

BECOATS, JOCELYN B., Ed.D. Determining the Correlation of Effective Middle School Math Teachers and Math Student Achievement. (2009)  
Directed by Dr. Carl Lashley. 118 pp.

This study determined whether there was a significant relationship between teacher effect data in middle school mathematics and a teachers years of experience and whether there was a significant relationship between an effective teacher as measured by the total score on the Haberman instrument and teacher effect scores as measured by SAS EVAAS, value added data as noted by the End of Grade results in middle school mathematics. Teachers were grouped by their years of experience and placed into two different cohorts. Teachers were also grouped by their total scores on the Haberman interview tool and then these results were merged together to determine any significant relationship. All End of Grade test scores in 2008 for children in grades 6-8 in the Guilford County School System were used to match with specific teachers to determine the value added data for these individuals. This data is obtained from the SAS Institute in Cary, North Carolina. A few descriptive analyses were conducted to demonstrate the relationship between Haberman Scores and teachers' years of experience with the growth students made from that particular teacher. The results of this study hinged on the accuracy of the value added assessment formula and the assumptions of the general linear model. Although there was not a significant difference between the years of experience and teacher effect data, the data indicated that the more experience a teacher had the higher their mean of student growth. Additionally, the data indicates that there is no significant relationship between an effective teacher as measured by the score on the Haberman instrument and the teacher effect scores as measured by SAS EVAAS data.

While the average growth on the Math 2008 scores increased by 1 point or so, the difference is not significant enough to see that a teachers' score on the Haberman test has any validity to increasing student proficiency in middle school math. The results of this study could be used by this or any other school system to look at the process used to place teachers in the front of our classrooms. Our children are our most precious commodity and their future begins with us.

DETERMINING THE CORRELATION OF EFFECTIVE  
MIDDLE SCHOOL MATH TEACHERS AND  
MATH STUDENT ACHIEVEMENT

by

Jocelyn B. Becoats

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  
2009

Approved by

---

Committee Chair

© 2009 by Jocelyn B. Becoats

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 Chair \_\_\_\_\_

Committee Members \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_  
Date of Acceptance by Committee

\_\_\_\_\_  
Date of Final Oral Examination

## ACKNOWLEDGMENTS

*Giving all honor and glory to My Lord and Savior Jesus Christ, this dissertation is dedicated to my two girls, Iman and Jady, for your love and understanding during these last three years. Both of you are my heart and joy and this was done to show you that no matter what obstacles life brings your way, stay true to who you are and what you know you can do. In the words of Maya Angelou, “We may encounter many defeats, but we must not be defeated.”*

I cannot begin to thank my family for everything that you all have done to ensure my success in the doctoral program. Whether it was cooking meals, keeping the children, talking through difficult times, or even taking family trips, I thank you from the bottom of my heart for being with me through thick and through thin. To my parents, Ruby and Grady Bethel, thank you for your encouragement, guidance, patience, and love while staying this course. To my sister and brother-in-law, Jaleria and Myles Burt, thanks for being there for me through all of the tough times when I needed you the most.

Words cannot express my sincere thanks to Eric Becoats, for sticking with me through the writing of this dissertation. Even though our relationship has gone through ups and downs, you were always very supportive and encouraging through this process. I wanted to quit, thanks for making me persevere.

To my friends and family, I could not have done it without you. Lise’ Timmons, Angela Houston, Nikki Hill, Joye Henderson, and Danita Curry always had listening ears

and ways of encouraging me when times got tough. Thanks for whatever you did to allow me to complete this level of accomplishment.

To my colleagues in GCS, Beth Folger, Barbara Zwadyk, Phyllis Martin, Wanda Legrand, and Lewis Ferebee, thanks for your motivation and encouragement throughout this entire process. The pats on the back, words of wisdom or even the signs on the wall continued to remind me that I could finish this task. My dear friend Netta, thanks for listening to me rant and rave when times were tight and for your assistance in whatever I needed.

Last, but not least, to my dissertation chair, Dr. Carl Lashley, many thanks for allowing me to take this at my pace and not holding me back. Your guidance and support was invaluable. To the other members of my committee, Drs. Cooper, Gause, and Williams, thanks for reading over the material and responding so positive. I was truly honored to have each of you on my committee.

## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
CHAPTER	
I. INTRODUCTION .....	1
Problem Statement .....	2
Significance of the Study .....	5
Background and Overview of Methodology.....	9
Organization of the Study .....	11
II. REVIEW OF THE LITERATURE.....	12
Teacher Effectiveness .....	13
Effective Teachers and Current Accountability .....	18
Teacher Selection .....	21
The Late 1970s and 1980s.....	21
Teacher Selection in Guilford County .....	24
The Haberman Approach.....	25
Guilford County’s Use of the Haberman Interview Tool .....	28
Teacher Retention.....	30
Value Added Data as an Indicator of Teacher Effectiveness .....	37
History.....	37
Definition and Purpose.....	39
Challenges of Value Added Assessment.....	42
North Carolina: No Child Left Behind and Adequate Yearly Progress (AYP).....	48
Guilford County’s Use of Value Added Assessments .....	51
III. RESEARCH DESIGN AND METHODS .....	57
Purpose of Study .....	57
Research Questions .....	57
Research Approach.....	58
Description of Sample .....	60
Methodology and Data Analysis Plan .....	62
End of Grade Tests: Mathematics.....	64

Developmental Scale Scores .....	64
Teacher Effect Data and Teacher Years of Experience .....	65
Haberman Scores.....	66
Teacher Effect Data with Years of Experience Compared to Haberman Scores .....	67
Limitations .....	69
Validity of Research .....	70
Definition of Terms .....	71
 IV. RESULTS.....	 75
Introduction.....	75
Demographic Data.....	76
Teacher Effect Scores and Years of Experience .....	79
Teacher Haberman Scores and Teacher Effect Data .....	81
Teacher Effect Data by Years of Experience and Haberman Scores .....	82
Research Question 1 .....	85
Research Question 2 .....	86
 V. FINDINGS AND DISCUSSION.....	 91
Summary of Research Problem.....	91
Findings .....	92
Discussion .....	95
Conclusions.....	102
Recommendations for Future Study.....	105
Recommendations for Policy Makers .....	106
Recommendations for School Administrators .....	107
 BIBLIOGRAPHY .....	 109
 APPENDIX A. IRB APPROVAL .....	 117
 APPENDIX B. SAMPLE VALUE ADDED REPORT.....	 118

## LIST OF TABLES

	Page
Table 1. School and Teacher Effectiveness Impact on Learning Entering School at the 50th Percentile .....	15
Table 2. School Status Labels and Recognitions .....	50
Table 3. Number of Schools in Various State Recognition Categories .....	51
Table 4. Description of Data Included in the Study .....	61
Table 5. Teachers' Years of Experience and 2008 Mean Teacher Effect Scores Estimated Marginal Means .....	80
Table 6. Analysis of Teacher Effect Data by Haberman Scores for Sixth- to Eighth-Grade Teachers.....	81
Table 7. Analysis of Teacher Effect Data by Years of Experience and Haberman Scores for Sixth- to Eighth-Grade Teachers .....	83
Table 8. Analysis of Teachers' Years of Experience and Haberman Scores .....	88

## LIST OF FIGURES

	Page
Figure 1. Teacher Turnovers in North Carolina from 1998-2009.....	34
Figure 2. Number of Sixth- to Eighth-Grade Teachers by Years of Experience .....	77
Figure 3. Middle School Teacher Haberman Scores Based on Score Levels.....	78
Figure 4. Mean of Teacher Effect Data by Haberman Scores .....	82
Figure 5. Analysis of Teacher Effect Data by Years of Experience and Haberman Scores for Sixth- to Eighth-Grade Teachers.....	84

## **CHAPTER I**

### **INTRODUCTION**

Over the past 30 years, the “art” of teaching has rapidly become the “science” of teaching (Marzano, 2003). This is not to imply that effective teaching strategies were absent before 1970. At the beginning of the 1970’s, researchers began to look at the effects of instruction on student learning (Stewart, 2006). The Coleman Report (1996) concluded that school really made little difference in the achievement of students. Coleman and colleagues, mathematical sociologists, concluded that the quality of schooling a child received only accounted for 10% of the variance in student achievement. What influenced the other 90%? Coleman and others believed that the differences in student achievement could be attributed to other factors such as a student’s natural ability or aptitude, socioeconomic status and home environment. These aspects cannot be changed by the school itself. What then can the school control which will affect student achievement? The answer to this question is having qualified teachers in our classrooms on a daily basis; however, selecting them is critical (Wise et al., 1987; Darling-Hammond, 1984).

Various factors can impact a school district’s ability to select quality teachers. Enrollment trends, retirements, and NCLB (No Child Left Behind) initiatives have caused an increase in the demand for qualified and effective teachers in our classrooms. Teacher supply has been on the decline and has created competition among school

districts to find teachers. In addition, finding teachers in critical areas such as math and science have been more challenging for districts.

According to various case studies there are four significant areas of research that are relevant to teacher selection: the quality of information about teaching candidates, teacher selection methods and criteria, the effects of the teacher selection processes, and teacher induction practices (Wise et al., 1987). NCLB requirements for highly qualified teachers refer specifically to the minimum qualifications and credentials a teacher brings to the position. However, research shows that classroom processes and outcomes are a better indicator of teacher effectiveness when trying to select a teacher. In New York City's The Equity Project, leaders agree that "a golden resume' and a well- run classroom are two different things" (Gootman, 2009).

