineffective teachers' estimates fell within the bottom 25% (low) of the teacher distribution; average teachers fell in the 25-75% category, and effective teachers were in the top 25% (high). Analyses showed the sequence of teachers was a highly significant predictor of a student's probability of passing achievement tests at all achievement levels. The students below the 50th percentile in fourth grade, however, were at a greater risk of failing the minimum competency test due to their teacher sequence than their peers at higher achievement levels.

**The State of Effective Teachers and Their Distribution**

The evidence from the previously described studies overwhelmingly indicates that good teachers definitely matter. The use of value-added data provides an opportunity to identify and equitably distribute effective teachers in schools that need them the most, and will assist the nation in eliminating the achievement divide. The achievement divide exists primarily between children of color and/or in poverty and their white counterparts. Given that reality, which teachers are then assigned to teach children that are stricken by poverty and of color?

In the Tennessee Value-Added study, for example, African American students were disproportionately assigned to ineffective teachers (Sanders & Rivers, 1996). The study was clear that, regardless of race, students who are assigned disproportionately to ineffective teachers will be academically behind their peers even with other teacher assignment patterns.

In analyzing value-added data from Dallas, Texas, the impact of the inequitable distribution of effective teachers is devastating. The study examined the performance of different middle school students, assigned to different teachers and how the teacher assignments affected the students' performance in mathematics (Babu & Mendro, 2003). The study analyzed the performance of two groups of seventh graders on the state's 2000 seventh grade mathematics test. One group was only assigned to effective math teachers during fifth, sixth, and seventh grades. By contrast, the other group was only assigned to ineffective teachers during those same years. The students' previous math performance categorized as low, middle, and high prior to the beginning of fifth grade was also examined. Almost twice as many or 77% of the previously high achieving students were assigned to a series of effective teachers compared to 40% of the low-achieving students. By contrast, more than twice as many low-achieving students were assigned to a sequence of ineffective teachers as high achieving students (81% vs. 30%).

In addition to the dramatic figures regarding assignment of teachers, each mid-to high achieving student passed the test, while only 42% of the previously low-achieving students who were taught by ineffective teachers passed the test. What is even more telling is that the previous low-achieving students who had effective teachers for three consecutive years experienced a 90% passing rate.

In addition, low performing students in every grade from third to eighth in both reading and math with effective teachers passed at much higher rates than low performing students taught by ineffective teachers for three years in a row.

Even with the compelling findings of the previous described studies, only a few states and districts are using value-added data to inform decisions regarding distribution of effective teachers. The sparse use of value-added data leads to limited information regarding the distribution of effective teachers with children at different readiness levels and other attributes. However, characteristics and practices of effective teachers have been the focus of much research over the past half century (Darling-Hammond, Wise, & Klein, 1995). Although the findings have been mixed, a few themes emerged as practices. Among the practices are content knowledge, education coursework, experience, certification, quality of learning in preparation programs, and test performance.

### ***Knowledge of the Content/Subject Matter***

Knowledge of the content or subject matter by the teacher has been found to be strongly related to student achievement. In a recent study, Hill, Rowan, and Ball (2005) noted that despite the growing interest and concern about the knowledge of the subject matter, there are few studies that address what counts as subject matter knowledge and how it relates to student achievement. As a result, the trio attempted to bridge the gap by analyzing teachers' scores on a measure of mathematical knowledge for teaching (Hill, Rowan, & Ball, 2005). An important purpose of the study was to separate the contribution of teachers' mathematical knowledge for teaching student achievement from other possible measures of teacher quality such as teacher certification, coursework, and experience. The study explored whether and how teachers' mathematical knowledge for teaching contributes to gains in students' mathematical achievement.

The researchers collected survey and student achievement data from first through third grade students and teachers in 115 elementary schools during the

2000-2001 through the 2003-2004 school years. The results of the study indicated that teachers' mathematical knowledge was significantly related to student achievement gains in both first and third grades after controlling for critical student and teacher level covariates. Although the results indicated a strong positive effect on content knowledge and student achievement, there were considerable limitations. Despite the researchers success in identifying a positive relationship between mathematical knowledge for teaching and student gain scores, the possibility remains that the gain could be attributed to general knowledge or aptitude versus content-specific knowledge. Even with the limitations, the study found that teachers' content knowledge for teaching mathematics was a significant predictor of student gains in both first and third grades. The average first and third grader gained close to 58 points and 39 points on the Terra Nova scale respectively.

While Hill et al. (2005) found that content knowledge is critical to the success of a student in mathematics, the findings of other previous work are not conclusive. Studies of teachers' scores on subject matter tests of the National Teacher Examinations have found no consistent relationship between the results of the exam and teacher performance as measured by student outcomes or supervisory ratings (Darling-Hammond et al., 1995). Most studies show statistically insignificant positive or negative relationships between teacher's knowledge and student outcomes (Andrews, Blackmon, & Mackey, 1980; Haney, Madaus, & Kreitzer, 1987; Summers & Wolfe, 1975).

The results of thirty studies relating to teachers' subject matter knowledge to student achievement were summarized by Byrne (1983). The results of the studies varied; 14 showed no relationship and 17 showed a positive relationship. Ashton and Crocker (1987) found only 5 of 14 studies they examined showed a positive relationship between measures of subject matter and teacher performance. In a multilevel analysis, Monk and King (1994) confirmed both positive and negative but insignificant effects of teachers' subject matter preparation on student achievement.

### *Education Coursework*

Unlike the lukewarm findings regarding subject matter knowledge, studies have found a positive relationship between education coursework and teachers' effectiveness. In a review of seven studies, Ashton and Crocker (1987) found significant positive relationships between education coursework and teacher performance in four of the studies. Begle (1979) in a study of the National Longitudinal Study of Mathematical Abilities found the number of credits a teacher had in mathematics methods courses was a stronger correlate of student performance than was the number of credits in mathematics courses. In a program-based study by Denton and Lacina (1984) a positive relationship between the extent of teachers' professional education coursework and their teaching performance and students' achievement. Based on these and similar findings, positive effects of subject matter knowledge are enhanced or off set by the knowledge of how to teach the subject to various kinds of students. In short, the degree to which one teaches combined with subject matter knowledge may augment or lessen teacher performance. Byrne (1983) summed this phenomenon up perfectly:

> It is surely plausible to suggest that insofar as a teacher's knowledge provides the basis for his or her effectiveness, the most relevant knowledge will be that which concerns the particular topic being taught and the relevant pedagogical strategies for teaching it to the particular types of pupils to whom it will be taught. If the teacher is to teach fractions, then it is knowledge of fractions and perhaps of closely associated topics which are of major importance. Similarly, knowledge of teaching strategies relevant to teaching fractions will be important. (p. 25)

There is some convincing evidence that the knowledge of the content and teaching and learning practices are positively correlated with student achievement. Yet, in this country, one out of four high school courses in the core subjects is being taught by teachers who did not major, and in many instances, minor in the field (Jerald & Ingersoll, 2002). In schools heavily impacted by poverty, the ratio is one in three. An out-of-field teacher is 77% more likely to be assigned to high poverty classrooms than to low poverty ones. A similar pattern is seen for minority students with 21% of the courses taught by teachers without a major or minor in the field in low-minority high schools compared to 29% in high-minority schools. For schools that have a high concentration of African American students (more than 90%), the statistic jumps to 35%.

***Experience in Teaching***

Studies have shown that inexperienced teachers are less effective than their peers (Carey, 2004; Rivkin, Hanuskek, & Kain, 2002). Yet, across this nation, students in high poverty or high-minority schools are almost twice as likely as other students (20% vs. 11%) to get an inexperienced teacher (National Center for Education Statistics, 2000). This trend is confirmed as we examine teacher distribution across states. In a high poverty school in Texas, for example, students are 20% more likely to have teachers with one year or less of experience than students in low-poverty schools (Carey, 2004; Rivkin et al., 2002). However, the Texas study did not investigate the correlation between teacher effectiveness estimates and years of experience. In 1999, 23% of teachers in New York City had fewer than three years of experience compared to 14% in nearby Lower Hudson and Long Island. California illustrates a much more dismal picture. Students in high-poverty, high minority schools are almost twice as likely to have a teacher in his/her first or second year of teaching as their counterparts attending majority white schools (Carey, 2004; Lankford, Wyckoff, & Papa, 2000).

***Teacher Certification and Undergraduate Experience***

Although all states tout a teacher certification program, many schools employ uncredentialed or teachers that are not fully certified. Uncredentialed teachers in high-poverty schools are hired at a rate of 61% higher than in all other districts nationwide (U. S. Department of Education, 2003). In high-poverty schools in California, more than 28% of the African American students are taught by uncertified teachers (Jepsen & Rivkin, 2002). In high-poverty schools in New York, 13,357 out of 114,638 or 12% of the teachers are not certified. In the rest of the state, 143 teachers out of 103,875 are uncertified. This means that 99% of the uncertified teachers in New York are teaching in high-poverty schools.

Not only are there large numbers of uncredentialed teachers teaching the nation’s most needy students, there are more teachers who graduated from “non-competitive” universities teaching in high poverty schools than teachers in low-poverty schools. Nationally, 21% of the teachers in the lowest poverty schools attended undergraduate schools deemed as “non-competitive,” compared to 39% in high-poverty schools (Carey, 2004; Wayne, 2002).

In conclusion, while many states’ have minimal requirements of certification, teachers that do not even meet the minimal standards are routinely assigned to teach in high-minority schools and/or students impacted heavily by poverty at a disproportionate rate. In addition, these teachers are more likely to have graduated from a “non-competitive” teacher preparation program or university.

### ***Achievement Performance as Measured by Standardized Tests***

When examining standardized test results, teachers who teach low-wealth and/or minority students are less likely to have performed well on teacher licensing tests, tests of basic skills, and college entrance exams (Kain & Singleton, 1996). In Illinois, for example, teachers who failed the state teacher licensure exam at least once are five times as likely to teach children in high-poverty schools. Teachers who failed the licensure test at least five times are 23 times as likely to teach children in poverty. In New York, 21% of the teachers who teach minority students failed one of the state's licensure exams, compared to 7% of those who teach white students. In addition, one study showed that 34% of the new teachers in New York high-poverty schools scored in the bottom quartile on the SAT compared to 8% in the top quartile. On the other hand, 9% of the teachers assigned to high wealth schools performed in the bottom quartile compared to 23% in the top quartile on the SAT (Carey, 2004; Shen, 2003).

In looking further into the distribution of high quality teachers, the National Center for Education Statistics (2000), reports that minority students get more inexperienced teachers with three years or less experience than white students (21% vs. 10%). In a study using 1999-2000 Schools and Staffing Survey, more mathematics classes were found to have been taught by teachers lacking a major in the field: 41% in comparison to 29% (Jerald & Ingersoll, 2002). One study examining middle schools, found that the higher the percentage of African American students the higher the percentage of teachers teaching out of field. In schools with 90% or higher African American students, 62% of the teachers were out of field; schools with 11-89% African American, 50% of the teachers were out of field, and schools with 10% or lower African American, the percentage was 44% (Jerald & Ingersoll, 2002).

The evidence is quite compelling: no matter how you define teacher quality or effectiveness, the pattern of distribution is generally similar. Low-wealth students, under achieving students and students of color are far more likely to have teachers who are non-credentialed, lack experience, educated poorly, and under performing (Lankford, Loeb, & Wyckoff, 2002).

# CHAPTER III

# RESEARCH DESIGN AND METHODS

## Purpose of Study

The goal of this study was to determine the degree to which teacher assignments affect students' performance on Algebra I and Eighth Grade Math End-of-Grade tests. The study provided estimates of the effect of a series of effective or ineffective teachers on the students' scores. Further, the study investigated the relationship between teacher effectiveness scores and teacher years of experience, and examined the distribution of teachers based on teacher effectiveness estimates across the Guilford County public school system.

## Research Questions

Based on the experiences of the researcher and review of the literature presented in Chapter II, three questions were put forth to investigate the relationship of teacher effectiveness and achievement:

Research Question 1: What is the relationship between sequences of teacher effectiveness as measured by SAS Educational Value Added Assessment System (EVAAS) or teacher effect scores and students' achievement as measured by their performance on the Algebra I End-of-Course and Eighth Grade Math End-of-Grade tests?

Research Question 2: What is the relationship between teacher effect data and teachers' years of experience?

Research Question 3: Based on teacher effectiveness estimates, how are teachers distributed throughout the district?

## Research Approach

This study is quantitative in design because of the desire to determine relationships between variables: sequence of teacher effectiveness and student achievement, and teacher effectiveness and years of teaching.

In addition, this study is non-experimental in design because the groups of students from which the data are derived are not assigned by the researcher; the groups of students were intact at the time of this investigation. In true experimental designs, students are randomly assigned to the treatment groups (Wiersma & Jurs, 2005). Non-experimental quantitative research is conducted in natural settings with many variables operating simultaneously. As a result, interpreting the results may be less clear cut than interpreting experimental research results. However, Wiesma and Jurs (2005) note that even with the ambiguous nature of non-experimental results, the research can be "designed to enhance not only completion of the research but also interpretation of the results. It is the research problem and the conditions of the research that determine the appropriate methodology" (p. 155).

Further, this study is described as ex post facto. When ex post facto research is conducted, variables are explored in retrospect to investigate possible relationships and effects (Wiersma & Jurs, 2005). The manipulation of variables by the researcher is absent from this type of research.

**Description of Sample**

Achievement scores of all eighth grade students who participated in 2005 Algebra I and Math End-of-Grade (EOG) testing and scores included in the computation of 2004-2005 AYP results were matched with teacher records in the value added database. The value added database contains achievement results in scale scores for every student who was tested in grades 3-12 and the students' respective teachers. For this study, two cohorts of data were examined:

1. Cohort 1 includes 2005 8$^{th}$ grade Algebra I scores matched with the students' seventh through fifth grade math teachers value added or teacher effect data, and fourth grade student math EOG scores.
2. Cohort 2 includes 2005 EOG scores matched with the students' seventh through fifth grade math teacher effect data, and fourth grade student math EOG scores.

In 2005, 3,533 students in Cohort 1 took the Algebra 1 test in 8$^{th}$ grade; 2118 scores were matched with teacher estimates and students have 4$^{th}$ grade math EOG scores in 2001. In Cohort 2, 5101 students took the math Eighth Grade EOG in 2005; 2900 scores were matched with teacher estimates and 4$^{th}$ grade math EOG scores in 2001. Students of teachers for whom SAS EVAAS estimates are not available will be excluded from the study. Table 2 captures the description of the cohorts.

**Table 2**

***Description of Data Included in Study***

| Cohort 1<br>2005 8th Grade Algebra I Scores | Cohort 2<br>2005 8th Grade End-of-Grade Scores |
|---|---|
| 2004 7th Grade Math Teacher Effect | 2004 7th Grade Algebra Teacher Effect |
| 2003 6th Grade Math Teacher Effect | 2003 6th Grade Math Teacher Effect |
| 2002 5th Grade Math Teacher Effect | 2002 5th Grade Math Teacher Effect |
| 2001 4th Grade Math Student EOG Scores | 2001 4th Grade Math Student EOG Scores |
| N = 2118 (matched or complete records) | N = 2900 (matched or complete records) |

## Methodology and Data Analysis Plan

This study was patterned after a Tennessee Value-added study conducted by Rivers (1999). The purpose of the 1999 study was to investigate the effect of a student's series of teachers on his/her mathematics competency score and predict the probability of a student passing the competency test by varying the cut-off scores. Although, estimates of teacher effectiveness were examined with mathematics achievement scores, the Rivers' study did not examine the relationship between teacher effectiveness and years of experience nor was there an examination of the distribution of effective teachers. This study examined the sequence of teacher effectiveness estimates, years of experience, and distribution of teachers. Therefore, the methodologies of the two studies were similar.

## Determining the Relationship between Sequences of Teacher Effectiveness and Student Achievement

To determine the relationship between sequences of teacher effectiveness and student achievement, SAS Institute used the mixed model, longitudinal process and achievement data from cohorts of students who took the 2005 Algebra I and Math End-of-Grade tests in the eighth grade to provide shrinkage estimates of teacher effects. After the teacher effects were obtained for each grade level, the distribution of teachers were arbitrarily grouped into three tertiles, with the teachers demonstrating the lowest degree of effectiveness in the first tertile and the teachers demonstrating the greatest degree of effectiveness in the third tertile. SAS EVAAS teacher effects were reported for individual years and in three-year averages. Each year's estimates were calculated using all available data from both current and previous years to provide the most accurate measurement of teachers' influence on student achievement (see Table 2). The inclusion of data from previous years ensured that teachers benefited from the most accurate estimate of their individual contribution. To minimize bias, the most recently calculated SAS EVAAS estimates of math teacher effects were computed using the most recent teacher's student data.

The achievement data for these analyses contained math scores, district and school numbers, the student's names, social security numbers, grades, birth dates, gender, ethnicity, and a special education indicator. The mathematics records at the student level (name, social security number and birth date) were matched with records in the SAS EVAAS database. To reduce the number of incomplete student records, the district's entire database was searched to locate records for students who might have moved from one school to another within the district prior to taking the math tests.

Tertile regressions were computed by ranking students within each school and assigning a tertile rank based on their fifth or fourth grade performance on the End-of-Grade mathematics test; the top 25% of students received the highest rank (T 3), students in middle were in the second tertile (T2), and the bottom 25% were ranked in the first tertile (T1). These tertile rankings were added to the student records as indicators of prior student achievement levels.

The SAS EVAAS numerical teacher identifiers for mathematics teachers for grade levels four through eight (beginning with 2001) were added to the individual student records. The grade and numerical teacher identifiers were used to match SAS EVAAS teacher effect estimates to the individual Algebra I and math EOG records by each student-year. Students with no teacher identifier were excluded from the analyses. The teacher effect estimates for each grade for each school in the district were ranked and divided into tertiles with the least effective teachers in Tertile 1, the middle 50% of the teachers in Tertile 2, and the most effective teachers in Tertile 3. The teacher effectiveness rankings for the appropriate grade/year were added to the individual student achievement records to facilitate both the analyses and the reporting of results.

Using fourth grade math achievement scores (dependent variable) and SAS EVAAS teacher estimates, it was possible to show the impact of different sequences of teacher effectiveness on the students' performance on Algebra I and Eighth Grade Math End-of-Grade scores. Least-squares means and contrast statements were used in Estimable Functions to examine teacher sequences.

The relationship between the variables was examined with correlations. Scatter plots are displayed in Appendix A. Additionally, a univariate procedure was used to determine the normality of each data set (see Appendix B). Multiple regression was used to ascertain the cumulative effect of the sequences of teacher effectiveness on student achievement as measured by success on Algebra I and Eighth Grade Math EOG tests. The models used to determine the relationship between the math achievement scores and teacher effectiveness included the following variables or covariates: (1) fourth grade mathematics End-of-Grade achievement scores, and student rankings, and (2) fifth, sixth, and seventh grade teacher tertile rankings (T1-bottom 25%, T2-middle 50%, T2-top 25%).

The distributions of the continuous variables were evaluated to determine the various combinations of student scores and levels of teacher effectiveness to be included in the analyses. Student gains averaged by achievement level of the students were cross tabulated with teacher effectiveness. Additionally, Estimable Functions were used to provide estimates of the cumulative effect of the sequence of teachers, for example, three consecutive effective teachers (top tertile), three consecutive average teachers (second tertile) and three consecutive ineffective teachers (bottom tertile) on Algebra I for each of the three groups.

## Determining the Relationship between Teacher Effectiveness Data and Teachers' Years of Experience

To determine the relationship between teacher effectiveness and teachers' years of experience, the second research question, the distribution of the frequency was examined to see if patterns emerged. Teacher identifiers used to ascertain teacher effectiveness data were matched with a teacher database with the same identifiers and years of experience of teachers employed in 2005 in Guilford County Schools. The years of experience were categorized as follows: 0-2, 3-5, 6-10, 11-20, and 21-99. To further understand this analysis, a Cochran-Mantel-Haenszel Chi Square statistical test was run. The test examined if there were a difference in the average classification of teacher effectiveness across the different years of experience groups.

The matching of the teacher data received from SAS Institute was completed by a member of the HR staff; the researcher was not privy to any teacher identifiers.

## Determining the Distribution of Teachers in the District

The final research question was also explored by generating frequency tables using 2005 teacher effectiveness data and schools' free/reduced price lunch percentages. The schools' free/reduced price lunch percentages were reported in the following categories: 0-19%, 20-39%, 40-59%, 60-79%, and 80-99%; the percentage of teachers by tertile rankings that were assigned to schools within the free/reduced categories were computed. For comparison purposes, a graph was generated to display the percentage of teachers by extreme tertiles in 0-19 and 80-99 percent schools. The free/reduced data was obtained from the district's Title I office.

## Limitations

Many of the students who took the Algebra I and Eighth Grade End-of-Grade tests did not have subsequent achievement data through the fourth grade. These student records were not included in the analyses.

In terms of incomplete data, there were also a number of teachers for which years of experience could not be ascertained due to incomplete teacher identifiers. To that end, trends should be interpreted with caution. Of the 427 teachers with available data, only 255 had known years of experience. While it appeared that the 'unknown' teachers were distributed across the different tertiles, this large proportion of missing data, combined with a marginally significant p-value of 0.0327 may call into question any conclusions regarding statistical relationships.

# CHAPTER IV

# RESULTS

This chapter presents the results of the analyses used in this study. The results are described in three sections that coincide with the three research questions. First, the results from the base models which illustrate the impact of a series of teachers as teacher sequences on student achievement will be presented. Second, the relationship between years of experience and teacher effectiveness will be described. Finally, the analysis of the distribution of teachers across the district will be presented.

## Research Question 1

*What is the relationship between sequences of teacher effectiveness as measured by SAS Effectiveness Value Added Assessment System (EVAAS) value-added or teacher effect scores and students' achievement as measured by their performance on the Algebra I End-of-Course and Eighth Grade Math End-of-Grade tests?*

The first analysis, as presented in Table 3, examined the relationship between the effectiveness of fifth grade teachers and Eighth Grade End-of-Grade (EOG) Algebra I End-of-Course (EOC) scores. EOC scores were matched with 4$^{th}$ grade math EOG scores and teacher estimates. After adjusting for the entering achievement of fourth grade students, the teacher effectiveness estimates of the fifth grade teachers were significant ($p < .01$). In other words, the fifth grade teachers influenced the achievement of eighth grade Algebra I students even after adjusting for the starting point of the students in the fourth grade.

Table 4 depicts the second analysis which examined the influence of the fifth and sixth grade teachers on the achievement of eighth grade Algebra I students. After adjusting for entering achievement of the fourth grade math End-of-Grade tests, the fifth and sixth grade effect scores were both significant ($p < .01$).

**Table 3**

***Analysis of Algebra I Scores and Grade Five Teacher Effectiveness Estimates***

| Dependent Variable: Alg I EOC Source | DF | Sum of Squares | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Model | 4 | 76997.5933 | 19249.3983 | 396.45 | <.0001 |
| Error | 2114 | 102645.1287 | 48.5549 | | |
| Corrected Total | 2118 | 179642.7220 | | | |
| R-Square | Coeff Var | Root MSE | Alg I Mean | | |
| 0.428615 | 11.58295 | 6.968137 | 60.15857 | | |
| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
| Student Ranks | 2 | 491.94394 | 245.97197 | 5.07 | 0.0064 |
| EOG Math 4 | 1 | 17756.77424 | 17756.77424 | 365.70 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 641.51936 | 641.51936 | 13.21 | 0.0003 |

**Table 4**

***Analysis of Algebra I Scores and Grades Five and Six Teacher Effectiveness Estimates***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 522.36476 | 261.1823 | 5.47 | 0.0043 |
| EOG Math 4 | 1 | 17671.23370 | 17671.23370 | 370.41 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 836.64417 | 836.64417 | 17.54 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 1838.80060 | 1838.80060 | 38.54 | <.0001 |

The effects of teachers in grades five, six, and seven were significant ($p < .01$; see Table 5). These findings confirmed the importance of the influence that teachers have on students three years after they leave the fifth grade. There is a residual effect of the fifth and sixth grade teacher on the End-of-Course Algebra I score, even after the effectiveness of the seventh grade teacher is taken into account.

An analysis was conducted to ascertain if there were interactions between the teachers for different grade levels (e.g. fifth and sixth). No additional effects were found beyond the individual teacher (see Table 6). These findings provide some evidence that there does not appear to be a “catch-up” effect for students who had a bad teacher and then a good one. There is also no magnifying effect of having two consecutive good or bad teachers above and beyond the fact that those teachers were good or bad.

**Table 5**

***Analysis of Algebra I Scores and Grades Five, Six, and Seven Teacher Effectiveness Estimates***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 462.24243 | 231.12122 | 4.89 | 0.0076 |
| EOG Math 4 | 1 | 17040.66512 | 17040.66512 | 360.40 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 771.97311 | 771.97311 | 16.32 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 1808.95915 | 1808.95915 | 38.54 | <.0001 |
| Grade 7 Teacher Estimates | 1 | 891.91996 | 891.91996 | 18.85 | <.0001 |

To corroborate the findings of the significance of the teacher effect scores, an analysis was run to examine if the teacher effect scores were impacted by the entering achievement levels of the students. As shown in Table 7, there were no significant interactions found between fourth grade math scores and teacher effect scores for fifth and sixth grade teachers. A seventh grade teacher may have a different effect on student achievement dependent on the entering achievement of the fourth grade students ($p < .01$).

**Table 6**

***Grades Five, Six, and Seven Teacher Interactions (Algebra 1, Grade 8)***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 457.79566 | 228.89783 | 4.83 | 0.0080 |
| EOG Math 4 | 1 | 17035.03220 | 17035.03220 | 359.69 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 780.52986 | 780.52986 | 16.48 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 1776.70772 | 1776.70772 | 37.52 | <.0001 |
| Grade 7 Teacher Estimates | 1 | 801.97739 | 801.97739 | 16.93 | <.0001 |
| Grade 5 Teacher Estimates *Grade 6 Teacher Estimates | 1 | 2.13551 | 2.13551 | 0.05 | 0.8319 |
| Grade 5 Teacher Estimates *Grade 7 Teacher Estimates | 1 | 2.23680 | 2.23680 | 0.05 | 0.8280 |
| Grade 6 Teacher Estimates *Grade 7 Teacher Estimates | 1 | 0.13449 | 0.13449 | 0.00 | 0.9575 |
| Grade 5 Teacher Estimates*Grade 6 Teacher Estimates *Grade 7 Teacher Estimates | 1 | 72.20115 | 72.20115 | 1.52 | 0.2171 |

The next analysis included student ranks, fourth grade math EOG scores and the combinations of teacher effectiveness experienced by students in grades five through seven in terms of teacher tertiles or rankings. There were 27

**Table 7**

***Analysis of Teacher Effect Scores and Entering Achievement of Students***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 386.74116 | 193.37058 | 4.10 | 0.0166 |
| EOG Math 4 | 1 | 16742.66618 | 16742.66618 | 355.28 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 844.93585 | 844.93585 | 17.93 | <.0001 |
| EOG Math 4*Grade 5 Teacher Estimates | 1 | 24.11769 | 24.11769 | 0.51 | 0.4744 |
| Grade 6 Teacher Estimates | 1 | 1759.60701 | 1759.60701 | 37.34 | <.0001 |
| EOG Math 4*Grade 6 Teacher Estimates | 1 | 60.90258 | 60.90258 | 1.29 | 0.2557 |
| Grade 7 Teacher Estimates | 1 | 918.50511 | 918.50511 | 19.49 | <.0001 |
| EOG Math 4*Grade 7 Teacher Estimates | 1 | 439.99787 | 439.99787 | 9.34 | 0.0023 |

possible teacher sequences or combinations (3x3x3) according to teacher ranks or tertiles (1=bottom 25%, 2 = middle 50%, 3 = top 25%). The implications of the significant teacher effects were explored using contrast statements and least-squares means (LS means) of the Algebra scores. As presented in Table 8, students who had teachers in the lower tertiles generally scored lower than the students who had teachers in the higher tertiles.