School districts within our state must focus on teacher selection efforts by finding the positives of the district and selling those to potential candidates while reducing barriers for teachers to accept employment. Mostly importantly, districts must sell their schools to the right candidates. We do not have time to waste in the hiring process because our students are sitting without teachers in the classroom. Our children are our future of tomorrow and we must continue to prepare them for the twenty-first century.

### **Problem Statement**

North Carolina has a need to hire new teachers especially in the fields of math and science. This need of hiring new teachers was supported by Darling-Hammond as early as 1994 when she found that more than 200,000 new teachers will be needed annually across the nation in response to teacher retirements and rising student enrollments at a

time when teacher-education preparation in programs were only producing 100,000 to 150,000 new teachers per year. Further, this is a problem nationally as “the nation needs to recruit 2.2 million new teachers by the end of the decade, including 200,000 specialists in mathematics and science” (Hoff, 2000, p. 7). According to the North Carolina Center for Public Policy Research, there are currently 86,000 teachers in North Carolina’s public schools and the state must hire about 10,000 teachers each year just to staff existing classrooms (McLaughlin, 2004). Yet, all the state’s public and private universities combined produce only about 3,100 teachers a year (McLaughlin, 2004). The Council of Great City Schools indicated that there is an immediate need for new teachers in urban schools in almost every teaching area, and especially for teachers of color and language minorities (The Urban Teacher Collaborative, 1996). The need for good teachers continues to grow and is reaching critical proportions, especially with math teachers in urban secondary schools. Experts at the National Teacher Recruitment Clearinghouse predict that over the next ten years the nation will need 2.2 to 2.4 million teachers. That means there will be 150,000 to 250,000 openings for teachers in the nation's elementary and secondary schools (Darling-Hammond & Berry, n. d.).

The Guilford County school district in Greensboro, NC experiences turnovers each year and to find replacements for teachers is a task of the human resource department. With the large number of vacancies within districts, specifically within the Guilford County School district, a screening of some sort is used in order to hire the most effective and knowledgeable teachers according to the human resource division. The Guilford County School System (GCS) used an interview tool to select highly qualified

effective individuals to serve in the urban schools. As the district hires teachers, there is a need to determine whether they are using the best interview tool, using the scores appropriately and hiring the teachers that will produce results in terms of math achievement.

Martin Haberman, professor of education and researcher of successful teachers, declares that teachers are at the root of whether the instruction in the classroom can be deemed as effective or ineffective (1987, 1995). He believes that teachers are not adequately prepared for teaching and the selection process of faculty is significant in the retention and effectiveness of classroom teachers. Haberman's focus is mainly on urban school districts and children in poverty. The process of identifying, refining, and explaining what an effective teacher of children considered to be at risk, in poverty, and in urban schools was his lifetime study. After determining the specific characteristics of effective teachers, Haberman developed an interview to predict which teachers would succeed with at risk students and continue teaching these youth and which would fail, or leave the profession. This tool is known as the Star Teacher Selection Interview.

The Guilford County School district has been a leader in using national models for interviewing. Most major models include both on-line screeners and a follow-up face to face structured interview as does the Haberman instrument. GCS sees the purpose of selecting a model to be used as a district-wide interview tool is to provide a systematic solution that is research-based, aligned to core values and best practices, and statistically reliable and valid. Since principals hire up to 600 new teachers each year for GCS, the human resources department finds it is important to have an interviewing tool that best

identifies the qualities and attributes needed to teach in an urban setting (Guilford County Schools, 2008). The question is whether this Haberman tool really selects effective middle school math teachers that produce growth with students in the classroom as measured by a statistical approach to determining student growth within a classroom. This research study investigated whether there was a relationship between teacher effect data and years of experience. Additionally, this study determined if there was a relationship between an effective teacher as measured by the hiring instrument (Haberman Interview tool) used in the Guilford County school district against a teacher's value added data as noted by the End of Grade test scores in middle school mathematics. Permission to utilize Haberman interview data, value added data, and end of grade test scores was granted in the summer of 2008 from the Guilford County School district. For this study, the researcher received no identifiable data on teacher's names, value added scores, or Haberman interview scores. All information was kept confidential from the Human Resource office and the Chief Accountability Officer.

### **Significance of the Study**

For several years the Guilford County School District had an instructional position, Instructional Improvement Officer (IIO), which assisted principals with curriculum strategies and management techniques. The role of the IIO was to also help with principal hiring decisions as well as evaluating the principals. Upon being hired by the district in 2006, I assumed the role of one of the middle school instructional improvement officers and remained in this role for two years. During this time, I assisted a cohort of middle school principals in making decisions regarding teacher applicants.

Although teachers interviewed with the individual principals, all candidates interested in coming into the system had to go through an on-line Haberman screener and then, if passed, had a face-to-face interview from a principal group using the Haberman interview instrument. Many times, principals indicated that there were candidates that after tallying their scores did not fall into the category set by GCS that they wanted to hire, but principals felt strongly that through their interview the candidate had a deep care and concern for children in poverty that would fit into the personnel at their school site. This posed a dilemma for me because my job was to support the principal. The principal was also stuck between a rock and a hard place because they were not able to hire any candidates who did not have a certain score on this instrument. Consequently, teachers (with scores considered as “high” on the Haberman) were hired by the district in the hard to staff areas and placed in the schools with vacancies. Since the hard to staff areas consisted of mathematics and these are the scores that determine a schools’ effectiveness or not, this posed a dilemma for me. As a former math teacher, I had a sincere desire to determine if the Haberman instrument was really an instrument that ultimately produced student growth on the end of grade tests.

After two years of serving as an IIO, I became the middle school curriculum and Instruction Officer where I served for one year. As the middle school curriculum officer, my department was responsible for facilitating curriculum, instruction and professional development for all middle schools. In my current role as the K-8 curriculum officer, I continue to see the need to collect data on teachers who have high scores and correlate those to what the teachers are actually doing in the classroom and to their value added

data. For teachers who are not showing growth as determined by the End of Grade results (EOG's), my department is responsible for determining if the curriculum needs revision or if more professional development is needed with the teachers. The content specialists within my division carefully study the test score data to determine whether the curriculum needs to be revised or teachers need more professional development. Since the teachers were employed based on a projection of their effectiveness as evidenced by the Haberman scores, there is an assumption that their teaching is adequate to the task. Therefore, if students are not scoring at proficient levels, the problem must lie either in the curriculum or in undeveloped skills that the teachers need to be provided through professional development. The question that still remains unanswered is whether professional development will work with particular teachers, and does the Haberman process accurately forecast teacher effectiveness? Since I work with middle school curriculum on a daily basis, I wanted to determine if the teachers in our classrooms that were deemed as "effective" based on the qualities of Martin Haberman (through his interview) are showing growth with students in their classrooms. This study will examine the relationship between the teacher selection process (use of Haberman Interview tool) and teacher effectiveness in math achievement. Since our state tests our students in reading and mathematics, we have data on students from one year to the next. Therefore, I chose to look at middle school mathematics teachers in my research.

The implications of this study could mean that the teachers who have scored high on this instrument are not producing great results when it comes to student growth. In this case, if a teacher is not producing a certain amount of growth, he/she could be taken out

of the classroom or even the school. Additionally, this research could prove that as a district we need to use more than just this one method of hiring individuals for our schools or not use this product at all. Having effective teachers in the classroom, which we know will produce high student achievement and engagement, is at the root of this research.

As I look within the Guilford County school system, we have continuously made incremental growth in math performance with middle school students; however, we have not met the goal of the state. In 2005-2006, the percent of sixth-grade students scoring at or above grade level in mathematics was 59.87 percent, in 2006-2007 was 66.04 percent, in 2007-2008 climbed to 71.9 percent and then as recent as 2008-2009 climbed to 79.2 percent. In 2005-2006, the percent of seventh-grade students scoring at or above grade level was 59.78 percent, in 2006-2007 was 62.2 percent, in 2007-2008 was 69.4 percent and in 2008-2009 rose to 79.8 percent. The percent of eighth-grade students scoring at or above grade level in 2005-2006 was 57.4 percent, in 2006-2007 was 65.3 percent, in 2007-2008 was 67.7 percent and finally in 2008-2009 rose to 80.3 percent. While the system continues to make growth at each grade level, the state department has indicated a goal of 77.2 percent for North Carolina school districts. The question left standing is do we have the right people before our children on a daily basis to make such growth with our students. Since the Guilford County School district uses an interview tool by Martin Haberman, the correlation of the high scores on this instrument with the academic growth of students could be done.

## **Background and Overview of Methodology**

Quantitative research examines phenomenon through the numerical representation of observations and statistical analysis and quantitative analysis uses various statistical methods for deriving results. For this research study, a quantitative method of methodology was used. This approach was used to determine the connection between Haberman scores and statistical data as a representation of teacher effectiveness.

Quantitative research is generally approached using scientific methods, which include: the generation of models, theories and hypotheses; the development of instruments and methods for measurement; experimental control and manipulation of variables; collection of empirical data; modeling and analysis of data; and the evaluation of results (Wiersma & Jurs, 2005).