**Table 8**

***Estimable Functions Used to Examine Teacher Sequences (Algebra 1 in Grade 8)***

| Teacher Ranks | | | | No T1 vs. One T1 | No T1 vs. Two T1 | No T1 vs. Three T1 |
|---|---|---|---|---|---|---|
| Grade 5 | Grade 7 | Grade 6 | LS Mean | Coeff | Coeff | Coeff |
| 1 | 1 | 1 | 58.149 | 0 | 0 | 8 |
| 1 | 2 | 1 | 59.720 | 0 | 4 | 0 |
| 1 | 3 | 1 | 61.070 | 0 | 4 | 0 |
| 1 | 1 | 2 | 59.618 | 0 | 4 | 0 |
| 1 | 2 | 2 | 60.159 | 2 | 0 | 0 |
| 1 | 3 | 2 | 60.173 | 2 | 0 | 0 |
| 1 | 1 | 3 | 57.822 | 0 | 4 | 0 |
| 1 | 2 | 3 | 60.877 | 2 | 0 | 0 |
| 1 | 3 | 3 | 61.701 | 2 | 0 | 0 |
| 2 | 1 | 1 | 59.105 | 0 | 4 | 0 |
| 2 | 2 | 1 | 59.121 | 2 | 0 | 0 |
| 2 | 3 | 1 | 59.683 | 2 | 0 | 0 |
| 2 | 1 | 2 | 59.127 | 2 | 0 | 0 |
| 2 | 2 | 2 | 62.089 | -3 | -3 | -1 |
| 2 | 3 | 2 | 60.836 | -3 | -3 | -1 |
| 2 | 1 | 3 | 58.179 | 2 | 0 | 0 |
| 2 | 2 | 3 | 61.563 | -3 | -3 | -1 |
| 2 | 3 | 3 | 61.127 | -3 | -3 | -1 |
| 3 | 1 | 1 | 58.107 | 0 | 4 | 0 |
| 3 | 2 | 1 | 60.875 | 2 | 0 | 0 |
| 3 | 3 | 1 | 62.382 | 2 | 0 | 0 |
| 3 | 1 | 2 | 60.705 | 2 | 0 | 0 |
| 3 | 2 | 2 | 60.540 | -3 | -3 | -1 |
| 3 | 3 | 2 | 61.942 | -3 | -3 | -1 |
| 3 | 1 | 3 | 59.557 | 2 | 0 | 0 |
| 3 | 2 | 3 | 61.700 | -3 | -3 | -1 |
| 3 | 3 | 3 | 61.930 | -3 | -3 | -1 |
| | | | | Divisor: 24 | Divisor: 24 | Divisor: 8 |
| | | | | Mean1: 60.211 | Mean 1: 59.240 | Mean 1: 58.149 |
| | | | | Mean 2: 61.466 | Mean 2: 61.466 | Mean 2: 61.466 |
| | | | | Diff in Means 1.255 | Diff in Means 2.226 | Diff in Means 3.317 |

Students who had no teachers in the bottom 25% for three years scored about 1.25 points higher on the Algebra test when compared with students who had one teacher in the bottom 25%. Similar changes for the rest of the groups that had a teacher in the bottom tertile tended to cause the students to score lower than students who had no teacher in the bottom 25%. Students who had two teachers in the bottom 25% scored 2.23 points lower compared with students with no teachers in the bottom 25% for three years. Similarly, students who had three teachers in the bottom tertile for three years compared with no teachers in the lowest tertile scored 3.32 points lower on the Algebra I test. Further, when examining the extremes, students who had all teachers in the bottom tertile versus all teachers in the highest tertile scored 3.78 points lower on the Algebra I test.

As shown in Table 9, the effect size (Cohen's *d*) for having one ineffective teacher versus no ineffective teachers was considered very small  and not very meaningful (.2 is small, .5 is medium and .8 is large). However, having two or three ineffective teachers was considered small but meaningful. In addition, the differences would have improved the students' percentile ranking in Algebra I considerably (see Appendix C).

A final test was run to determine whether the effect of teacher sequences differed based on the achievement levels of the students. As displayed in Table 10, there was no interaction between achievement levels of students and the

**Table 9**

***Teacher Effectiveness Effect Sizes (Algebra 1, Grade 8)***

| Teacher Tertile | Cohen's *d* |
|---|---|
| No T1 vs. One T1 | 0.124 |
| No T1 vs. Two T1 | 0.220 |
| No T1 vs. Three T1 | 0.328 |

**Table 10**

***Analysis of Algebra I Scores, Grades Five, Six, and Seven Teacher Effectiveness Estimates and Student Ranks***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 165.98407 | 82.99203 | 1.73 | 0.1778 |
| EOG Math 4 | 1 | 15995.43906 | 15995.43906 | 333.20 | <.0001 |
| GrdRnk*GrdRnk*GrdRnk | 26 | 2381.53337 | 91.59744 | 1.91 | 0.0038 |
| Std*GrdR*GrdR*GrdRn | 52 | 2326.95351 | 44.74911 | 0.93 | 0.6127 |

effect of sequence of teachers (p = 0.61). All students, low or high achieving, benefit from good teachers.

A similar analysis as the one used for examining teacher effects and Algebra I scores was conducted to examine the relationship between the effectiveness of fifth grade teachers and eighth grade student math achievement as measured by math End-of-Grade (EOG) tests (see Table 11). After adjusting for the entering achievement of fourth grade students, the teacher effectiveness estimates of the fifth grade teachers were significant (p<.01).

**Table 11**

***Analysis of Eighth Grade Math End-of-Grade Scores and Grade Five Teacher Effectiveness Estimates***

| Dependent Variable: End-of-Grade Math 08 | | | | | |
|---|---|---|---|---|---|
| Source | DF | Sum of Squares | Mean Square | F Value | Pr>F |
| Model | 4 | 237088.0940 | 59272.0235 | 1631.85 | <.0001 |
| Error | 2896 | 105188.5513 | 36.3220 | | |
| Corrected Total | 2900 | 342276.6453 | | | |
| R-Square | Coeff Var | Root MSE | EOG 08 Math Mean | | |
| 0.692680 | 2.208287 | 6.026775 | 272.9162 | | |
| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
| Student Ranks | 2 | 641.65887 | 320.82943 | 8.83 | 0.0001 |
| EOG Math 4 | 1 | 41293.03525 | 41293.03525 | 1136.86 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 2057.08202 | 2057.08202 | 56.63 | <.0001 |

An analysis depicted in Table 12 examined the influence of the fifth and sixth grade teachers on the achievement of eighth grade math students using

EOG scores. After adjusting for entering achievement of the fourth grade math EOG tests, the fifth and sixth grade effect scores were both significant (p<.01).

**Table 12**

***Analysis of Eighth Grade Math End-of-Grade Scores and Grades Five and Six Teacher Effectiveness Estimates***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 674.67538 | 337.33769 | 9.57 | <.0001 |
| EOG Math 4 | 1 | 40934.18930 | 40934.18930 | 1161.72 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 2414.80583 | 2414.80583 | 68.53 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 3181.16058 | 3181.16058 | 90.28 | <.0001 |

In examining the results from the next analysis (see Table 13), the effects of teachers in grades five, six, and seven were significant (p<.01). In addition, there was an effect of the fifth and sixth grade teacher on the End-of-Course Algebra I score even after the effectiveness of the seventh grade teacher is taken into account.

An analysis was conducted to ascertain if there were interactions between fifth and sixth, fifth and seventh, sixth and seventh, fifth, sixth and seventh, etc. There were no interactions found between the teacher effects at the three grade

**Table 13**

***Analysis of Eighth Grade Math End-of-Grade Scores and Grades Five, Six, and Seven Teacher Effectiveness Estimates***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 524.31095 | 262.15547 | 7.61 | 0.0005 |
| EOG Math 4 | 1 | 38844.00566 | 38844.00566 | 1128.07 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 2289.34360 | 2289.34360 | 66.48 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 2952.66400 | 2952.66400 | 85.75 | <.0001 |
| Grade 7 Teacher Estimates | 1 | 2355.07109 | 2355.07109 | 68.39 | <.0001 |

levels (see Table 14). Similar to the previous analysis, the results indicate that the order in which the students encountered their teachers did not matter. In addition, there were no additional effects found beyond the individual teacher. These findings suggest that there is no “catch-up” effect for students who had a bad teacher and then a good one. There is also no magnifying effect of having two consecutive good or bad teachers above and beyond the fact that those teachers were good or bad.

To corroborate the findings of the significance of the teacher effect scores, an analysis was run to examine if the teacher effect scores were impacted by the entering achievement levels of the students. Table 15 shows the interaction of

**Table 14**

***Grades Five, Six and Seven Teacher Interactions (Math EOG, Grade 8)***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 529.86241 | 264.93121 | 7.69 | 0.0005 |
| EOG Math 4 | 1 | 38668.90867 | 38668.90867 | 1122.95 | <.0001 |
| Grade 5 Teacher Estimates | 1 | 2246.68988 | 2246.68988 | 65.24 | <.0001 |
| Grade 6 Teacher Estimates | 1 | 2837.11629 | 2837.11629 | 82.39 | <.0001 |
| Grade 7 Teacher Estimates | 1 | 2437.63451 | 2437.63451 | 70.79 | <.0001 |
| Grade 5 Teacher Estimates *Grade 6 Teacher Estimates | 1 | 1.77498 | 1.77498 | 0.05 | 0.8204 |
| Grade 5 Teacher Estimates *Grade 7 Teacher Estimates | 1 | 2.97613 | 2.97613 | 0.09 | 0.7688 |
| Grade 6 Teacher Estimates *Grade 7 Teacher Estimates | 1 | 97.05473 | 97.05473 | 2.82 | 0.0933 |
| Grade 5 Teacher Estimates*Grade 6 * 7 Teacher Estimates | 1 | 32.71824 | 32.71824 | 0.95 | 0.3298 |

teacher effect and student achievement levels in tertiles. According to the estimates, higher achieving students experienced bigger jumps in their performance with better teachers.

**Table 15**

***Analysis of Teacher Effect Scores and Entering Achievement of Eighth Grade Math Students***

| Parameter | Estimate | Standard Error | t Value | Pr>\|t\| |
|---|---|---|---|---|
| Intercept | 273.4249164 B | 0.38290786 | 714.07 | <.0001 |
| Student Ranks 1 | -0.3437645 B | 0.73253000 | -0.47 | 0.6389 |
| Student Ranks 2 | -0.9592042 | 0.43996044 | -2.18 | 0.0293 |
| Student Ranks 3 | 0.0000000 B | | | |
| EOG Math 4 | 1.0082331 | 0.03034439 | 33.23 | <.0001 |
| Grade 5 Teacher Estimates* Student Ranks 1 | 0.9759547 | 0.24318600 | 4.01 | <.0001 |
| Grade 5 Teacher Estimates* Student Ranks 2 | 0.8599846 | 0.14876131 | 5.78 | <.0001 |
| Grade 5 Teacher Estimates* Student Ranks 3 | 0.9599495 | 0.21263273 | 4.51 | <.0001 |
| Grade 6 Teacher Estimates* Student Ranks 1 | 0.6252315 | 0.27100467 | 2.31 | 0.0211 |
| Grade 6 Teacher Estimates* Student Ranks 2 | 0.9327604 | 0.14996710 | 6.22 | <.0001 |
| Grade 6 Teacher Estimates* Student Ranks 3 | 1.14161460 | 0.20087366 | 7.05 | <.0001 |
| Grade 7 Teacher Estimates* Student Ranks 1 | 0.6231655 | 0.26086083 | 2.39 | 0.0170 |
| Grade 7 Teacher Estimates* Student Ranks 2 | 0.8134203 | 0.14906866 | 5.46 | <.0001 |
| Grade 7 Teacher Estimates* Student Ranks 3 | 1.4302186 | 0.22076430 | 6.48 | <.0001 |

The next analysis included student ranks, fourth grade math EOG scores and the combinations of teacher effectiveness experienced by students in grades five through seven in terms of teacher tertiles. There were 27 possible teacher sequences or combinations (3x3x3) according to teacher ranks. As depicted in Table 16, students who had teachers in the lower tertiles generally scored lower than the students who had teachers in the higher tertiles.

Students who had no teachers in the bottom 25% for three years scored 1.875 points higher on the math EOG test than students who had one teacher in the bottom 25%. Similar changes for the rest of the groups that had a teacher in the bottom tertile tended to cause the students to score lower than students who had no teacher in the bottom 25%. Students with two teachers in the bottom 25% scored 3.571 points lower than students who had no teachers in the bottom 25% for three years. Students who had three teachers in the bottom tertile scored 6.535 points lower on the math EOG test than students who had no teachers in the lowest tertile for three years. Further, when examining the extremes, students who had all teachers in the bottom tertile versus all teachers in the highest tertile scored 8.018 points lower on the Math EOG test. The effect sizes were small but meaningful for students who had one or two ineffective teachers. Having three ineffective teachers had a medium effect; however, having all ineffective teachers versus all effective teachers was close to a large effect (Cohen’s d = 0.74). In sum, the differences or increased scores would have improved the students’ percentile ranking on their math EOG even more substantially than the Algebra 1 rankings (see Appendix C).

As with the Algebra I results, a final test was run to consider the achievement levels of the students and their impact on teacher sequencing. As

**Table 16**

***Estimable Functions Used to Examine Teacher Sequences (Eighth Grade Math End-of-Grade Tests)***

| Teacher Ranks | | | | No T1 vs. One T1 | No T1vs. Two T1 | No T1 vs. Three T1 |
|---|---|---|---|---|---|---|
| Grade 5 | Grade 7 | Grade 6 | LS Mean | Coeff | Coeff | Coeff |
| 1 | 1 | 1 | 268.1531 | 0 | 0 | 8 |
| 1 | 2 | 1 | 270.5819 | 0 | 4 | 0 |
| 1 | 3 | 1 | 272.2281 | 0 | 4 | 0 |
| 1 | 1 | 2 | 270.6402 | 0 | 4 | 0 |
| 1 | 2 | 2 | 272.3176 | 2 | 0 | 0 |
| 1 | 3 | 2 | 273.6296 | 2 | 0 | 0 |
| 1 | 1 | 3 | 270.4969 | 0 | 4 | 0 |
| 1 | 2 | 3 | 273.1174 | 2 | 0 | 0 |
| 1 | 3 | 3 | 274.0388 | 2 | 0 | 0 |
| 2 | 1 | 1 | 270.6385 | 0 | 4 | 0 |
| 2 | 2 | 1 | 271.3354 | 2 | 0 | 0 |
| 2 | 3 | 1 | 272.7352 | 2 | 0 | 0 |
| 2 | 1 | 2 | 271.823 | 2 | 0 | 0 |
| 2 | 2 | 2 | 273.2823 | -3 | -3 | -1 |
| 2 | 3 | 2 | 274.4083 | -3 | -3 | -1 |
| 2 | 1 | 3 | 272.2632 | 2 | 0 | 0 |
| 2 | 2 | 3 | 274.0216 | -3 | -3 | -1 |
| 2 | 3 | 3 | 274.5156 | -3 | -3 | -1 |
| 3 | 1 | 1 | 271.9942 | 0 | 4 | 0 |
| 3 | 2 | 1 | 273.2232 | 2 | 0 | 0 |
| 3 | 3 | 1 | 273.9944 | 2 | 0 | 0 |
| 3 | 1 | 2 | 272.747 | 2 | 0 | 0 |
| 3 | 2 | 2 | 274.1534 | -3 | -3 | -1 |
| 3 | 3 | 2 | 276.2124 | -3 | -3 | -1 |
| 3 | 1 | 3 | 272.297 | 2 | 0 | 0 |
| 3 | 2 | 3 | 274.5797 | -3 | -3 | -1 |
| 3 | 3 | 3 | 276.1708 | -3 | -3 | -1 |
| | | | | Divisor: 24 | Divisor: 24 | Divisor: 8 |
| | | | | Mean 1: 272.793 | Mean 1: 271.097 | Mean 1: 268.153 |
| | | | | Mean 2: 274.668 | Mean 2: 274.668 | Mean 2: 274.668 |
| | | | | Diff in Means 1.875 | Diff in Means 3.571 | Diff in Means 6.535 |

displayed in Table 18, there was no interaction between achievement levels of students and the effect of sequence of teachers.

**Table 17**

***Teacher Effectiveness Effect Sizes (Grade 8 Math EOG)***

| Teacher Tertile | Cohen's *d* |
|---|---|
| No T1 vs. One T1 | 0.172 |
| No T1 vs. Two T1 | 0.328 |
| No T1 vs. Three T1 | 0.600 |

**Table 18**

***Analysis of Eighth Grade Math End-of-Grade Scores, Grades Five, Six, and Seven Teacher Effectiveness Estimates and Student Ranks***

| Source | DF | Type III SS | Mean Square | F Value | Pr>F |
|---|---|---|---|---|---|
| Student Ranks | 2 | 156.60170 | 78.30085 | 2.24 | 0.1070 |
| EOG Math 4 | 1 | 35678.52175 | 35678.52175 | 1019.37 | <.0001 |
| GrdRnk*GrdRnk*GrdRnk | 26 | 4775.90618 | 4775.90618 | 5.25 | <.0001 |
| Std*GrdR*GrdR*GrdRn | 52 | 2160.99694 | 41.55763 | 1.19 | 0.1698 |

For the results of the first research question to be statistically valid, the following assumptions have to be true at a minimum:

1. The residuals (the difference between the actual Algebra I scores and the fitted values, the EOG Math Grade 4 score and the teacher effects in grades 5-7 to estimate the same Algebra scores) should be normally distributed. The histograms of the residuals appear to have a bell-shaped curve and the results of the normality tests were all not significant which confirmed normality (see Appendix B).
2. The residuals should have a constant variance. This was checked by creating a plot of the residuals and the predicted values. The plot did not show a spreading out or compacting. The residuals were found to have a constant variance as evidenced by the plot or “blob” (see Appendix B).

**Research Question 2**

*What is the relationship between teacher effect data and teachers’ years of experience?*

In examining the relationship of years of experience and teacher effectiveness of teachers in grade five, there was a higher percentage of ineffective than effective teachers with 0-2 years of experience (see Table 19). The highest percentage of teachers fell in the middle 50% across years. In sum, there tended to be a higher percentage of teachers in Tertile 1 with fewer years of experience than in Tertiles Two and Three. Generally, as the experience level of the teachers increased, the percentage of effective teachers increased.

**Table 19**

***2005 Grade Five Teachers' Years of Experience by Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Years of Experience | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-02 Years | 6<br>31.58 | 10<br>52.63 | 3<br>15.79 | 19 |
| 03-05 Years | 4<br>36.36 | 6<br>54.55 | 1<br>9.09 | 11 |
| 06-10 Years | 4<br>23.53 | 8<br>47.06 | 5<br>29.41 | 17 |
| 11-20 Years | 5<br>17.24 | 18<br>62.07 | 6<br>20.69 | 29 |
| 21-99 Years | 7<br>18.42 | 19<br>50.00 | 12<br>31.58 | 38 |
| Unknown | 20<br>28.99 | 31<br>44.93 | 18<br>26.09 | 69 |
| Total | 46 | 92 | 45 | 183 |

As with fifth grade teachers, as the experience level of grade six teachers increased, the percentage of effective teachers rose (see Table 20). The 11-20 years of experience category posted the lowest percentage of ineffective teachers than in all other categories. Although small numbers of teachers in this category prevent us from drawing strong conclusions, half of the teachers with 21 or more years of experience were considered ineffective teachers as were half the teachers with 0-2 years of experience.

**Table 20**

***2005 Grade Six Teachers' Years of Experience by Teacher Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Years of Experience | Bottom Tertile<br>*Count*<br>*Percent* | Inter Tertile<br>*Count*<br>*Percent* | Top Tertile<br>*Count*<br>*Percent* | Total |
| 00-02 Years | 2<br>50.00 | 2<br>50.00 | 0<br>0.00 | 4 |
| 03-05 Years | 2<br>13.33 | 10<br>66.67 | 3<br>20.00 | 15 |
| 06-10 Years | 2<br>16.67 | 6<br>50.00 | 4<br>33.33 | 12 |
| 11-20 Years | 1<br>9.09 | 6<br>54.55 | 4<br>36.36 | 11 |
| 21-99 Years | 4<br>50.00 | 1<br>12.50 | 3<br>37.50 | 8 |
| Unknown | 11<br>26.83 | 21<br>51.22 | 9<br>21.95 | 41 |
| Total | 22 | 46 | 23 | 91 |

As presented in Table 21, 50% of the seventh grade math teachers with 0-2 years of experience fell in the bottom tertile compared to 0% in the 21-99 years category. In fact, 67% of the teachers with 21-99 years of experience were considered highly effective although small numbers of teachers in this category prevent us from drawing strong conclusions.

**Table 21**

***2005 Grade Seven Teachers' Years of Experience vs. Teacher Rankings***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Years of Experience | Bottom Tertile<br>*Count*<br>*Percent* | Inter Tertile<br>*Count*<br>*Percent* | Top Tertile<br>*Count*<br>*Percent* | Total |
| 00-02 Years | 9<br>50.00 | 6<br>33.33 | 3<br>16.67 | 18 |
| 03-05 Years | 3<br>33.33 | 4<br>44.44 | 2<br>22.22 | 9 |
| 06-10 Years | 0<br>0.00 | 5<br>100.00 | 0<br>0.00 | 5 |
| 11-20 Years | 1<br>25.00 | 3<br>75.00 | 0<br>0.00 | 4 |
| 21-99 Years | 0<br>0.00 | 2<br>33.33 | 4<br>66.67 | 6 |
| Unknown | 6<br>16.67 | 19<br>52.78 | 11<br>30.56 | 36 |
| Total | 19 | 39 | 20 | 78 |

Fifty-five percent of the eighth grade teachers with 6-10 years of experience were deemed highly effective. Of all categories, this category contained the highest percentage, of highly effective teachers (see Table 22).

As summarized in Table 23, there were a total of 427 teachers in grades 5-8 with teacher effectiveness estimates. When examining the data across all

**Table 22**

***2005 Grade Eight Teachers' Years of Experience by Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Years of Experience | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-02 Years | 2<br>22.22 | 5<br>55.56 | 2<br>22.22 | 9 |
| 03-05 Years | 3<br>25.00 | 6<br>50.00 | 3<br>25.00 | 12 |
| 06-10 Years | 2<br>18.18 | 3<br>27.27 | 6<br>54.55 | 11 |
| 11-20 Years | 2<br>22.22 | 3<br>33.33 | 4<br>44.44 | 9 |
| 21-99 Years | 2<br>25.00 | 3<br>37.50 | 3<br>37.50 | 8 |
| Unknown | 8<br>30.77 | 17<br>65.38 | 1<br>3.85 | 26 |
| Total | 19 | 37 | 19 | 75 |

grades instead of individual grades, the results are more definitive. Teachers with more years of experience tended to be Tertile 3 or the most effective teachers, while teachers with only a few years of experience tended to be Tertile 1 or the least effective teachers (see Figure 1). The greatest percentage of teachers across all years, tended to be average teachers. Generally, the more experience teachers gained, the more effective they became (?$^2$=10.5102; p=.0327).

**Table 23**

***2005 Grades Five-Eight Teachers' Years of Experience by Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Years of Experience | Bottom Tertile<br>*Count*<br>*Percent* | Inter Tertile<br>*Count*<br>*Percent* | Top Tertile<br>*Count*<br>*Percent* | Total |
| 00-02 Years | 19<br>38 | 23<br>46 | 8<br>16 | 50 |
| 03-05 Years | 12<br>25.53 | 26<br>55.32 | 9<br>19.15 | 47 |
| 06-10 Years | 8<br>17.78 | 22<br>48.89 | 15<br>33.33 | 45 |
| 11-20 Years | 9<br>16.98 | 30<br>56.60 | 14<br>26.42 | 53 |
| 21-99 Years | 13<br>21.67 | 25<br>41.67 | 22<br>36.67 | 60 |
| Unknown | 45<br>26.16 | 88<br>51.16 | 39<br>22.67 | 172 |
| Total | 106 | 214 | 107 | 427 |

When examining the extremes, the least experienced teachers tended to be less effective than the teachers with the most experience (see Figure 2).


**Figure 1.** ***2005 Teachers' Years of Experience by Teacher Effectiveness***


**Figure 2.** ***2005 Teachers' Years of Experience by Teacher Effectiveness***

## Research Question 3

*Based on teacher effectiveness estimates, how are teachers distributed throughout the district?*

The first analysis conducted examined the teacher effect estimates of teachers who taught fifth, sixth, seventh and eighth grade math in 2005 across tertiles. As shown in Table 24, of the 427 math teachers, ¼ of them were deemed ineffective, ½ average, and ¼ were considered highly effective. In addition, there were unknown years of experience due to the inability of the Guilford County's Human Resources Department to match teacher identifiers generated from SAS Institute.

**Table 24**

***2005 Fifth, Sixth, Seventh, and Eighth Grade Math Teacher Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| Subject/Grade | Bottom Tertile<br>*Count*<br>*Percent* | Inter Tertile<br>*Count*<br>*Percent* | Top Tertile<br>*Count*<br>*Percent* | Total |
| EOG Math 05 | 46<br>25.14 | 92<br>50.27 | 45<br>24.59 | 183 |
| EOG Math 06 | 22<br>24.18 | 46<br>50.55 | 23<br>25.27 | 91 |
| EOG Math 07 | 19<br>24.36 | 39<br>50.00 | 20<br>25.64 | 78 |
| EOG Math 08 | 19<br>25.33 | 37<br>49.33 | 19<br>25.33 | 75 |
| Total | 106 | 214 | 107 | 427 |

Table 25 displays the 2005 school assignments or distribution of fifth grade math teachers based on the percentage of students who received free/reduced price lunch (FRPL) and the teachers' effectiveness. In the 0-59% FRPL category or low poverty category, 20% of the teachers were deemed ineffective and 80% mid to highly effective. In contrast, in the high poverty or 60-99% schools, 31% of the teachers were considered ineffective and 69% mid to highly effective. In essence, a greater percentage of Tertile 1 teachers and lesser percentage of Tertile 3 teachers taught in high poverty schools than in low poverty schools. When examining the odds ratio of having a less effective grade five teacher in a high poverty school than in a low poverty school, one finds that a high poverty school was 1.879 times more likely to get an ineffective teacher than a low poverty school.

Grade six results are presented in Table 26. As the FRPL percentage increased, the number of Tertile 1 teachers increased with one exception; the percentage of Tertile 1 teachers was lower in 80-99% FRPL schools than in 60-79 percent schools. Correspondingly, there was a higher percentage of highly effective teachers in the 80-99 percent schools than in the 60-79 percent schools. A lower percentage of ineffective teachers were assigned to low poverty schools (0-59% FRPL) than in high poverty schools (60-99% FRPL), 21% compared to 28% respectively. A higher percentage of average and effective teachers were also assigned to low poverty as opposed to the high poverty schools; 79% average to highly effective teachers were in low poverty schools

**Table 25**

***2005 Grade Five Teacher Effectiveness and Free/Reduced Price Lunch Percentages***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| FRPL Percentage | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-19 Percent | 3<br>15.79 | 6<br>31.58 | 10<br>52.63 | 19 |
| 20-39 Percent | 10<br>20.83 | 24<br>50.00 | 14<br>29.17 | 48 |
| 40-59 Percent | 6<br>20.00 | 15<br>50.00 | 9<br>30.00 | 30 |
| 60-79 Percent | 9<br>29.03 | 18<br>58.06 | 4<br>12.90 | 31 |
| 80-99 Percent | 18<br>32.73 | 29<br>52.73 | 8<br>14.55 | 55 |
| Total | 46 | 92 | 45 | 183 |

compared to 72% in high poverty schools. Both low and high poverty schools had less than 30% of its teachers designated as highly effective, 29% and 21% respectively. The odds ratio indicated that ineffective grade six teachers were 2.08 times more likely to have taught in a high poverty school than a low poverty school.

**Table 26**

***Free/Reduced Price Lunch Percentages by 2005 Grade Six Teacher Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| FRPL Percentage | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-19 Percent | 1<br>10.00 | 7<br>70.00 | 2<br>20.00 | 10 |
| 20-39 Percent | 3<br>15.00 | 11<br>55.00 | 6<br>30.00 | 20 |
| 40-59 Percent | 7<br>31.82 | 8<br>36.36 | 7<br>31.82 | 22 |
| 60-79 Percent | 7<br>29.17 | 14<br>58.33 | 3<br>12.50 | 24 |
| 80-99 Percent | 4<br>26.67 | 6<br>40.00 | 5<br>33.33 | 15 |
| Total | 22 | 46 | 23 | 91 |

Table 27 displays grade seven teachers and their distribution across the county according to their effectiveness rankings. When examining patterns, one finds that 18% of teachers deemed ineffective were in low poverty schools (0-59% FRPL) compared to 33% in high poverty schools (60-99% FRPL). Seventy-eight percent of the teachers deemed mid level to highly effective were assigned to low poverty schools compared to 67% in the high poverty schools. A high poverty school was 2.17 times more likely to have had ineffective grade seven teachers than a low poverty school.

**Table 27**

***Free/Reduced Price Lunch Percentages by 2005 Grade Seven Teacher Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| FRPL Percentage | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-19 Percent | 0<br>0.00 | 6<br>75.00 | 2<br>25.00 | 10 |
| 20-39 Percent | 1<br>5.00 | 10<br>50.00 | 9<br>45.00 | 20 |
| 40-59 Percent | 8<br>40.00 | 9<br>45.00 | 3<br>15.00 | 20 |
| 60-79 Percent | 8<br>44.44 | 9<br>50.00 | 1<br>5.56 | 18 |
| 80-99 Percent | 2<br>16.67 | 5<br>41.67 | 5<br>41.67 | 12 |
| Total | 19 | 39 | 20 | 78 |

Table 28 displays the 2005 school assignments of eighth grade math teachers based on the percentage of students who received free/reduced price lunch (FRPL). In the low poverty schools (FRPL < 60%), 23% of the teachers were ineffective compared to 29% in high poverty schools. Seventy-seven percent of the teachers assigned to low poverty schools were considered average to highly effective compared to 71% in high poverty schools.