Quantitative research primarily occurs to determine relationships, effects, and causes (Wiersma & Jurs, 2005). This study utilized modeling and analysis of value added data compared to an analysis of the Haberman scores. The evaluation of these two analyses helped to determine if there was a positive relationship between effective teachers and student achievement. This study was quantitative in design because of my interest in determining relationships among variables: sequence of teacher effectiveness and math student achievement, and teacher effectiveness and years of teaching. The effectiveness of middle school math teachers in an urban school district was measured by students' math achievement as determined by their performance on the End of Grade tests in grades 6-8 from students. The North Carolina End-of-Grade Test of mathematics, which is designed to meet the need for higher standards of achievement for North

Carolina students, assesses grade-level knowledge and skills established by the North Carolina *Standard Course of Study*. The *Standard Course of Study* identifies the following seven strands for mathematics: numeration, geometry, patterns and pre-algebra, measurement, problem-solving, data analysis and statistics, and computation. The test contains both an applications section and a computation section with a total of 80 multiple-choice questions at each grade level. The two parts of the test are combined to provide one total score. The applications section of the test examines a student's ability to solve problems and explain mathematical processes. Students may use calculators and rulers for this part of the test. The computation portion of the test examines a student's ability to complete computations without technical assistance.

Student performance in mathematics is reported in a variety of contexts, including developmental scale scores and achievement levels. Group and subgroup scores are also reported. Although the mathematics tests were developed as accountability tests, curriculum information is provided on the goal summary reports. In addition, developmental scale scores depict the amount of growth that actually occurs in student achievement over time. These scores can be used to compare the expected rate of growth to actual student or group performance as the student(s) move from grade 3 to grade 8. Achievement levels show how a student or group of students performed in relation to predetermined grade-level standards set using teacher judgment linked to student performance.

Achievement scores of all students taught in middle school math were matched against the value added data of the teachers. In 1982, Bill Sanders, a Tennessee statistics

professor, created a technique called value-added modeling, which uses a complicated statistical system to compute data. Value-added modeling works by attributing growth patterns higher or lower than average to teachers' skills—or the lack of them. Using this model, information was gathered from the SAS Institute with the permission of Guilford County Schools. Teacher effect data was determined for each of these middle school math teachers. The relationship between teacher effect data and years of experience was also examined. By using a multiple discriminant method, teachers were then categorized by their years of experience in teaching. From this correlated data, teachers were categorized based on their score on the Haberman interview tool using the STAR teacher qualities as denoted by Martin Haberman, education professor, for highly effective teachers in an urban district. The results of this study helped to determine if the sole selection criteria for placement of teachers in this urban district had a positive relationship to student growth.

### **Organization of the Study**

This dissertation was designed into five chapters. The first chapter includes an introduction, a problem statement, significance of the study, and a description of the study (including the methodology). Chapter II consists of a review of the literature on teacher effectiveness, teacher selection, the Haberman Approach, teacher retention, and Valued Added Data as an indicator of teacher effectiveness. The research design and methodology is described in Chapter III. An analysis and discussion of the results is presented in Chapters IV and V, respectively.

## **CHAPTER II**

### **REVIEW OF THE LITERATURE**

In recent years, assessment data played a major role in education policy and practice. The public, the States, and the Federal Government want ways to determine if schools are making an increase in student achievement and holding them accountable for results. Test-based accountability systems have motivated researchers and policymakers to explore many methods for analyzing student achievement data in order to evaluate the effectiveness of public schools. Most commonly, school effectiveness decisions have hinged upon levels and changes over time in achievement measures for successive cohorts of different students and ranks of schools based on such measures (Meyer, 1997). We have been giving students achievement tests in one form or another for many years, but not until recently with the No Child Left Behind Act (2002) have we looked at the instruments that measure the achievement level of individual children as well as schools in a given group. Additionally, it is the quality of the teacher in the classroom who has a large effect on the results of the assessment data (Wong, 2007).

In this chapter a review of literature and research is presented. The chapter is divided into five major sections: Teacher Effectiveness, Teacher Selection, The Haberman Approach, Teacher Retention and Value Added Data as an indicator of Teacher Effectiveness.

## Teacher Effectiveness

The definition of effective teaching really depends on who you ask. When I was in school, effective teachers were those whose desks were neatly arranged in rows and where the students were always quiet and on task. Teachers may say that effective teachers are those who have good content knowledge, plan thorough lessons, and have good classroom management. If students were asked to describe an effective teacher, they would probably say one who is friendly, fun, and doesn't give too much homework. Effective teachers have a combination of those characteristics; they have a positive student rapport, strong content background, well planned lessons, and engaging instructional delivery. In addition, effective teachers differentiate their classroom instruction in an effort to meet the needs of all students. One would expect a teacher with all of these qualities to ultimately raise student achievement (Wong, 2007).

A research synthesis for the National Council on Teacher Quality (Goe, Bell, & Little, 2008) takes teacher effectiveness and breaks it down into five critical points:

- Effective teachers have high expectations for all students and help them learn, as demonstrated on value-added, test-based, or alternative measures.
- Effective teachers contribute to positive academic, attitudinal, and social outcomes for students such as regular attendance, on-time promotion to the next grade and graduation, self-efficacy, and cooperative behavior.
- Effective teachers use diverse resources to plan and structure engaging learning opportunities; monitor student progress formatively, adapting instruction as needed; and evaluate learning using multiple sources of evidence.
- Effective teachers contribute to the development of classrooms and schools that value diversity and civic-mindedness.
- Effective teachers collaborate with other teachers, administrators, parents, and education professionals to ensure students' success, particularly the success of students with special needs and those at high risk of failure. (Goe, Bell, & Little, 2008, p. 8)

According to Harry Wong (1996), the greatest effect on student achievement is the effectiveness of the teacher. The teachers we hire today will become the teachers for the next generation. Their success will determine the success of an entire generation of students. Over 200 studies (“What Matters Most,” 1996) have said that the most significant factor in improving student learning is with a knowledgeable and skillful teacher. This is something we already know. We’ve known it for years, but we will not implement the obvious.

**We have known the following about teachers for decades:**

- Teacher expertise accounts for more difference in student performance—40 percent—than any other factor (Ferguson, 2001).
- Students who have several effective teachers in a row make dramatic achievement gains, while those who have even two ineffective teachers in a row lose significant ground (Sacks, 1999).
- Based on research in Texas, the importance of having an effective teacher instead of an average teacher for four or five years in a row could essentially close the gap in math performance between students from low-income and high-income households (Hanushek, Kain, & Rivkin, 2001).
- The difference in teacher effectiveness is the single largest factor affecting academic growth of populations of students (Sanders, 2000a or b).

The achievement gap facing poor and minority students is due not to poverty or family conditions, but to systematic differences in teacher quality (Wong, 2007). A student who is taught by an ineffective teacher for two years in a row can never recover the learning

lost during those years (Marzano, 2003). As a teacher's effectiveness increases, the first group to benefit from this improvement is the lower achieving students. Additionally, one set of research proposed that assigning great teachers five year in a row to a class of disadvantaged children could close the achievement gap between these students and their peers (Hanushek, Kain, O'Brien, & Rivkin, 2005). Currently, the odds that a child will be assigned a great teacher five years running are 1 in 17,000 (Walsh, 2007).

The ineffective teacher affects little, if any growth in students. The effective teacher, even in an ineffective school, produces improved student learning and increased student achievement. A study of the Denver Public Schools found that teachers accounted for more than twice the total variation in student test score change than did the schools (Meyer, 2001).

Marzano's research indicates that if a child is at the 50th percentile and is placed in one of the following situations, after two years he believes as stated in Table 1 that the following will happen:

**Table 1**

*School and Teacher Effectiveness Impact on Learning Entering School at the 50<sup>th</sup>*

*Percentile*

School and Teacher	Percentile after Two Years
Ineffective school and Ineffective teacher	3 <sup>rd</sup>
Effective school and Ineffective teacher	37 <sup>th</sup>
Average school and Average teacher	50 <sup>th</sup>
Ineffective school and Effective teacher	63 <sup>rd</sup>
Effective school and Average teacher	78 <sup>th</sup>

Marzano (2003) states that it is better to train ineffective teachers and let the teacher leave than to not train a teacher and allow them to stay. So what does this mean as we hire teachers? Are there certain characteristics we look for in teachers that show their effectiveness?

A number of characteristics have been identified in effective teachers. Teachers should be friendly, helpful and congenial. They should be able to empathize with students, understand their world and listen to them (Wubbels, Brekelmans, van Tartwijk, & Admiral, 1999). The establishment of an atmosphere of caring, warmth, and understanding is essential to effective teaching. Students perceive teachers who are responsive and assertive as caring. Not only did effective urban teachers possess the requisite of academic knowledge and pedagogical strategies, but they also possessed a belief system that was more sensitive to the specific needs of the students.

A review of the literature cited by Stronge (2002) listed the following four teachers' affective characteristics as related to their effectiveness and perceived effectiveness. Social interactions with students give teachers the opportunities to demonstrate caring, fairness, and respect which assists in cultivating a positive learning environment and promoting student achievement. Effective teachers demonstrate caring by listening, understanding and knowing their students. Enthusiasm for learning, teaching and subject matter supports positive relationships with students and encourages student achievement. Finally, a teacher's dedication to students and the job of teaching provides evidence that the teachers view themselves as responsible for the success of their students (Strong, 2002).

According to Collinson's (1994) literature review, three characteristics of an exemplary teacher appear across the literature. Exemplary teachers have "a love of learning, an ethic of care, and commitment to teaching" (p. 2). Teachers who see themselves as having a positive capacity to cope with stress may be less likely to suffer from burnout than those who approach stress negatively. These teachers' professional renewal was supported by a disposition to seek alternatives, to take an active role in adapting curriculum and learning about their students, and by a deep belief that they could make a difference and that education is important.

Treadway (1999) indicated that urban teachers should examine their attitudes, teaching practices, and their affect on students instead of blaming students when handling complex issues. She cited other research indicating prior work experience, a conscious choice to pursue teaching, advocacy experiences relative to social justice, diversity and morality issues as desirable for urban teachers. Urban teachers must be able to visualize what their classroom will look like and envision how they will participate in the process of educational transformation. Teacher quality is the most important school related factor influencing student achievement (Rice, 2003). Effective teachers spend more time working with small groups throughout the day. Teachers who have higher rates of communication with parents are viewed as more effective. They also have more students in their classes on task and engaged in learning throughout the day.

The Center for Public Education (2005) monitored state and district efforts on practices that provide children with highly qualified and effective teachers. It noted that the following were teacher qualities that showed a positive relationship to student

achievement: content knowledge, teacher experience, teacher training and credentials, and the overall academic ability. There were two studies (Texas and Tennessee) that provide empirical evidence that teachers make a substantial difference in student achievement.

### *Effective Teachers and Current Accountability*

It is important to note that considerable research has been conducted on the characteristics of effective teachers, but today the effectiveness of the teacher depends heavily on student performance, because of No Child Left Behind. School and district administrators of today would most likely describe teachers as effective if, their students show academic growth on the standardized tests they take at the end of the semester or school year. The No Child Left Behind (NCLB) legislation emphasizes school accountability for performance on standardized tests, almost to the exclusion of other indicators of school effectiveness.

The most compelling evidence for the importance of teaching came initially from economists who adapted value-added models from business to measure the effect of teachers on student learning. While the statistical methods are complex, the definition of effective teaching is not. Simply, researchers looked for the change in students' test scores according to the teacher they were assigned. A highly effective teacher, therefore, is one whose students show the most gains from one year to the next. By using this approach, researchers were able to isolate the effect of the teacher from other factors related to student performance, for example, students' prior academic record, student's socioeconomic status or school they attend.

One initiative in Tennessee, the Tennessee Value Added Assessment System (TVAAS), provided an explanation of how much an effective teacher had on student outcomes. TVAAS was the first data-tracking system in the country to measure individual teacher performance according to annual gains in student test scores. Initiated in 1990, this system provided extensive data on state achievement tests for all students in grades 2-8 in Tennessee and allowed for comparisons of teacher effects on students' learning. Other states, such as North Carolina, Arizona, and Florida, have since adopted similar models; additional states are expected to follow suit. Analyses of TVAAS data indicated that teachers had a substantial effect on student achievement. The differences in student performance were more heavily influenced by the teacher than by student ethnicity or class, or by the school attended by the student (Darlington, 1997).

The positive effects associated with being taught by a highly effective teacher, defined as a teacher whose average student score gain is in the top 25 percent, were stronger for poor and minority students than for their white and affluent counterparts. For example, a study by Nye, Konstantopoulos, and Hedges (2004) found that low-income students were more likely to benefit from instruction by a highly effective teacher than were their more advantaged peers. A study by Sanders and Rivers (1996) found that the achievement gains from having a highly effective teacher could be almost three times as large for African American students as for white students, even when comparing students who start with similar achievement levels.

A second important finding from the work of Sanders and Rivers was that the positive effects of teacher quality appear to accumulate over the years. That is, students

who were enrolled in a succession of classes taught by effective teachers demonstrated greater learning gains than did students who had the least effective teachers one after another. For example, fifth-grade math students who had three consecutive highly effective teachers scored between 52 and 54 percentile points ahead of students who had three consecutive teachers who were least effective, even though the math achievement of both groups of students was the same prior to entering second grade (Sanders & Rivers, 1996). This work mirrors the work of Marzano and his study on effective teachers.

In conclusion, effective teachers are measured by the results they have with the students. Building relationships, caring about individual students, understanding and knowing content collectively create an effective teacher. It is evident that effective teachers are needed in our schools to instruct our students. We still have a problem finding those effective teachers, placing them in the appropriate grades and schools and retaining them.

Additionally, research consistently shows that teacher quality, whether measured by content knowledge, caring attitude, experience, or intellect, is strongly related to student achievement. Simply put, skilled teachers produce better student results. With NCLB affecting school districts around the country, more emphasis is being placed on the selection of teachers in the classroom.

## **Teacher Selection**

### ***The Late 1970s and 1980s***

Growing enrollments, pending retirements, and NCLB initiatives have given a rise to the increasing demand for qualified and effective teachers in our public schools. The supply of teachers has been on decline and therefore the competition among school districts to find teachers is on the rise.

Numerous factors affect a school district's ability to find and hire quality teachers. Researchers Boardman, Darling-Hammond, and Mullin (1982) have found that school district characteristics such as geographic location, climate, neighborhood and student characteristics, cost of living, class size, and other variables affect teacher supply. Secondly, as state policy tightens accreditation and certification standards for schools, it reduces the number of eligible candidates for teaching positions. More stringent certification requirements make this situation worse given the decline in the academic ability of those who enter teacher education programs and then the teaching force (Weaver, 1979). Teacher certification testing has been implemented throughout most of the nation. The Educational Testing Service (ETS) and National Evaluation Services (NES) provide testing programs for the states in basic skills, subject content, and pedagogy and professional knowledge. Since these tests are in most states, they can have a dramatic affect on teacher supply. There has been a growing movement to treat teaching as a licensed profession by creating professional standards and practices board. Finally, the effectiveness of the teacher selection process depends on the type of teachers sought by the school district. The evaluation of potential teachers entails assessments of

their qualification, interpersonal characteristics, and even performance. However, such assessments vary depending on the goal of the district.

According to case studies conducted by Wise et al. (1987), while districts have similar selection procedures, there are differences in (a) the criteria in the selection tools, (b) the weights placed on the different indicators of teacher ability, (c) the extent to which selection processes are centralized or decentralized, and (d) the manner in which teachers are treated before and after the hiring. Although there were differences, there were five significant areas of research that are relevant to teacher selection: the quality of information about teaching candidates, teacher selection methods and criteria, the effects of the teacher selection processes, and teacher induction practices.

Research that has sought to identify specific teaching behaviors linked to teaching effectiveness has produced some promising results (Gage, 1978; Rosenshine, 1979). However, these behaviors are more likely to be high-inference variables representative of teaching patterns such as clarity, variability, and ability to engage students in learning activities than of discrete measurable actions that can be assessed in short-term situation tests. Other researchers have found that effective teaching behaviors vary for students of different socio-economic, mental, and psychological characteristics (Cronbach & Snow, 1977). The literature on teacher characteristics has not identified any certain set of skills, attitudes, interests, or abilities that consistently discriminates between effective and ineffective teachers (King, 1981).

There were several methods by which school districts select teachers. The interview was one instrument that school administrators prefer to use during the selection

process (Bredeson, 1982). Young (1984) noted that most researchers have addressed the issue of interview validity by correlating an interview score with some criterion of job performance. He asserted that their findings were disappointing for several reasons—but primarily because “validity coefficients were generally low and accounted for only a small portion of the variance in job performance and the correlation coefficients failed to provide any insight relative to why the interview was found to be valid or in-valid” (p. 43). To compensate for some of the deficiencies in the “typical” selection interview, some school districts are relying on structured interview systems. Interviewers were trained on asking specific predetermined questions that are subsets of theme areas related to what are believed to be the characteristics of effective teachers. Interviewers scored and interpreted a prospective teacher’s responses according to predetermined criteria and standards. Some developers of these instruments claim that the instruments, when used in conjunction with other hiring strategies, could increase the probability of hiring more highly rated teachers (e.g. see Muller, 1981). Other researches have concluded that these structured interview systems did not enhance the effectiveness of the teacher selection process (e.g. Smith, 1980). Regardless of this, the interview revealed a lot of information about prospective teachers that other selection strategies could not. Lastly, school systems must determine their own selection criteria. Smith (1980) has noted three critical criteria: school systems must determine if the candidate has mastery of fundamental knowledge, has mastery of instructional skills necessary to be an effective teacher, and will use their skills and knowledge to best fit the needs of the school and the system.

Smith noted that the degree to which school systems mix the three criteria together will determine the quality of the teacher hired.

Even though many school districts may differ in their selection process, the Guilford County School (GCS) system elected to use a teacher interview tool that was believed to select candidates that have qualities and attributes for teaching students in an urban setting.

### ***Teacher Selection in Guilford County***

The teachers in Guilford County are hired from many in-state, out-of-state, and international sources. These teachers are carefully recruited, screened and selected using a multi-tiered approach that addresses federal and state requirements, aligns to district priorities and values, and includes key stakeholders.

Teachers are recruited from in and out of state college and university campuses that yield the greatest production of teachers, particularly in hard to staff areas. The school system also conducts recruitment/hiring fairs and advertises through on line search companies. GCS provides differentiated salary schedules for certain licensure areas and for teachers in hard to staff schools. Each year, approximately 10,500 teachers apply to teach with GCS according to records.

The GCS school district implemented a district-wide interview using the Ventures for Excellence Instrument in 2002. That instrument assisted the district in screening for teachers who had unique dispositions to teach. Specifically, the Ventures process screened candidates for their ability to express their unconditional positive regard for people, their ability to define the importance of building positive relationships, their

articulation of teaching/learning strategies to meet student needs, and their ability to make content meaningful to learners. This information was important to me because the value added data for teachers that were hired using the Haberman instrument was compared to other teachers, the other teachers may have been hired using the Ventures instrument. The Ventures process was used from 2002 until January 2005. At that time, GCS transitioned from Ventures for Excellence to the Haberman Star Teacher and Principal (2005) process. The process is currently in its second full year of implementation.

### **The Haberman Approach**

Teacher attrition and retention have been credited in part to shortsighted recruitment and selection processes. School leaders tend to narrowly focus on selecting a qualified candidate to fill a vacancy (Norton, 1999). Is there a match between the candidate's personal characteristics and philosophy and those of the district?