**Table 28**

***2005 Grade Eight Teacher Effectiveness and Free/Reduced Price Lunch Percentages***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| FRPL Percentage | Bottom Tertile *Count* *Percent* | Inter Tertile *Count* *Percent* | Top Tertile *Count* *Percent* | Total |
| 00-19 Percent | 2<br>28.57 | 1<br>14.29 | 4<br>57.14 | 7 |
| 20-39 Percent | 4<br>19.05 | 11<br>52.38 | 6<br>28.57 | 21 |
| 40-59 Percent | 5<br>26.32 | 8<br>42.11 | 6<br>31.58 | 19 |
| 60-79 Percent | 7<br>43.75 | 7<br>43.75 | 2<br>12.50 | 16 |
| 80-99 Percent | 1<br>8.33 | 10<br>83.33 | 1<br>8.33 | 12 |
| Total | 19 | 37 | 19 | 75 |

Low poverty schools housed 34% highly effective teachers compared to 11% in high poverty schools. Ineffective eighth grade teachers were 1.29 times more likely to have worked in high poverty schools than in low poverty schools.

Table 29 summarizes the distribution of teachers based on rankings across grades five through eight. Consistent with the specific grade level trends, there were fewer ineffective teachers teaching in low poverty schools than in high poverty schools. Conversely, there were a higher percentage of effective teachers teaching in low poverty schools than in low poverty schools. The greatest percentage of teachers in both categories fell in the average range. In sum, as students became poorer, they were more likely to experience less effective teachers in their schools (see Figures 3 and 4).

**Table 29**

***Free/Reduced Price Lunch Percentages by 2005 Grades Five through Eight Teacher Effectiveness***

| | Teacher Effectiveness | | | |
|---|---|---|---|---|
| FRPL Percentage | Bottom Tertile<br>*Count*<br>*Percent* | Inter Tertile<br>*Count*<br>*Percent* | Top Tertile<br>*Count*<br>*Percent* | Total |
| 00-19 Percent | 6<br>13.64 | 20<br>45.45 | 18<br>40.91 | 44 |
| 20-39 Percent | 18<br>16.51 | 56<br>51.38 | 35<br>32.11 | 109 |
| 40-59 Percent | 26<br>28.57 | 40<br>43.96 | 25<br>27.47 | 91 |
| 60-79 Percent | 31<br>34.83 | 48<br>53.93 | 10<br>11.24 | 89 |
| 80-99 Percent | 25<br>26.60 | 50<br>53.19 | 19<br>20.21 | 94 |
| Total | 106 | 214 | 107 | 427 |

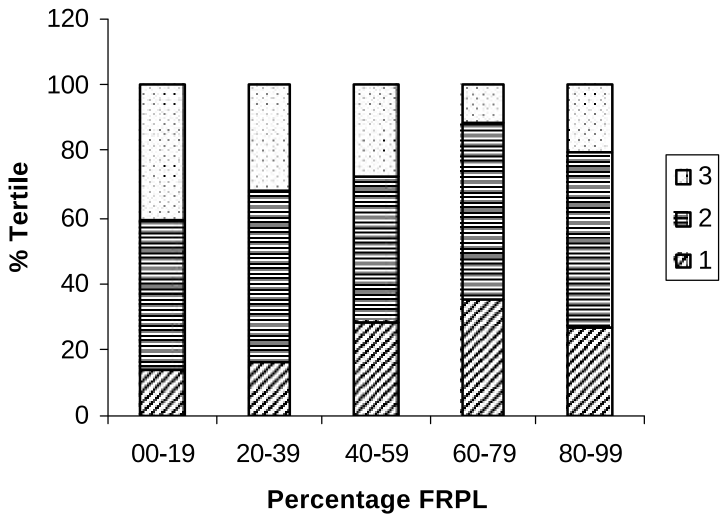



**Figure 3.** ***Free/Reduced Price Lunch Percentages by 2005 Teacher Effectiveness***

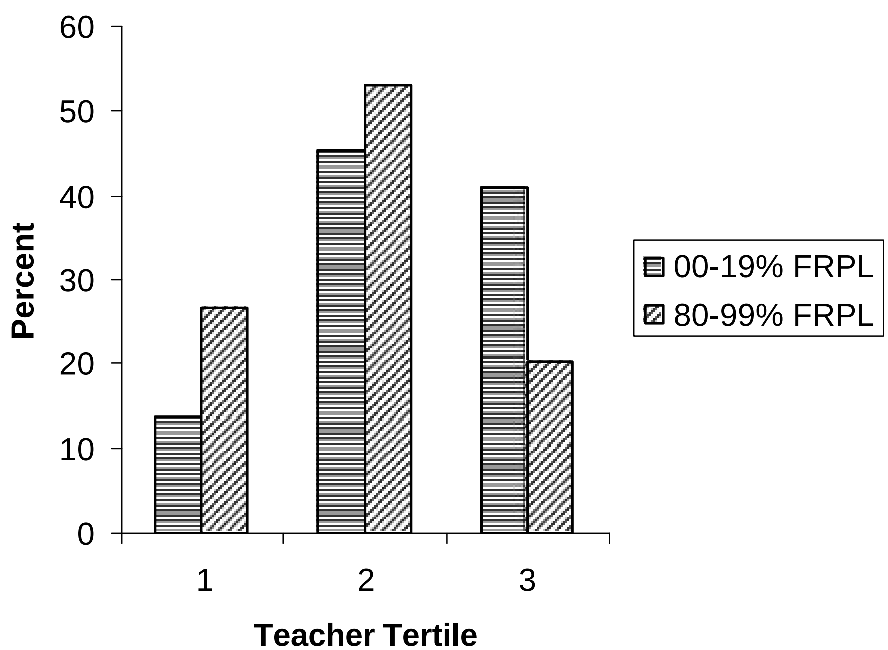



**Figure 4.** ***Free/Reduced Price Lunch by 2005 Teacher Effectiveness***

In examining extremes, the lowest poverty schools against the highest poverty schools, one finds a consistent theme: ineffective teachers were disproportionately assigned to high poverty schools as were effective teachers to low poverty schools. Further, when computing odds ratios using the data contained in Table 30, inexperienced teachers were 7.47 times more likely to have been assigned to 70-99% FRPL schools than 00-30%. Likewise, inexperienced teachers were 4.29 times more likely to have been assigned to 30-70% FRPL schools than in 00-30% schools.

**Table 30**

***2005 Free/Reduced Price Lunch Percentages and Teachers' Years of Experience***

| FRPL Percentage | Years 00-02 Years | of 03-05 Years | Experience 06-99 Years | Total |
|---|---|---|---|---|
| 00-30 Percent | 4<br>5.63 | 11<br>15.49 | 56<br>78.87 | 71 |
| 30-70 Percent | 21<br>20.39 | 22<br>21.36 | 60<br>58.25 | 103 |
| 70-99 Percent | 25<br>30.86 | 14<br>17.28 | 42<br>51.85 | 81 |
| Total | 50 | 47 | 158 | 255 |

The final analysis examined the relationship between schools according to their free/reduced price lunch percentages and teachers years of experience.

These data indicate that poorer schools were more likely to have less experienced teachers in them. As displayed in Table 30, in 0-30 percent FRPL schools, 6% of the teachers had 0-2 years of experience, 15% with 3-5 years of experience, and 79% with 6 or more years of experience. In 30-70 percent schools, 20% of the teachers had 0-2 years, 21% with 3-5 years and 58% with 6 or more years of experience. In 70-99 percent FRPL schools, 31% had teachers with 0-2 years of experience, 17% with 3-5 years and 52% with 6 or more years of experience.

# CHAPTER V

# FINDINGS AND DISCUSSION

The purpose of this final chapter is to provide a review of the entire research study with an emphasis on a discussion of the results as they relate to closing the achievement divide among students in schools. The discussion section includes a review of some of the achievement gap research and how this study informs this work. Further, recommendations or possible next steps are included for educators and policy makers.

## Summary of Research Problem

The purpose of this study was to determine the degree to which teacher assignments affect students' math performance as measured by their Algebra I End-of-Course (EOC) and Eighth Grade Math End-of-Grade (EOG) tests. The study provided estimates of the effect of a series of effective or ineffective teachers on the students' scores. The student data necessary for these analyses were acquired from Guilford County Schools, Greensboro, North Carolina and prepared for analyses by the SAS Institute, Cary, North Carolina. This preparation required the merging of year 2001 and 2005 achievement scores from the district with achievement and teacher effectiveness for the same students. A variety of descriptive analyses were conducted to demonstrate the relationship between the cumulative effects of teacher quality and student achievement as measured by students' performance on Eighth Grade Math End-of-Grade and Algebra I tests. The results of this study hinged on the accuracy of the value added assessment formula and the assumptions of the general linear model. While the system is not perfect, many researchers who have examined the system agree that the system is far superior over the use of simple raw averages to reach conclusions regarding district, school, and teacher effectiveness (Sanders, 2000; Wang et al., 1993).

Further, the purpose of the study was to determine the relationship between teacher effectiveness, years of experience and the distribution of teachers in low and high poverty schools.

**Findings**

Three research questions were explored in this study and a summary of the results of the explorations follows.

*I. What is the relationship between sequences of teacher effectiveness as measured by SAS Effectiveness Value Added Assessment System (EVAAS) or teacher effect scores and students' achievement as measured by their performance on the Algebra I and Eighth Grade Math End-of-Grade tests?*

The models examined the relationship between the effectiveness of fifth, fifth and sixth, and fifth, sixth and seventh grade teachers and the eighth grade Algebra I EOC and eighth grade Math EOG scores respectively. The findings using both achievement tests were similar; however, the EOG data were more robust due to 782 more matched scores, therefore, even more definitive than the Algebra I findings.

1. Even after adjusting for the entering achievement of the students in fourth grade, the impact of the previous fifth, sixth and seventh grade teachers, was quite meaningful on how eighth grade students performed on the Algebra I EOC and the EOG tests. These findings confirmed the importance of the influence that teachers have on students three years after they leave them.
2. When examining the interactions between grade level teachers, the order in which the students encountered the teachers did not have a significant bearing on the results. The findings suggest that there is no “catch-up” effect for students who had a “bad” teacher and then a “good” one. There is also no magnifying effect of having two consecutive effective or ineffective teachers above and beyond the fact that the teachers were effective or ineffective.
3. Another important finding indicated no interaction between the achievement levels of students and the effect of sequence of teachers; all children, low or high achieving, benefit greatly from having good teachers. Assuming that students were comparable across assignment of teachers, students who had one, two or three ineffective teachers in the bottom 25% scored lower than students who had no ineffective teachers. The effect size for having three ineffective Algebra I teachers

versus no ineffective teachers was small but meaningful. However, the effect size for three ineffective eighth grade math teachers was medium, close to large. In addition, when examining Algebra I and EOG percentiles or distribution of scores, the results are staggering (see Appendix C). In 2004-2005, in the state of North Carolina, the median Algebra 1 Scale Score was 62.96 or 63 which is at the $59^{th}$ percentile. Had the Algebra students not experienced 3 ineffective teachers, they could have scored in the $74^{th}$ percentile instead of the 59th. Similarly, had the eighth grade students who took the $8^{th}$ grade math EOG not had three ineffective teachers, they could have scored in the $79^{th}$ percentile instead of the $59^{th}$, only one point from the $81^{st}$ percentile.

II*. What is the relationship between teacher effect data and teachers' years of experience?*

1. Teachers who were more experienced and with known years of experience tended to be among the top 25% of the teachers in grades 5-8 in 2005, while teachers with only a few years of experience tended to be the least effective teachers. The greatest percentage of teachers across all years, tended to be average teachers. Generally, the more experience teachers gained, the more effective they became.

There were a number of teachers for which years of experience could not be ascertained due to incomplete teacher identifiers. To that end, trends should be interpreted with caution. Of the 427 teachers with available data, only 255 had 'known years' of experience. While it appeared that the 'unknown' teachers were distributed across the different tertiles, this large proportion of missing data, combined with a marginally significant p-value of 0.0327 may call into question any conclusions regarding statistical relationships.

III. *How are teachers that are rated high for adding value distributed throughout the district?*

1. When exploring trends in 2005 assignment of teachers in grades 5-8 based on FRPL percentages, one finds that the highest percentage of the most effective teachers were assigned to schools that were least impacted by poverty.
2. The highest percentage of teachers in both low and high poverty schools were mid-level or average teachers.
3. The trend was consistent across all grade levels, as students became poorer; they were more likely to experience less effective and less experienced teachers in their schools.

The FRPL status was a poverty indicator for the school and was not ascertained for each student taught by the teachers for which data were available. Analyses were made using the schools' FRPL status. As a result, the data in this study were not ideal for making cause and effect conclusions but more for identifying trends or patterns.

## Discussion and Implications

When examining the findings from each of the research questions, it is apparent that teachers do matter the most in this complex phenomenon called "educating America's children." Even after adjusting for prior achievement levels, the students were influenced positively or negatively by their teachers at least 3 years after they had them. In addition, experienced teachers tended to be more effective than novice teachers, yet the neediest schools, those most impacted by poverty were disproportionately assigned ineffective novice teachers. Even with this finding, the question that must be addressed by building level administrators: 'Are the most effective teachers in your building teaching the students who need them the most?' If the answer is no, then the question becomes 'why?' Given these findings, it is crucial for administrators in Guilford County or other districts that use teacher effectiveness estimates or value-added data to use the teacher data to help make decisions regarding class rosters or teaching assignments. As a result, decisions may not resonate with some teachers or the community; however, administrators are challenged to embrace an opportunity for which they have control over to increase student achievement and close the achievement divide: teacher assignments.

Eliminating the achievement divide is the most critical problem facing public schools in America today. According to Pollock (2001), although the achievement divide is well documented as a national crisis, as a local issue, the subject is not well publicized. The public, educators, social scientists, and researchers alike have neglected the issue; therefore, there has been limited research on effective practices and programs (Jencks & Phillips, 1998b). Elmore (2001) indicated that high-poverty, low-performing schools lacked the internal capacity for accountability as well as improvement strategies primarily because of lack of staff and district staff capacity. Ferguson (1998) argued that teachers' perceptions and expectations are paramount: "My bottom line conclusion is that teachers' perceptions, expectations, and behaviors probably do help to sustain, and perhaps even to expand, the black-white test score gap" (p. 313). Grissmer, Flanagan, Kawata, and Williamson (2000) posited that the achievement divide could effectively be addressed by providing resources to disadvantaged families and schools, lowering class size in the early grades, improving early childhood programs and improving teacher education and professional development. They also called for further research. Rothstein (2001) championed that significant progress toward closing the achievement gap could be made if strengthening families and communities, attending to health and nutrition needs, and improving family housing and income were a foci of public policy. Still others suggested that early childhood education as the most effective intervention (Ramey & Ramey, 1998; Thomas & Bainbridge, 2001). Jencks and Phillips (1998b) stated that "If we want equal outcomes among twelfth graders, we will also have to narrow the skill gap between black and white children before they enter school" (p. 46).