The Urban Teacher Collaborative report (1996) indicated support for innovative approaches to selection such as urban teacher interview instruments and other experience based criteria. Eighty-seven percent of the Council of Great City School Districts that responded to the 1994-95 survey indicated that structured interviews that assessed aptitudes and characteristics were very useful as teacher selection criteria.

There was a general agreement that the "highly qualified" teacher referred to in the "No Child Left Behind" Act of 2001 would possess the subject-matter knowledge and the experience to influence student achievement positively (Ansell & McCabe, 2003).

But what personal characteristics will assist teachers in transmitting that subject-matter knowledge to students and to staying in the teaching profession?

Martin Haberman, professor of education and researcher on the study of successful teachers, believes that teachers are at the root of what actually occurs in the classroom. He believes that teachers were not adequately prepared for teaching and the selection process for teachers is significant in the retention and effectiveness of teachers in the classroom. His focus was mainly on urban school districts and children in poverty. The process of identifying, refining, and explaining what an “effective” teacher of children considered to be at risk, in poverty, and in urban schools was his lifetime study. Determining the specific characteristics of effective teachers turned into an interview to predict which teachers would succeed and stay with at-risk youth and which would fail or leave. This tool is known as the Star Teacher Selection Interview.

Haberman stated that only teachers who possess a particular set of attributes and ideology could successfully offer urban students the multicultural curriculum and support that they need (Haberman & Post, 1998). He has developed and refined an instrument that purports to identify seven important characteristics:

<u>Question #</u>	<u>Characteristic</u>
1	Persistence
2	Response to Authority
3	Application of generalizations
4	Approach to at-risk students
5	Personal vs. professional orientation to teaching

6 Burnout

7 Fallibility

(Haberman et al., 1993). The instrument does not discriminate against any ethnic or gender group. However, the instrument does tend to sort out individuals who are immature, unfamiliar with poverty or who lack resilience.

The ability to develop interpersonal relationships and the communication skills developed by teacher candidates with extensive life experiences enable them to exhibit several of the characteristics enumerated by Haberman as necessary in effective urban teachers. Although age is not a factor, candidates who have greater life experience tend to score higher on the Urban Teacher Selection Instrument (UTSI) (Haberman, 1995b).

“The primary goal of a successful urban teacher is to teach pupils effectively whether such pupils are lovable or not” (Haberman, 1987, p. 47). This attitude exemplifies persistency and a professional orientation towards teaching on the part of the teacher. These teachers must believe that they can make a difference in the lives of their students. This belief causes them to approach their work differently than those who believe that “the locus of control over their efforts and their students’ learning resides outside themselves” (Haberman, 1987, p. 24). Haberman contended that the strategies and efforts that result from this factor are often the difference between the success and failure of an at-risk student.

A teacher’s response to authority is comprised of a critical set of behaviors that illustrate a teacher’s willingness to work in, with and around the school authority, doing whatever is required to ensure the success of the students. A teacher’s definition of at-risk

students and an understanding of how they become at-risk are critical to what a teacher thinks that he can do to teach effectively. A teacher who believes that students from a low socio-economic background are not as capable academically will not put forth the efforts required, or not respond appropriately, to ensure the success of these students. Teachers who are able to see the applications of generalizations, principles and research findings will be able to translate abstract concepts to their students in a concrete format. They will be able to create multiple examples from generalizations. Both of these techniques allow urban students to examine concepts at multiple levels and from various points of reference from within their lives, while gaining a more complete understanding of the concept.

### ***Guilford County's Use of the Haberman Interview Tool***

The Guilford County School (GCS) District has been a leader in using national models for interviewing. Most major models include both on line screeners and a follow-up face to face structured interview as does the Haberman instrument. GCS saw the purpose of selecting a model to be used as a district-wide interview tool was to provide a systematic solution that is research-based, aligned to core values and best practices, and statistically reliable and valid. Since principals hire up to 600 new teachers each year for GCS, the human resources department felt as though it was important to have an interviewing tool that best identifies the qualities and attributes needed to teach in an urban setting (Guilford County Schools, 2008).

Describing the process of GCS, district personnel are involved in screening and selection strategies which include licensing/transcript analysis, criminal records checks,

reference checks, on-line screening and interviews. As teacher applicants become aware of available positions within the county, they may submit an application on line, via the mail, or in person. They are then directed to take an on-line screener which is the initial screening phase of the Haberman process. Of the 10,500 GCS teacher applicants annually, 54% meet initial screening standards (i.e. licensing requirements, screening requirements). Upon successful completion and sufficient score on the on-line screener, applicants are then chosen to come in for a face to face interview. Face to face interviews are conducted by principals within the district. To support the common language and strengthen implementation, interviews are conducted in teams of two or three persons. These teams enable greater inter-rater reliability; greater opportunity to probe and examine candidate strengths and improvement needs, thus supporting more accurate interview results. Of the 54% of the teachers that meet the initial screening standards, approximately 20% of those are selected from the interview as prospective teacher candidates for GCS. Names of candidates who have met requirements are available for principal review. This process assisted principals in their final selection process by providing a pool of eligible candidates that have been interviewed and identified by fellow GCS principals.

In conclusion, school districts undertake a series of activities for getting new teachers in the classroom like recruiting qualified applicants, screening them according to evaluation criteria, hiring the most qualified, and placing them where their skills are most needed. Districts also try to evaluate, strengthen, and keep their new hires during the first years especially when they are likely to leave.

### **Teacher Retention**

According to the National Commission on Teaching and America's Future (Rebora, 2003), the United States did not actually have a shortage of qualified new teachers. The real school-staffing problem was retention. Qualified teacher shortages stemmed from teachers either leaving or moving from their jobs (Ingersoll, 1997). Approximately one third of America's new teachers left teaching sometime during their first three years of teaching; almost half left during the first five years (Rebora, 2003). Private schools have a higher turnover rate (12%-16%) than do public schools (8%-9%). Public and private schools that have enrollment under 300 have had higher turnover rates (17%) than do schools with larger enrollment, (9.3%). Public schools with half or more of their students receiving free lunch had a higher turnover rate (10%) than do schools with a lower concentration of students receiving free lunches, 8% (Ingersoll and Rossi, 1995).

According to human capital theory (Kirby & Grissmer, 1993) a decision to stay or leave a job or career was often based on how much a person had invested in it. An individual who stayed in a profession accumulated specific human capital that was relevant to that profession only. The greater the amount of specific human capital, the less likely the individual would consider leaving the profession. Therefore the more complex the initial training and the longer a person held a teaching position, the less likely he/she would see leaving as an option (Kirby & Grissmer, 1993).

*Teacher Magazine* (2002) described the average teacher as a 42-year-old, white female, with 15 years of teaching experience. This statement was supported by data

found in the 1993-94 *Schools and Staffing Survey (SASS): Selected Results* (National Center for Education Studies, 1995) and the *Schools and Staffing Survey: 1999-2000* (National Center for Educational Studies, 2002). SASS 1999-2000 reported percentages of teachers who stayed in teaching, disaggregated by age, for each state. In most cases teachers under 30 years of age were the most likely to leave teaching when compared with other age groups.

According to Andrew (1998), long-term studies showed a higher rate of entry and retention of teachers from programs without undergraduate majors in education. Human capital theory also supported the concept of retention being related to the length and complexity of the training (Kirby & Grissmer, 1993). Only 10% to 12% of the extended program graduates left the profession within five years as compared to twice that amount of four-year program graduates. Most of the four-year graduates who quit did so within the first two years. Self-selection based upon commitment may have explained this retention discrepancy. Those who were committed to teaching were more likely to stay in teaching longer than individuals without a commitment to teach (Miller, Brownell, & Smith, 1999). Some universities and even some companies had a way of determining the commitment level of those wanting to teach through selection tools or internships or on the job training.

The Urban Initiative Project of the George Washington University's Graduate School of Education and Human Development used the Urban Teacher Selection Instrument (UTSI) as part of the entrance requirements to their graduate education program. The program required a full-year internship, or student teaching experience, in

an urban setting as opposed to the common two placements in a 15-week period. The extended placement allowed student teachers to develop the skills and experiences that formed a strong foundation to support them in their early teaching careers. The Urban Initiative Project produced graduates with the characteristics identified by the UTSI. The four-year retention rate for the members of the first cohort of George Washington University program was 87.5%. After three years teaching, 73.7% of cohort two remained, and 100% of cohort three remained in teaching two years after graduating.

Murnane's study (as cited in Stevenson, Dantley, & Holcomb, 1999) relating to teacher attrition and retention indicated that older female teachers new to the profession tended to remain in teaching longer, while younger female teachers were more likely to leave. Elementary teachers had the lowest attrition rate, while chemistry and physics teachers had the highest. Teachers with high scores on qualifying or credentialing exams were more likely to leave earlier than were teachers whose scores were low or average. Teachers from large urban districts tended to have shorter teaching careers than did those who worked in smaller suburban districts. African American teachers were less likely to leave teaching than were European American teachers. When placed in diverse schools, teachers who are not committed to a multicultural competence, as needed in urban classrooms, reported that they did not enjoy their jobs and generally left teaching within the first two years (Haberman & Post, 1998).

There is a mismatch in the supply and demand of quality teachers. Many students preparing to become teachers are not majoring in subject areas that match the needs of schools and districts experiencing shortages. Even in states that produce enough new

teachers to meet overall demand, there remain acute shortages in mathematics, science, bilingual, and special education. Moreover, many teaching candidates prefer suburban districts. They are unwilling to teach in many urban and rural schools (Stevenson, 1999).

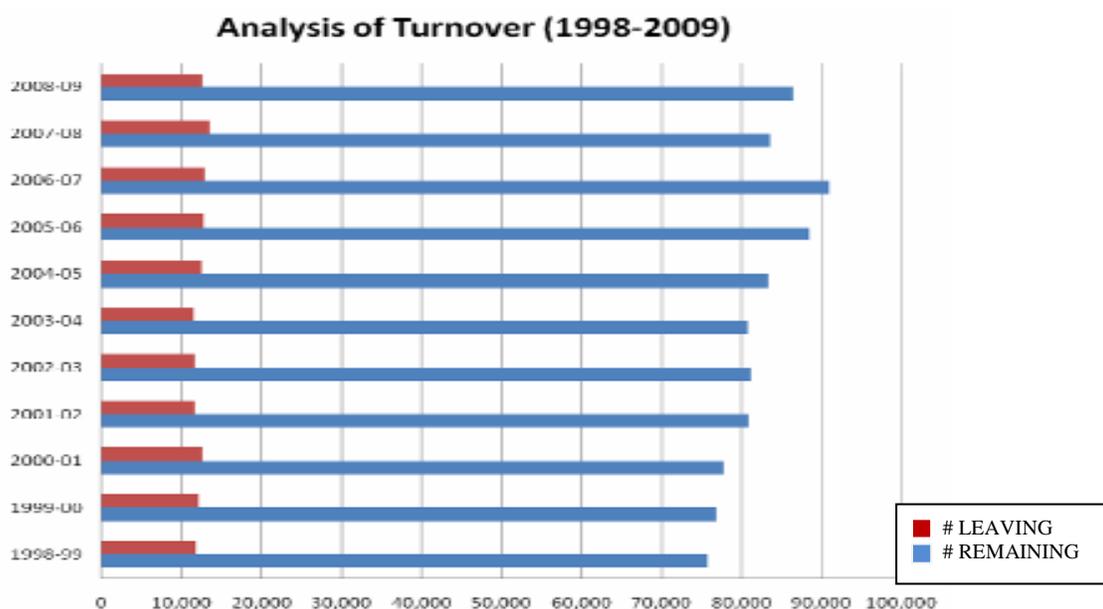
In many school districts because of the shortages of teachers, they now have various kinds of assistance programs directed at attracting and training teachers especially in mathematics (Urban Teacher Collaborative, 1996). The efforts put into these programs have resulted, for the most part, in increased teacher retention. Federal and state mandates have forced school districts to look into this issue as well. For example, the No Child Left Behind (NCLB) legislation calls for “highly qualified” teachers in all Title I schools where historically many of the math teachers are provisionally licensed or where the turnover rate for qualified teachers is skyrocketing. As a result, many schools are overflowing with first year teachers or lateral entry teachers with experience ranging from limited to none at all.

The need for a support system is even more critical for hard to find math teachers in our secondary schools. Research studies have shown that beginning teachers, in particular, benefit most from intense instructional support. Mathematics teachers especially need critical support during their initial years in the teaching profession. Some math teachers were lateral entry teachers who had previous work experience in corporate careers. Often teacher retention was very difficult, especially in low performing schools in urban areas (Stevenson, 1999).

In 2008-2009, North Carolina's average teacher turnover rate was nearly 13%, ranging from a high of 26% to a low of .75%. Turnover among teachers in low-

performing schools was substantially higher, with a low of 12% and a high of 57% (<http://www.ncpublicschools.org/newsroom/news/2009-10/20090903-01>).

Figure 1 shows the comparison of teacher turnovers in North Carolina over the past 10 years. It is apparent that the number of teachers leaving stayed the same or increased slightly until 2007-2008. Fewer North Carolina teachers left their jobs last year, according to a North Carolina Department of Public Instruction annual Teacher Turnover Report presented to the State Board of Education in September of 2009. According to the statewide report, local school systems had an average teacher turnover rate of 12.72% for 2008-09, slightly less than the 2007-08 rate of 13.85%. North Carolina's teacher turnover rate also remained lower than the 16.8% rate of turnover among all teachers in the United States.



North Carolina Department of Public Instruction (2009) Teacher Turnover Report.

**Figure 1. Teacher Turnovers in North Carolina from 1998-2009**

North Carolina put strategies in place to address teacher retention but the question remains, will these strategies impact retention at low-performing schools? The North Carolina Education Research Data Center at Duke University's Center for Child and Family Policy in 2006 researched this issue and the following summarizes three studies that address issues related to teacher retention. A study on the evaluation of \$1800 for teacher bonuses in North Carolina to certified math, science and special education teachers working in high poverty or academically failing public secondary schools was conducted. The study found that:

- (1) The bonus payment was sufficient to reduce mean turnover rates of the targeted teachers by 12 percent;
- (2) Responses to the program were concentrated among experienced teachers; and
- (3) In 2003-04, 17 percent of principals in schools with the program did not know their schools had ever been eligible.

Implications of the study indicated that:

- (1) Supplemental pay may have been a promising approach to retaining teachers in hard to staff subjects and schools; and
- (2) Greater efforts must be made to promote such programs.

A second study examined 272 hard-to-staff schools in North Carolina and looked specifically at teacher perceptions of the work environment. This study found that:

- (1) Minority, disadvantaged, and academically struggling students were more likely to be in hard-to-staff schools and less likely to have experienced, effective teachers;

- (2) In 2000-01, in hard-to-staff schools, 71 percent of students performed at grade level on End of Grade or End of Course tests, compared with 80 percent of students in other schools;
- (3) In hard-to-staff schools, 62 percent of the students were ethnic minorities, compared to 39 percent of the students in other schools;
- 4) In hard-to-staff schools, 47 percent of students were eligible for free/reduced price lunch compared to 35 percent of those in other schools;
- (5) Forty-two percent of hard-to-staff schools are middle schools, while only 18 percent of other schools were middle schools; and
- (6) Teachers in hard-to-staff schools were less satisfied with every aspect of the school environment than their peers.

The study stated that these findings indicate that:

- (1) Addressing working conditions will be essential to reducing teacher turnover;
- and
- (2) Efforts to reduce teacher turnover should target conditions in hard-to-staff schools.

A review of teacher retention from the SERVE Center at the University of North Carolina at Greensboro and the North Carolina Education Research Data Center at the Center for Child and Family Policy (December, 2006), including both quantitative and qualitative studies found:

- (1) The issue of retaining teachers was one of retaining quality teachers who positively influence student learning, not just retaining all teachers;

- (2) Teachers who feel effective with their students were more likely to stay;
- (3) Teachers in collaborative, collegial environments were more likely to stay;
- (4) Increased pay was positively associated with retention;
- (5) Turnover was highest among high poverty, high minority schools;
- (6) Teachers entering the classroom through Alternative Certification Programs were more likely to leave the classroom;
- (7) Teachers teaching out-of-field and teaching courses requiring many different preps have lower job satisfaction;
- (8) Late hiring and lack of information in the hiring process could negatively influence retention; and
- (9) Poor facilities were associated with increased turnover.

The review concluded that many factors contributed to increasing teacher retention, so single-pronged approaches would have much less chance of success.

### **Value Added Data as an Indicator of Teacher Effectiveness**

#### ***History***

To be fair to schools and to children, any instrument that is used to determine mastery on achievement tests must be “standardized,” uniformly administered, made equivalent, offered in different locations and years, and scores must be delivered in a manner that is fair to compare (Evergreen Freedom Foundation, 2004). In recent years, all 50 states have implemented accountability measures in response to increasing concerns about the quality of our American education. With an increased emphasis from the policy of No Child Left Behind (NCLB), states look at ways of assessing how

effective schools and teachers are doing in serving their students. A recent comparison of student achievement in mathematics and reading literacy at the eighth grade level found that student performance in the United States averaged below the mark in performance in 14 countries (National Conference of State Legislators, 2008). Why does this occur? Are there disparities that exist within our education system that creates this problem? As an accountability model, there have been three types of achievement tests given in the past. The traditional “norm-referenced” achievement tests were given to students and the results were expressed by percentile rank, comparing students to each other meaning that half of all students scored by definition “below standard” (Evergreen Freedom Foundation, 2004). Comparing students with one another has limited usefulness in determining a school’s effectiveness or a teacher’s effectiveness. Furthermore, a school’s scores most of the time reflect the socio-economic status of its students rather than the work of the teaching staff. Indicators such as race and ethnicity, economic status, geography and parental education may also affect the quality of education students are receiving. Students from economically disadvantaged backgrounds, urban environments, or racially diverse families typically do not perform as well on assessments when compared to their peers.

Criterion-referenced tests were popular among the educators who wanted to look at test in relation to actual achievement. When taking a criterion-referenced test, students are expected to demonstrate selected knowledge and skills, with scores reflecting their own standing in comparison to selected levels of achievement. In other words, this type of assessment compared a student with an established benchmark of achievement

(Evergreen Freedom Foundation, 2004). This again was very controversial for many reasons and one of them was the standards themselves. Some people thought that the standards were too high which caused students to fail and others believed that the standards did not reflect the content we really wanted students to know and understand (Evergreen Freedom Foundation, 2004). Additionally, to help close the "achievement gap," states began seeking more accountability from their schools. Earlier attempts at reform focused on school inputs and processes, such as complying with regulations and funding allocations. A 1995 report by the National Education Association (NEA), however, found that states were beginning to shift their focus towards student outcomes. States began adopting standards and tools to assess student performance with the expectation that all children can achieve a certain performance level. This trend in education reform became known as standards-based accountability. This is where the real issue started. A true assessment tool that evaluates students, teachers, and school progress was most needed. For this reason, value-added assessment was discovered which compared a student's current level of performance to his or her past level of performance. The policy of No Child Left Behind (NCLB) requires that states look at ways of assessing how effective schools and teachers are doing in serving their students. North Carolina has begun to look at value added assessments as means to determining school and teacher effectiveness.