With all of the research and positions written regarding the achievement divide, there is very limited data to support or inform what happens to children once they enter the school house doors. Based on the results of this study, and other value-added studies such as the ones conducted in Dallas and Tennessee, for increased achievement to be realized by all, accountability is paramount and must be embraced by the district, school, principal, and most important the teacher. Student work and achievement must be tied to the work of the teacher. Using value-added data is a step towards bridging the accountability gap and the achievement divide.

In this study as in the study conducted by Sanders and Rivers (1996) and Rivers (1999), the teacher was a significant contributor to a student's achievement, especially in math. The impact of the teacher was quite significant and influenced student achievement outcomes three years after a student was assigned to a teacher. In addition, the sequence of teachers was a tremendous factor in a student's performance. Some students who had three ineffective teachers could have scored almost seven points higher on a math End-of-Grade test had they not had any ineffective teachers, going from the 59$^{th}$ to 79$^{th}$ percentile. Unfortunately, many of the ineffective teachers were disproportionately assigned to high-poverty schools. Conversely, many of the highly effective teachers (top 25%) were disproportionately assigned to high wealth schools in the district. Based on these findings, the researcher is convinced that the assignment of effective teachers in high poverty schools is a viable approach to reducing the achievement divide.

The deployment and retention of highly effective teachers as a solution for closing the achievement divide is a complex one. Many highly effective teachers choose to work in schools where there are fewer problems that are primarily related to socio-economics. Local school boards, including the Guilford County Board of Education, could choose to eliminate the existence of high-poverty schools and create middle-class schools (Kahlenberg, 2001). However, if high-poverty schools are here to stay, then local policy makers have the authority to deploy and retain highly effective teachers in high poverty schools.

Deployment and retention of highly effective teachers will require the implementation of several strategies. First, effective teachers must be identified; therefore, there has to be a fair and objective way to estimate the effects of teachers on the academic growth of students; SAS EVAAS or any value-added data system will fulfill that requirement. While the use of value-added data is not perfect, it is a great deal better than the use of raw averages or subjective pencil-paper evaluations by superiors. Teachers must be recruited, hired or retained based on their proven record of helping students learn. More experienced teachers were generally found to be more effective than novice teachers. Second, credible data on the results of teaching efforts must be provided to teachers. Schmoker (1999) noted, “Data and results can be a powerful force for generating an intrinsic desire to improve” (p. 70). Without tangible feedback on the results of their work, teachers can hardly hope to improve. Third, teachers must be deployed to work in school communities where the fundamental purpose of school is learning and not teaching. Teachers must be equipped to administer, interpret and act on formative assessments to influence student learning. Teachers must work with a team of professionals committed to supporting each other and improving their craft to increase the achievement of their students. When teachers learn, students learn. In short, more effective teachers must spring forth from the ranks. This is an extremely important point given that a preponderance of the teaching force in this study was considered mid-level teachers. Fourth, the teacher results should be used to make teaching assignments within a school. The use of teacher effectiveness data could change the entire culture of a school: elementary school children may no longer be taught by their same homeroom teacher all day in every subject. The students should have teachers who have demonstrated strengths in particular disciplines teach them regardless of their homeroom assignments. The strongest teachers must teach the children who need them the most. From this study, we found that all students, regardless of their achievement level benefit from having an effective teacher. Finally, teachers and administrators must be adequately compensated for their performance or the quality or difficulty of their work. The face of the teaching job market must change and become competitive, attracting the best and most committed persons into the classrooms. Teaching in a high poverty school is very different from teaching in a low poverty or high wealth school. I speak from experience; I've taught and administered in both settings. The work in a high poverty school is intense and constant. The student to student and student to teacher interactions are not always ideal and tend to cause one's stress level to rise. Yet the work is rewarding and absolutely necessary. Effective educators must be compensated for choosing to devote their lives to working in environments that are not always commensurate with the traditional ways of "doing school."

In sum, the deployment and retention of effective teachers to high-poverty schools are two means of addressing the achievement divide based on the influence and residual effects teachers have on student learning as evidenced by the results of this study. School systems and law makers across the nation must be poised to increase the effectiveness of its teaching workforce and deploy more effective teachers to high-poverty schools. The implementation of these strategies hinges on the use of measurement methodologies that can objectively estimate the effects of schools and teachers on the academic growth of students.

## Educational Research

Further research is appropriate in three areas. First, teacher effectiveness studies using value-added data in other curriculum areas other than mathematics should be conducted. Will similar findings spring forth using reading, science or history achievement scores and teacher effectiveness estimates? Second, while there has been a great deal of research on the effective practices of teachers, there continues to be limited research on the use of value-added data and the distribution of effective teachers. A more statistically rigorous study should be conducted that explores the correlation of specific teacher effectiveness scores with individual student FRPL status. Lastly, a study on the impact of a new recruitment initiative recently launched by Guilford County Schools, Greensboro, North Carolina on student learning is warranted in two to three years. The initiative known as *Mission Possible* involves the use of value-added data to identify and recruit highly effective teachers within the district to teach in high-poverty schools. The teachers will be monetarily compensated for choosing to teach at the identified schools and getting positive value-added results at the close of each year.  In addition, a longitudinal study is warranted to investigate if the teachers who were deemed highly effective at the start of this initiative remain in the highly effective category after years of working in a school heavily impacted by poverty. This researcher applauds this school system's efforts to seek learning for its children with ardor and diligence.

**BIBLIOGRAPHY**

Andrews, J. W., Blackmon, C. R., & Mackey, J.A. (1980). Preservice performance and the national teacher examinations. *Phi Delta Kappan*, 61, 358-359.

Ashton, P. & Crocker, L. (1987). Systematic study of planned variations: The essential focus of teacher education reform. *Journal of Teacher Education*, 38, 2-8.

Babu, S. & Mendro, R. (2003). *Teacher accountability: HLM-based teacher effectiveness indices in the investigation of teacher effects on student achievement in a state assessment program*. Paper presented at the meeting of the American Education Research Association.

Barton, P. (2005). Why does the gap persist? *Education Leadership, 62,* 9-13.

Begle, E.G. (1979). *Critical variables in mathematics education*. Presented at meeting of the Mathematical Association of American and National Council of Teachers of Mathematics, Washington, DC.

Boehner, J. (2001). *News from the committee work force.* Washington DC.

Bock, D., & Wolfe, R. (1996). A review and analysis of Tennessee value-added assessment system. Knoxville, TN: Tennessee Comptroller of the Treasury.

Byrne, C. J. (1983). *Teacher knowledge and teacher effectiveness: A literature review, theoretical analysis and discussion of research strategy*. Paper presented at the meeting of the Northwestern Educational Research Association, Ellenville, NY.

Carey, K. (2004). *The real value of teachers.* Washington, DC: The Education Trust.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.

Comer, J. (2001). Closing the gap. Retrieved September, 25, 2005 from Fafrican.com. Koret Communications Limited. Microsoft Corporation.

Darling-Hammond, L. (1997). *Doing what matters most: Investing in quality teaching*. Paper presented at the meeting of the New York: National Commission on Teaching and America's Future, New York.

Darling-Hammond, L., Wise, F. & Klein, S. (1995). A License to teach: Building a profession for 21st century schools. Boulder: Westview Press.

Darling-Hammond, L. (2001). The challenge of staffing our schools. *Educational Leadership*, 12-17.

Darlington, R. (1997). Grading teachers, grading schools. In J. Millman (Ed.), *The Tennessee value-added assessment system: A challenge to familiar assessment* (pp. 163-181). Corwin Press, CA: Thousand Oaks.

Denton, J. J., & Lacina, L. J. (1984). Quantity for professional education coursework linked with process measures of student teaching. *Teacher Education and Practice*, 39-64.

Donlevy, J. (2002). No child left behind: In search of equity for all children. *International Journal of Instructional Media*, 29, 257-259.

Elmore, R. (2001). *Psychiatrists and lightbulbs: Educational accountability and the problem of capacity*. Paper presented at the annual meeting of the American Educational Research association, Seattle, WA.

Ferguson, R.F. (1991). Paying for public education: New evidence on how and why money matters. *Harvard Journal on Legislation,* 28, 465-498.

Ferguson, R. F. (1997). Evidence that schools can narrow the black-white test score gap*. Harvard Journal on Legislation*, 32.

Ferguson, R. F. (1998). Teacher perceptions and expectations and the black-white test score gap. In C. Jencks & M. Phillips, *The black-white test score gap* (pp.273-317). Washington, DC: Brookings Institution.

Glass, G.V; Cayhen, L.S.; Smith, M.L.; & Filby, N.N. (1982). *School class size: Research and policy.* Beverly Hills, CA: SAGE Publications.

Gould, S. J. (1981). *The mismeasure of man.* New York: W. W. Norton.

Grissmer, D., Flanagan, A., Kawata, F., & Williamson, S. (2000). *Improving student achievement: What state NAEP scores tell us*. Santa Monica, CA: RAND.

Haney, W., Madaus, G., & Kreitzer, A. (1987). Charms talismanic: Testing teachers education for the improvement of American education. *Review of Research in Education,14.*

Hanushek, E. (1986). The economics of schooling: Production and efficiency in the public schools. *The Journal of Economic Literature, 24*(3) 1141-77.

Haycock, K. (1997). Education Watch: The 1996 education state and national data book. Washington DC: The Education Trust.

Haycock, K. (1998). Good teaching matters. Washington, DC: The Education Trust.

Haycock (2001). Closing the achievement gap. Washington, DC: The Education Trust.

Herrstein, R., & Murray, C. (1994). The bell curve: Intelligence and class structure in American life. New York: Free Press.

Hershberg, T. (2005). Value-added data and systemic reform: a response to the challenge of human capital development*. Phi Delta Kappan*, *87,* 276-283.

Hill, R., Rowan, B., & Ball, D. (2005). Effects of teachers' mathematical knowledge on student achievement. *American Educational Research Journal*, *42,* 371-406.

Howard, J. (1995). You can't get there from here: The need for a new logic in education reform. *Journal of the American Arts & Sciences*, *124.*

Howell, D. (2004). *Fundamental statistics for the behavioral sciences (5th ed.).* California: Brooks/Cole.

Jencks, C. & Phillips, M. (1998a). America's next achievement test: Closing the black, white achievement gap. *The American Prospect, 40,* 44-53.

Jencks, C., & Phillips, M. (1998b). *The black-white test score gap*. Washington, DC: Brookings Institution.

Jepsen, C., & Rivkin, S. (2002). Class size reduction, teacher quality and academic achievement in California public elementary schools. California: Public Policy Institute.

Jerald, C. & Ingersoll, R. (2002). All talk, no action: putting an end to out of field teaching. Washington, DC: The Education Trust.

Jordan, H. R., Mendro, R. L., & Weersinghe, D. (1997). *Teacher Effects on longitudinal student achievement: A preliminary report on research on teacher effectiveness.* Paper Presented at the National Evaluation Institute, Indianapolis, IN.

Kahlenberg, R. (2001). *All tgether now: Creating middle-class schools through public school choice plans*. Washington, DC: Brookings Institution.

Kain, J., & Singleton, C. (1996). Equality of educational opportunity revisited. *New England Economic Review*, 109.

Lankford, H., Wyckoff, J., & Papa (2000). *The labor market for public school teachers: a descriptive analysis of New York state's teacher workforce*. Albany, NY: University at Albany.

Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. Educational Evaluation and Policy Analysis.

Monk, D., & King J. (1994). Multilevel teacher resource effects on pupil performance in secondary mathematics and science. In R. G. Ehrenberg (Ed.), *Choices and Consequence*. Ithaca NY: ILR Press.

National Assessments Education Progress. (2000). Retrieved October 15, 2005 from http://nces.ed.gov/nationsreportcard/reading/results2003/raceethnicity.asp

National Center for Education Statistics 2000. Monitoring quality: an indicators report. Washington, DC: U. S. Department of Education.

National Center for Educational Statistics. (2001). Washington DC: U. S. Department of Education.

North Carolina Department of Education. (2005). Retrieved October 3, 2005 from http://www.ncpublicschools.org/.

Pollock, M. (2001). How the question we ask most about race in education is the very question we most suppress. *Educational Researcher, 30,* 2-12.

Ramey, C., & Ramey, S. (1998). Early intervention and early experience. *American Psychologist, 53,* 109-120.

Rivers, J. C. (1999). *The impact of teacher effect on student math competency achievement.* Unpublished doctoral dissertation, University of Tennessee, Knoxville.

Rivkin, S., Hanushek, E., & Kain, J. (2002). Teachers, schools and academic achievement. Dallas, TX: University of Texas-Dallas Schools Project.

Rothstein, R. (2001). Finance fungibility: Investigating relative impacts of investments in schools and non-school educational institutions to improve student achievement. Retrieved September 15, 2004 from http//www.financeproject.org/achievement.htm.

Sanders, W. (2000). Value-Added assessment from Student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14,* 329-339.

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee value-added assessment system (TVAAS) database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education*, *12,* 247-256.

Sanders, W., & Rivers, J. (1996). Cumulative and residual effects of teachers on future academic achievement. *Research Progress Report*. University of Tennessee Value Added Research and Assessment Center.

Sanders, W., & Wright, P. (2004). Controlling for student background in value added assessment of teachers. *Journal of Educational and Behavioral Statistics*, *29,* 37-65.

Sanders, W. L. (1998). Value-added assessment. *School Administrator, 55,* 24.

Sanders, W. L., Saxton, A. M., & Horn, S. P. (1997). The Tennessee value-added system: A quantitative outcomes-based approach to educational assessment. *Grading teachers, grading schools: Is student achievement a valid measure?* Thousand Oaks, CA: Corwin Press.

Sanders, W. L., & Topping, K. J. (1999). *Teacher effectiveness and computer assessment of reading: Relating value-added and learning information system data.* Knoxville, TN: University of Tennessee Value Added Research and Assessment Center.

SAS Institute. (2006). 9.1 User's Guide. Retrieved April 10, 2006 from http://support.sas.com/documentation/onlinedoc/91pdf/sasdoc91/stat_ug_7313.pdf.

Schmoker, M. (1999). *Results: The key to continuous improvement.* Alexandria, VA: Association for Supervision & Curriculum Development.