### ***Definition and Purpose***

Value-added assessment is a method of analyzing and reporting student test results based on improvement (growth) in standardized test scores over two or more

points in time. According to Doran and Fleishman (2005), value added assessment was one technique that researchers identify effective and ineffective teachers and schools. This is different from the traditional methods of determining school effectiveness (including the adequate yearly progress under NCLB) in that value-added models do not look only at current levels of performance of children but it looks at measuring a child's improvement from one year to the next. This again is in contrast to analyzing test data to measure students against an absolute standard of achievement, or to rank them against each other, or to evaluate a school's performance. Value added assessment is not a different type of testing program. Standardized tests, whether they are normed or criterion referenced, are used to obtain the scores. The act of comparing the scores to the past scores is what distinguishes value added assessment. "The idea of value-added modeling is to level the playing field by using comparisons statistical procedures that allow direct comparisons between schools and teachers—even when whose schools are working with quite different populations" (Doran & Fleischman, 2005).

The Evergreen Freedom Foundation (2004) says that value-added assessment has the potential to provide many benefits when analyzing student scores:

**Focus:** Value-added assessment changes the focus of education statistics from quibbling over demographic factors to asking the essential question: How well are students progressing?

**Equitable comparison:** By focusing on student growth, value-added provides a way to recognize outstanding student growth accomplished by teachers. This is especially important in schools with high populations of learning disadvantaged students.

**Accountability:** Because value-added assessment provides results that are less tied to student demographics and more tied to teacher effectiveness, they provide a fairer accountability measure for schools and teachers.

**Diagnostic:** Value-added assessment alone cannot identify the cause of poor student achievement. But where the data is sufficiently detailed, it can identify where failures and successes are occurring, giving staff a starting place to begin asking questions and making data-driven decisions. The more extensive forms of value-added provide a gold mine of data for education research. (Evergreen Freedom Foundation, 2004, p. VA-3)

The Wisconsin Education Association Council (2004) concurred with the beliefs of the Evergreen Foundation (2004) and believed that value-added assessments are the following:

- They move the attention away from the “limitations” that demographic factors place on student learning (e.g. social class, parents’ education level, and race/ethnicity).
- They are based on the assumption that the teacher is the most important influence on student achievement. “Proponents argue that this is justified because value-added data show that teacher quality is the most important factor determining how much students learn.”
- Valued added assessments steer people away from focusing on the percent or number of students who reach an arbitrary standard. Instead it focuses at improvements (Allen, Galinat, Leedham, & Leverich, 2004).

In sum, states have moved toward this standards based system to measure the effectiveness on the academic gains of students from the beginning of the year to the end. North Carolina has begun to look at this process especially since the passage of the No Child Left Behind Act.

### *Challenges of Value Added Assessment*

Advocates of value-added modeling call value added assessment the single most promising development in this era of academic accountability—perhaps the best hope for improving America’s mediocre-to-poor public schools. One of the more prominent studies of valued added assessment was the 2006 Carnegie Report. This report raised the issue of value-added assessments and the challenge it causes educators. While the report, written by Barbara Stewart, clearly saw the value-added method having the potential of becoming a powerful tool in improving schools, especially struggling low-income, minority schools it also sees it as a mechanism for attracting the best teachers to those most challenging schools. One of the main challenges according to the Carnegie Report was evaluation. The belief was

that value-added modeling could return power and credit to the frontline of education: the teachers in the classroom. It could do away with subjective, partisan teacher evaluations. It could, for the first time, provide a clear way to figure out who are the best, the better, the average and the ineffective teachers. It could also offer a way for principals to know which teachers need help and which have the skills to help others. (Stewart, 2006, p. 13)

Adam Urbansky, a Rochester teacher’s union official, said that educators must be careful thinking that this model is a quick fix for successful schools. There is also danger in placing too much emphasis on the value-added scores of teachers when student achievement rates may also be the result of socioeconomic and other factors beyond a teachers’ control. The most die-hard advocates agreed that value-added modeling was useless if it is not acted on. Regardless of these thoughts, the most pressing question

arising today is what are the most effective teachers doing so well? The answer to this question, according to Martin Haberman, lies within the human being.

In November, 2005, Harold Doran and Steve Fleischman researched value added assessment and released an article in *Educational Leadership* entitling three challenges of this model. They were the complex statistical difficulty of the model, the test scores being vertically scaled, and the whether or not the model can actually be called a teacher effect. Many educators have expressed that some models are too statistically complex and difficult to understand (Darlington, 1997). It is true that this value added model has a complex statistical analysis component that is highly technical and complex. Doran and Fleischman believed that those that use the model should collaborate with professional organizations that have experience with methods such as this one.

Secondly, Doran and Fleischman (2005) believed that in order for value added modeling to work, the tests must be vertically scaled. This means that there is a process that connects the many tests given and places them on the same scale so that it is possible to measure growth over a period of time.

To connect different tests and measure student growth, designers of value-added models commonly assume that the curriculum in higher grades is nothing more than a harder version of that in the previous grade; in other words, 8<sup>th</sup> grade math is the same as 7<sup>th</sup> grade math, just more difficult. Therefore, one can measure a student's increase in math knowledge by measuring his or her academic growth over time. (Doran & Fleischman, 2005, p. 86)

Other researchers suggested that this year to year curricula discrepancy was major. To correlate one test to another over the course of a year may create errors in longitudinal analyses (Doran & Cohen, 2005).

Thirdly, as in the Carnegie report, Doran and Fleischman were concerned as to whether the estimate given from the model can be called a teacher effect. The difficulty they believed was how educators could extract out all of the other factors that may have influenced a teacher's score such as socio-economic status.

Similarly, a Policy Report (Pearson Education, Inc., 2004) on value added assessments denoted many issues and challenges to this approach. The most commonly cited concerns include the statistical methods used to account for bias from non-educational factors, some fundamental issues of scientific research, the effect of missing data, the limitation of subject areas that can be evaluated, and finally questions about the available research published. This research mentioned that one of the main issues from value-added data is that using a student's race or socio-economic status contains a bias that is not accounted for when given a standardized test.

Researchers believed that these non-educational biases may be attributed mistakenly to teachers with no way of determining the error. Ballou (2002) suggested that non-educational factors have an impact on the evaluation of teachers regardless of whether they try to control those variables or not. Another non-educational factor was the possibility that a student's aptitude played a role in his or her learning growth. Some believed (Kupermintz, 2003) that teachers who were assigned more challenging students were likely to be evaluated as less effective than other teachers who were equally effective but have students with high aptitudes for learning.

Another concern from this report was that the data did not tell which specific practices made a teacher's instruction more effective at providing students with

achievement growth. Some value added models used an average teacher from the district to compare others to and other models didn't use anything. This, according to this report, could have been a challenge and researchers must find an answer to this concern. Thirdly, not having data on individual students from previous achievement records seemed to be an issue. Some school districts that have a high mobility rate of student population are sometimes faced with incomplete records. For teachers of these students, how can they be evaluated when there is no data to determine whether a child has shown growth or not. Finally, NCLB requires testing for grades 3-8. Schools are only required to test in English language arts, mathematics, and fifth and eighth-grade science, for example in North Carolina. This means that only teachers of these subjects have value added data that they are held accountable for where as teachers in other subjects are not held to the same standard. Hence, the use of value added data for evaluative purposes may be of limited value for some districts.

In conclusion, the report clearly supported the use of a value-added model and did not see their findings to invalidate any of the studies but as an opportunity to advance and strengthen value added models through further research.

The Evergreen Freedom Foundation (2004), a vocal critic of public education, found there were several challenges in implementing value added assessments including-statistical issues, political challenges, time investment, and delay in data. A statistically sound model brought with it more variables that only those who are statically literate can understand but difficult for parents and educators. This was a problem when it came to analyzing teacher proficiency. Some districts were not able to examine data at the

teacher/classroom level only at the building level. In order for value added data to be useful, it must be examined at the teacher level because research shows that the primary influence on student growth comes from the quality of the teacher. The political challenges of determining which students should be included in the data poses an issue for some districts. When students have only been in schools for a short period of time or for schools that have a large number of students who are low achieving, should the expectation of those teachers and building level educators be different from those in other schools? This issue still remains a constant question and one that has no resolution as of yet. The bottom line is that all teachers are held to the same expectation because all students are expected to achieve at the same level. Another issue discussed was the issue of time. To take standardized tests required frequent testing times in order to plot student growth. Some people believed that too much time is spent on testing rather than on teaching.

Lastly, this report indicated a lapse between administering any test and receiving the results so that it makes a difference. Most times, since value-added data uses student data from the previous year along with data from the current year, it takes until the fall before the results are processed and analyzed. If this is the case, the value added data cannot be used with teacher evaluations since the data is not available at the time evaluations are conducted. Additionally, with the Tennessee Value Added system (TVAAS), teacher data was further delayed because three years' longitudinal data is required to ensure statistical accuracy which means that the data is unavailable for the first two years of teaching.

The Wisconsin Education Association Council (2004) addressed two main issues related to value-added assessments. They were based on a very narrow measure of student learning and the value added data was used to evaluate teachers, administrators, and schools. In most school districts, schools are asked to provide a wide range of services and only judged by the test scores. In doing so, judging a school only based on test scores will negatively impact the students, teachers, administrators, parents and even the public. The thought then became whether teachers were teaching the test or doing what is best for children. Secondly the report found the other issue was the use of value-added data to evaluate teachers, administrators, and schools. Just as the Evergreen Freedom Foundation (2004) mentioned earlier suggested that value-added data could be used to make a variety of personnel decisions whether good or bad, the (WEAC) believed this could be an issue as well.