Shen, J. (2003). The distribution of quality of new teaching force: results from the baccalaureate and beyond longitudinal study, 1993-97. Washington, DC: The Education Trust.

Strauss, R. & Sawyer, E. (1986). Some new evidences on teacher and student competencies. *Economics of Educational Review*, *5,* 41-48.

Summers, A. A., & Wolfe, B. L. (1975). Do schools make a difference? *American Economic Review, 67,* 639-652.

Thomas, M., & Bainbridge, W. (2001). All children can learn: facts and fallacies. *Education Research Service Spectrum*, 1-4.

U. S. Department of Education. (2003). *Meeting the highly qualified teacher challenge: The secretary's second annual report on teacher quality.* Washington, DC.

U. S. Department of Education. (2005). *Stronger accountability: How No Child Left Behind benefits African Americans.* [Online]. Retrieved October 2005 from http://www.ed.gov/nclb/accountability/achieve/nclb-aa.html.

Wang, M. C., Haertel, G., & Walberg, H. J. (1993). Toward a knowledge base of school learning. *Review of Educational Research, 73,* 249-294.

Wayne, A. (2002). Teacher inequality: new evidence on teachers' academic skills. SRI International.

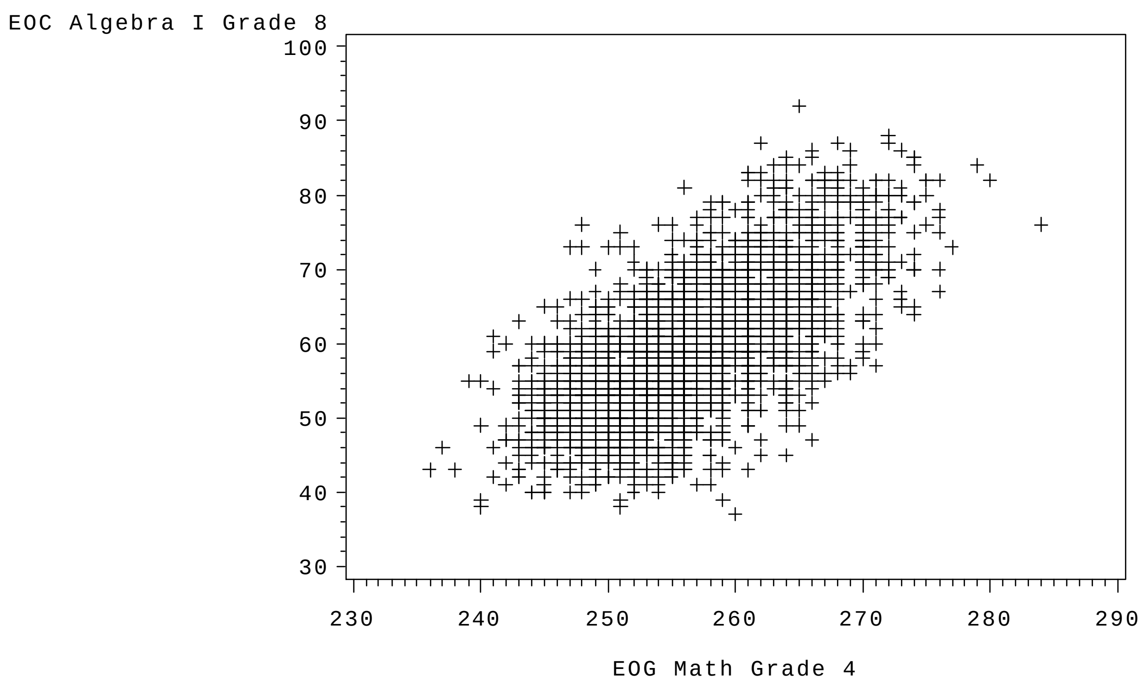
Wiersma, W., & Jurs, S. (2005). *Research methods in education* ($8^{th}$ ed.). New York: Allyn & Bacon.

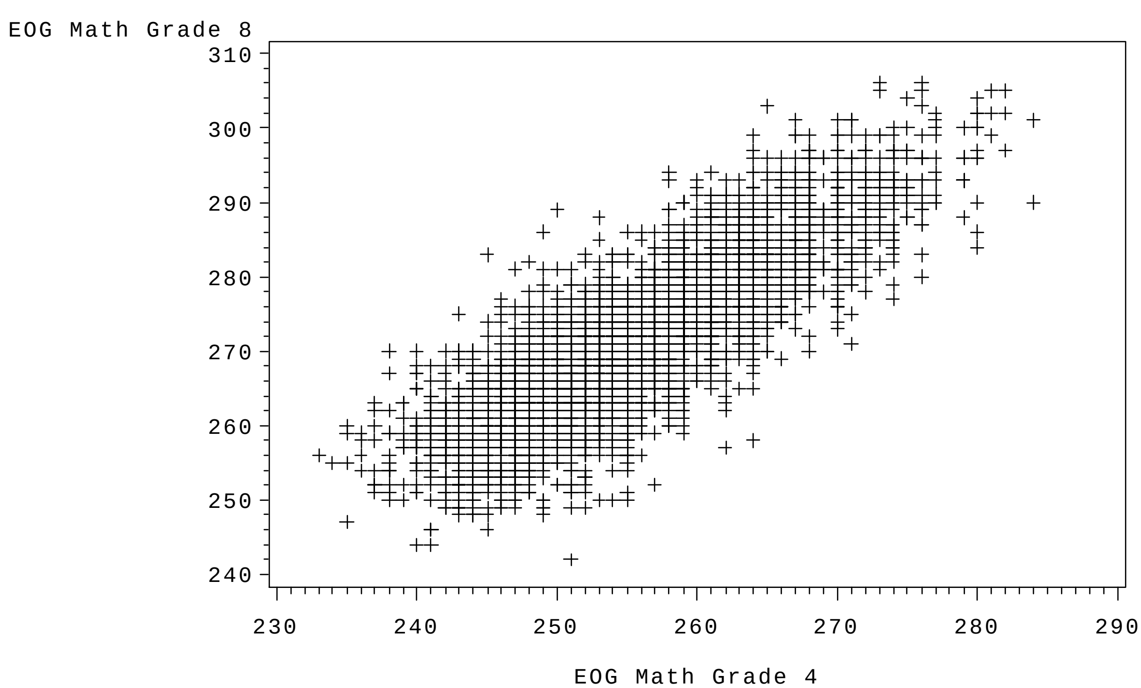
Wright, S. P., Horn, S. R., & Sanders, W. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation*, *11,* 57-67.

Wright, P. (2004, July). *Advantages of a multivariate longitudinal approach to educational value-added assessment without imputation*. Paper presented at the National Evaluation Institute, Colorado Springs, CO.

# APPENDIX A

# SCATTER PLOTS


EOC Algebra I Grade 8
100
90
80
70
60
50
40
30
230
240
250
260
270
280
290
EOG Math Grade 4


EOG Math Grade 8
310
300
290
280
270
260
250
240
230
240
250
260
270
280
290
EOG Math Grade 4

# APPENDIX B

# NORMALITY DETERMINATION

Univariate Procedure

Variable: Residual Algebra I Grade 8

| Moments | | | |
|---|---|---|---|
| **N** | 2119 | **Sum Weights** | 2119 |
| **Mean** | 0 | **Sum Observations** | 0 |
| **Std Deviation** | 6.85018507 | **Variance** | 46.9250355 |
| **Skewness** | 0.06885596 | **Kurtosis** | 0.2032101 |
| **Uncorrected SS** | 99387.2252 | **Corrected SS** | 99387.2252 |
| **Coeff Variation** | | **Std Error Mean** | 0.14881161 |

| Basic Statistical Measures | | | |
|---|---|---|---|
| **Location** | | **Variability** | |
| **Mean** | 0.00000 | **Std Deviation** | 6.85019 |
| **Median** | -0.12531 | **Variance** | 46.92504 |
| **Mode** | -4.39010 | **Range** | 48.71926 |
| | | **Interquartile Range** | 8.91631 |

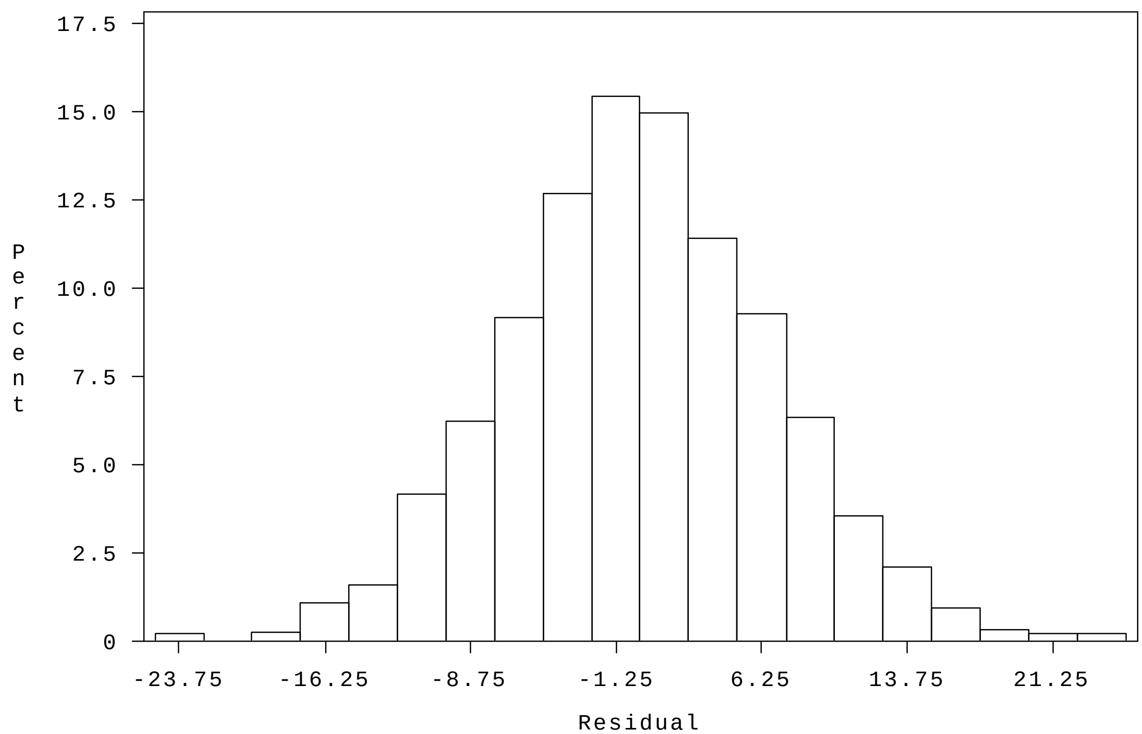
**Note:** The mode displayed is the smallest of 3 modes with a count of 2.

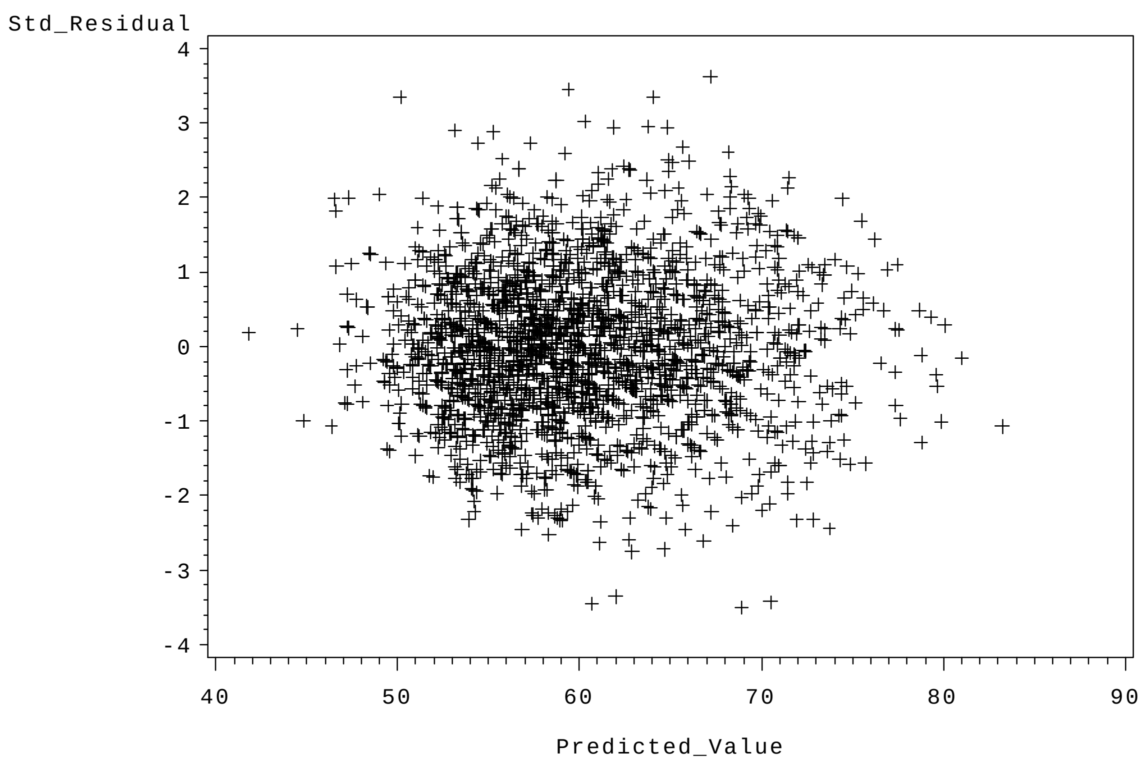
| Tests for Location: Mu0=0 | | | | |
|---|---|---|---|---|
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 0 | **Pr > \|t\|** | 1.0000 |
| **Sign** | **M** | -15.5 | **Pr >= \|M\|** | 0.5146 |
| **Signed Rank** | **S** | -8362 | **Pr >= \|S\|** | 0.7667 |

| Tests for Normality | | | | |
|---|---|---|---|---|
| **Test** | **Statistic** | | **p Value** | |
| **Kolmogorov-Smirnov** | **D** | 0.018778 | **Pr > D** | 0.0700 |
| **Cramer-von Mises** | **W-Sq** | 0.099641 | **Pr > W-Sq** | 0.1167 |
| **Anderson-Darling** | **A-Sq** | 0.533427 | **Pr > A-Sq** | 0.1801 |

| Quantiles (Definition 5) | |
|---|---|
| **Quantile** | **Estimate** |
| **100% Max** | 24.780629 |
| **99%** | 16.311051 |
| **95%** | 11.307495 |
| **90%** | 8.773021 |
| **75% Q3** | 4.470255 |
| **50% Median** | -0.125315 |
| **25% Q1** | -4.446050 |
| **10%** | -8.737113 |
| **5%** | -11.328107 |
| **1%** | -15.765676 |
| **0% Min** | -23.938629 |

| Extreme Observations | | | |
|---|---|---|---|
| **Lowest** | | **Highest** | |
| | | | |
| **Value** | **Obs** | **Value** | **Obs** |
| -23.9386 | 1821 | 20.7240 | 2062 |
| -23.6722 | 1738 | 22.8647 | 1369 |
| -23.4930 | 327 | 22.9350 | 795 |
| -22.9947 | 925 | 23.6084 | 1975 |
| -18.8808 | 458 | 24.7806 | 1834 |


Percent
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
-23.75
-16.25
-8.75
-1.25
6.25
13.75
21.25
Residual


Std_Residual
4
3
2
1
0
-1
-2
-3
-4
40
50
60
70
80
90
Predicted_Value

Univariate Procedure

Variable: Residual Math EOG Grade 8

| **Moments** | | | |
|---|---|---|---|
| **N** | 2901 | **Sum Weights** | 2901 |
| **Mean** | 0 | **Sum Observations** | 0 |
| **Std Deviation** | 5.85150927 | **Variance** | 34.2401608 |
| **Skewness** | 0.04521077 | **Kurtosis** | 0.28032159 |
| **Uncorrected SS** | 99296.4662 | **Corrected SS** | 99296.4662 |
| **Coeff Variation** | | **Std Error Mean** | 0.10864107 |

| **Basic Statistical Measures** | | | |
|---|---|---|---|
| **Location** | | **Variability** | |
| **Mean** | 0.00000 | **Std Deviation** | 5.85151 |
| **Median** | -0.03555 | **Variance** | 34.24016 |
| **Mode** | -4.93263 | **Range** | 45.07064 |
| | | **Interquartile Range** | 7.70594 |

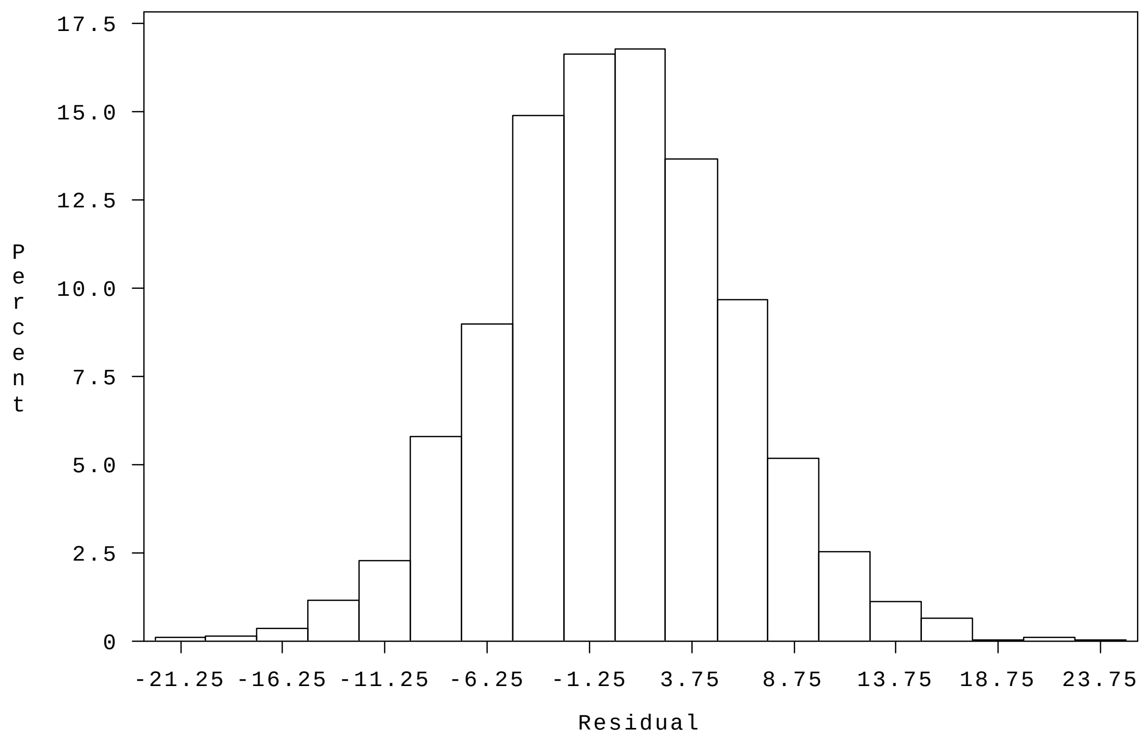
**Note:** The mode displayed is the smallest of 6 modes with a count of 2.

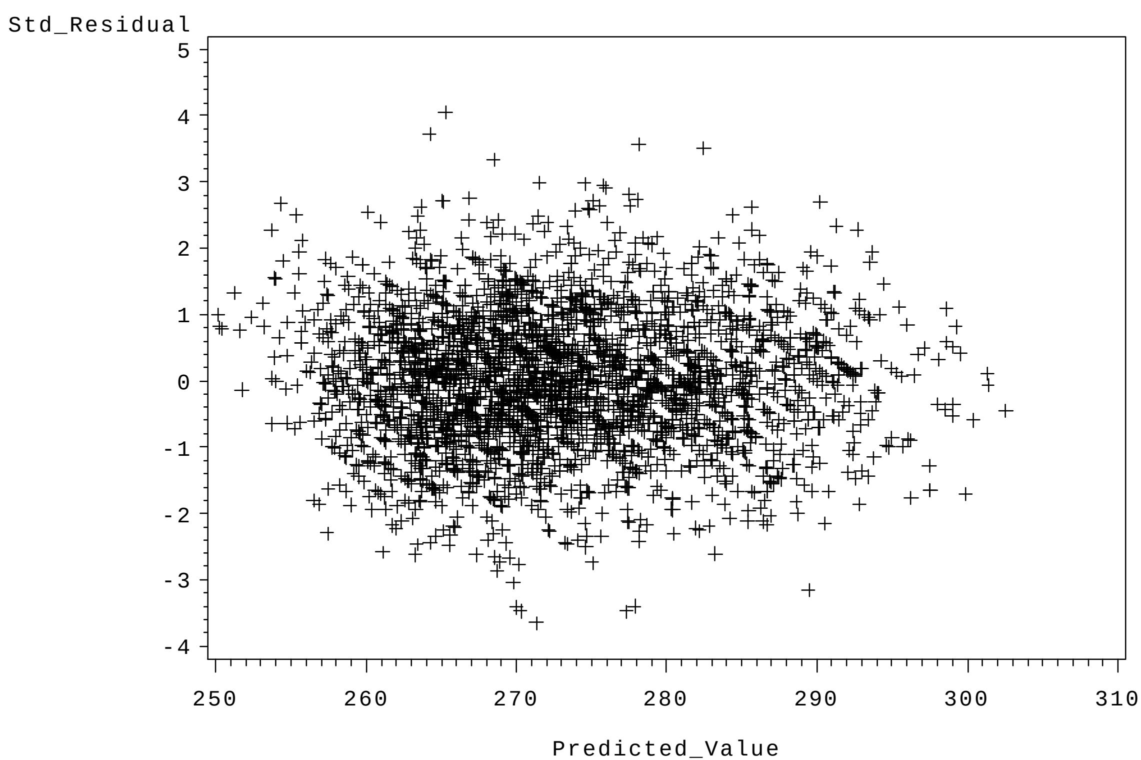
| **Tests for Location: Mu0=0** | | | | |
|---|---|---|---|---|
| **Test** | **Statistic** | | **p Value** | |
| **Student's t** | **t** | 0 | **Pr > \|t\|** | 1.0000 |
| **Sign** | **M** | -8.5 | **Pr >= \|M\|** | 0.7664 |
| **Signed Rank** | **S** | -8482.5 | **Pr >= \|S\|** | 0.8509 |

| Tests for Normality | | | | |
|---|---|---|---|---|
| **Test** | **Statistic** | | **p Value** | |
| **Kolmogorov-Smirnov** | **D** | 0.012836 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.074252 | **Pr > W-Sq** | 0.2484 |
| **Anderson-Darling** | **A-Sq** | 0.494248 | **Pr > A-Sq** | 0.2232 |

| Quantiles (Definition 5) | |
|---|---|
| **Quantile** | **Estimate** |
| **100% Max** | 23.7100312 |
| **99%** | 14.2007821 |
| **95%** | 9.5459311 |
| **90%** | 7.3907728 |
| **75% Q3** | 3.8549551 |
| **50% Median** | -0.0355475 |
| **25% Q1** | -3.8509875 |
| **10%** | -7.4460218 |
| **5%** | -9.6072664 |
| **1%** | -13.6158769 |
| **0% Min** | -21.3606038 |

| Extreme Observations | | | |
|---|---|---|---|
| **Lowest** | | **Highest** | |
| **Value** | **Obs** | **Value** | **Obs** |
| -21.3606 | 1191 | 19.5050 | 44 |
| -20.3332 | 1668 | 20.5548 | 73 |
| -20.2833 | 1683 | 20.8717 | 2170 |
| -19.9995 | 1437 | 21.7694 | 675 |
| -19.8569 | 1770 | 23.7100 | 415 |


17.5
15.0
12.5
10.0
7.5
5.0
2.5
0
Percent
-21.25 -16.25 -11.25 -6.25 -1.25 3.75 8.75 13.75 18.75 23.75
Residual


Std_Residual
5
4
3
2
1
0
-1
-2
-3
-4
250 260 270 280 290 300 310
Predicted_Value

# APPENDIX C

# DISTRIBUTION OF SCALE SCORES

## Table 33. 2004-05 End-of-Course Distribution of Scale Scores
## Algebra I

| | | | |
|---|---|---|---|
| NUMBER OF STUDENTS WITH VALID SCORES | 110,866 | HIGH SCORE | 96 |
| | | LOW SCORE | 31 |
| MEAN | 63.1 | | |
| STANDARD DEVIATION | 10.1 | | |
| VARIANCE | 101.2 | | |

| 2005 STATE PERCENTILES | SCALE SCORE |
|---|---|
| 90 | 76.61 |
| 75 | 69.94 |
| 50 (MEDIAN) | 62.96 |
| 25 | 56.09 |
| 10 | 50.22 |

FREQUENCY DISTRIBUTION

| SCALE SCORE | FREQUENCY | CUMULATIVE FREQUENCY | PERCENT | CUMULATIVE PERCENT | 2001 STATE PERCENTILE |
|---|---|---|---|---|---|
| 96 | 34 | 110866 | 0.03 | 100.00 | 99 |
| 95 | 37 | 110832 | 0.03 | 99.97 | 99 |
| 94 | 78 | 110795 | 0.07 | 99.94 | 99 |
| 93 | 74 | 110717 | 0.07 | 99.87 | 99 |
| 92 | 63 | 110643 | 0.06 | 99.80 | 99 |
| 91 | 151 | 110580 | 0.14 | 99.74 | 99 |
| 90 | 94 | 110429 | 0.08 | 99.61 | 99 |
| 89 | 205 | 110335 | 0.18 | 99.52 | 99 |
| 88 | 285 | 110130 | 0.26 | 99.34 | 99 |
| 87 | 169 | 109845 | 0.15 | 99.08 | 99 |
| 86 | 569 | 109676 | 0.51 | 98.93 | 99 |
| 85 | 454 | 109107 | 0.41 | 98.41 | 99 |
| 84 | 504 | 108653 | 0.45 | 98.00 | 99 |
| 83 | 552 | 108149 | 0.50 | 97.55 | 99 |
| 82 | 950 | 107597 | 0.86 | 97.05 | 98 |
| 81 | 1060 | 106647 | 0.96 | 96.19 | 98 |
| 80 | 1141 | 105587 | 1.03 | 95.24 | 97 |
| 79 | 1208 | 104446 | 1.09 | 94.21 | 97 |
| 78 | 1371 | 103238 | 1.24 | 93.12 | 96 |
| 77 | 2345 | 101867 | 2.12 | 91.88 | 95 |
| 76 | 1484 | 99522 | 1.34 | 89.77 | 94 |
| 75 | 2214 | 98038 | 2.00 | 88.43 | 93 |
| 74 | 2299 | 95824 | 2.07 | 86.43 | 91 |
| 73 | 2316 | 93525 | 2.09 | 84.36 | 90 |
| 72 | 3144 | 91209 | 2.84 | 82.27 | 88 |
| 71 | 2546 | 88065 | 2.30 | 79.43 | 86 |
| 70 | 4207 | 85519 | 3.79 | 77.14 | 83 |
| 69 | 3678 | 81312 | 3.32 | 73.34 | 80 |
| 68 | 2869 | 77634 | 2.59 | 70.03 | 77 |
| 67 | 4574 | 74765 | 4.13 | 67.44 | 74 |
| 66 | 3894 | 70191 | 3.51 | 63.31 | 70 |
| 65 | 3170 | 66297 | 2.86 | 59.80 | 67 |
| 64 | 4976 | 63127 | 4.49 | 56.94 | 63 |
| 63 | 5010 | 58151 | 4.52 | 52.45 | 59 |
| 62 | 3424 | 53141 | 3.09 | 47.93 | 55 |
| 61 | 4174 | 49717 | 3.76 | 44.84 | 51 |
| 60 | 4943 | 45543 | 4.46 | 41.08 | 46 |
| 59 | 4141 | 40600 | 3.74 | 36.62 | 42 |
| 58 | 3098 | 36459 | 2.79 | 32.89 | 38 |
| 57 | 4710 | 33361 | 4.25 | 30.09 | 34 |
| 56 | 2290 | 28651 | 2.07 | 25.84 | 30 |
| 55 | 4507 | 26361 | 4.07 | 23.78 | 26 |
| 54 | 2152 | 21854 | 1.94 | 19.71 | 23 |
| 53 | 2860 | 19702 | 2.58 | 17.77 | 20 |
| 52 | 2701 | 16842 | 2.44 | 15.19 | 17 |
| 51 | 2565 | 14141 | 2.31 | 12.76 | 14 |
| 50 | 1759 | 11576 | 1.59 | 10.44 | 12 |
| 49 | 1646 | 9817 | 1.48 | 8.85 | 10 |
| 48 | 1562 | 8171 | 1.41 | 7.37 | 8 |
| 47 | 1410 | 6609 | 1.27 | 5.96 | 6 |
| 46 | 1269 | 5199 | 1.14 | 4.69 | 5 |

Notes: Data received from LEAs and charter schools after July 22, 2005 are not included in this table.
Prepared by the NCDPI Division of Accountability Services/Test Development Section.