Value added assessment is a complex issue itself and the challenges of value added data are quite similar in nature. There will always be issues of time, money, statistical analysis, and even the correlation between the standardized tests and curricula decisions. The common variable in all reports and possibly the most important is that student achievement will only increase if the teacher in the classroom is effective. Value-added assessment assumes that teachers are the most important factor affecting student learning, and the amount of value that a teacher adds can be measured (Allen et al., 2004).

*North Carolina: No Child Left Behind and Adequate Yearly Progress (AYP)*

Making states, districts, and schools accountable for the performance of their students has become a popular topic among policymakers and the education community. After decades of focusing on inputs such as funding levels, curriculum offerings and resources, policymakers are now emphasizing student learning and achievement outcomes as the means of gauging an education system. This trend is referred to as standards-based accountability system. A standards-based accountability system sets goals in the form of standards, assigns responsibilities for meeting those goals, and holds the system accountable for its performance. Standards need to be rigorous enough to challenge students, without being set so high that they are unachievable or discouraging. They are to be aligned with the curriculum so that the material being taught in the classrooms allows students to achieve the desired goals.

Under this type of system, the state's role changes from ensuring compliance with regulations, to providing incentives and offering technical assistance to build school capacity. State officials prescribe the outcomes, but the choices about instructional methods and practices are left for the district professional educators to decide. The No Child Left Behind Act of 2001 (NCLB), adopted this accountability system as the foundation for reform. Over the past years, states receiving federal funds for education were required to continue to implement standards-based accountability measures into their education system.

The value-added type of assessment is a key part of North Carolina's "ABC's of Public Education" accountability system. In 1996 the State Board passed the ABC's in

response to the School-Based Management and Accountability Program. This ABC system contains both a growth (value-added) and a performance standard. All elementary and middle schools have been evaluated using this system since 1996-1997 and high schools since 1997-1998 (Evergreen Freedom Foundation, 2004).

The ABC growth measure examines whether a school has met either expected growth or high growth standards compared to a year selected as a baseline. Performance measurers look at what percentage of students in a school have met a certain benchmark. Together the two measures are used for a comprehensive school recognition and classification program. Recognition is provided to schools when they have a high percentage of students meeting the performance standards, and cash awards are given when schools make expected or high growth. Schools with high growth are awarded \$1,500 per certified staff and \$500 per teacher assistant. Schools making expected growth are awarded \$750 per certified staff and \$375 per teacher assistant.

Those schools that do not make expected growth and have less than 50% of their students reaching the performance benchmark are classified as “low performing”. The State Board of Education assigns an assistance team to some of these schools and offers others assistance on a voluntary basis. Data from the Department of Public Instruction indicate that there are 2,467 public schools in North Carolina assigned an ABCs growth status for the 2008-09 school year. These include regular public schools spanning combinations of grades K-12, charter schools, alternative schools, and charter schools evaluated as alternative schools. Eighty and eight-tenths percent of these schools met

either their expected or their high growth standards. Table 2 shows the schools and status labels and recognitions as indicated above.

**Table 2**

***School Status Labels and Recognitions***

<i>Performance Level</i> Based on Percent of Students' Scores at or above Achievement Level III	<i>Academic Growth</i>	
	Schools Making Expected Growth or High Growth	Schools Making Less than Expected Growth
90% to 100%	Met AYP	Honor School of Excellence
	AYP Not Met	School of Excellence
80% to 89%	School of Distinction	
60% to 79%	School of Progress	
50% to 59%	Priority School	
Less than 50%	Priority School	Low Performing

Looking at Guilford County Schools, the ABCs Highlights include:

- In 2008-2009, 75.4 percent of Guilford County schools attained Expected Growth while 35.6 percent attained High Growth. In 2007-08, 79.7 percent attained Expected Growth, compared with 55.9 percent in 2007-08.
- Eight schools fell in the No Recognition category in 2009 and which was consistent with the 2008 school year compared with 10 in 2007 and 35 schools in 2006. These schools failed to meet Expected Growth but had 60-79 percent overall proficiency.
- Additionally, in 2008 nine schools were noted as low performing schools which increased slightly in 2009 to 10 schools (GCS website). Table 3 shows

the Guilford County results of the schools in the various state recognition categories.

**Table 3**

*Number of Schools in Various State Recognition Categories*

<b>Recognition</b>	<b>2008</b>	<b>2009</b>
Honor Schools of Excellence	1	10
Schools of Excellence	0	0
Schools of Distinction	14	17
Schools of Progress	41	44
No Recognition	8	8
Priority Schools	42	24
Low-Performing Schools	9	10
Expected Growth*	1	3

\*Only for alternative schools

As one can tell, we have a lot of work to do in North Carolina as well as in Guilford County. To do this will require the help of effective teachers in the classroom and leadership capacity at the administrative level.

*Guilford County's Use of Value Added Assessments*

Guilford County Schools has participated in value-added assessments for the past five years. They have worked directly with Dr. William Sanders and June Rivers who were the pioneers of the Tennessee Value Added 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, 2000a or b).

Currently, Guilford County Schools, in conjunction with SAS, Inc., can generate value added data for teachers who are the teacher of record for a tested subject in grades four through twelve. Value added data are generated once annually and are distributed to principals and teachers in the form of a single page PDF report per tested subject per eligible teacher. For teachers who have taught tested subjects in Guilford County Schools for multiple years, reports will reflect longitudinal data.

Upon receipt of value added data (VAD) reports from SAS, Inc., Guilford County Schools determines the process of distributing VAD reports from the central office to the individual teacher. Distribution of these reports has historically been the responsibility of different departments. In 2007, the Department of Human Resources sorted and distributed VAD reports by the Instructional Improvement Officer. Instructional Improvement Officers received these reports from Human Resources and then distributed reports to the appropriate principals. The responsibility of sharing VAD reports with teachers has consistently rested with the school principal. In 2008, VAD reports were shared with teachers in December; six months after students completed the EOG and EOC tests from which the VAD was generated. In previous years, VAD reports were distributed in the spring, nearly one full calendar year after students completed EOG and EOC tests.

When access is given to the district and to schools, principal may access several reports. Principals are able to view reports about their school, by teacher, by student, by grade, and by subject. When viewing this information, data is sorted by grade level and by the year. The number of students in the report is given along with their mean score and mean percentile. An average of students predicted mean score and predicted percentile score is noted. The school effect size is gathered by subtracting the mean student score and the mean predicted score. If the mean predicted score is lower than the actual student score, there will be a positive school effect size. If the mean predicted score is higher than the actual student score, there will be a negative school effect size. From that information, a correlation between the school and the state is determined as being below the average school in the state (Below), not detectably different from the average school in the state (NDD), or above the average school in the state (Above). A teacher report is similar in nature in that student data is readily available. Teacher data is disaggregated by the performance levels of students giving teachers the levels of students that they make the most growth with. For example, students are grouped into three levels, low, middle, or high. These levels indicate where students performed in that particular class. Data are then summarized from prior years indicating a positive, neutral, or negative value which is derived from the difference between the mean scores of what students were predicted to have scored and their actual mean scores. If students scored better than they were predicted to score, a teacher will have a positive teacher effect. If the predicted score of students is less than the actual mean score of students' scores, then the teacher effect value is negative. When teachers' scores are within the standard error of measure and

there is no difference between the mean predicted score and the mean actual score, then the teacher effect has no detectable difference which correlates to an average teacher. Since the value-added reports show the groups of students by level that a teacher is most effective with, principals use these data to determine which students a teacher will have the most success with. Teachers are then assigned to specific classes by the administrators based on these data.

There are four major functions of the value-added data plan in Guilford County. They are data management, communication, training, and application. In function 1, data management, value added data is linked to teachers through a social security or another identification number. Data follows teachers as they move among schools, grade levels, and tested subjects. Each teacher record contains VAD for each year a teacher taught a tested subject, the corresponding school name, and the corresponding course name for each data point. As a district, human resources has access to a database that supports the tracking of longitudinal data for each eligible teacher to use in making tenure decisions, placement decisions, scheduling decisions, and school improvement plans. In function 2, communication, a comprehensive communications plan seeks a clear and transparent understanding of VAD by all users of data. The human resource division provides a power point presentation on the use of VAD that contains frequently asked questions (FAQs) that are used in training principals and for principal use in training teachers. The power point also displays research on uses of VAD, VAD timelines, sample VAD reports with interpretations, a video of sample VAD principal-teacher conferences, and general resources. Additionally, email updates, information sessions, and newsletter updates are

regular forms of communication. Finally, there is a point of contact at both the district level for schools to seek additional information and assistance. In function three, training, principals are trained by the human resource division on the interpretation of the VAD reports, how and when to conference with a teacher, and the applications of VAD in the school improvement process. All of these are critical to the interpretation and application of VAD for evaluating teacher performance and student achievement. Function four, application, is the processes and structures that are in place to support the application of VAD for measuring teacher performance and student achievement. Application of VAD may include using data for scheduling teachers, placement of teachers at schools, tenure decisions, school improvement processes, and district improvement processes. The human resource department in the GCS school systems says that each school should each have a faculty member that is trained to act as a VAD expert on campus. This person can serve as a resource to the other faculty members as well as being a point-of-contact for district level VAD communications. Of the four identified functions, this area is the least developed across the country and the area of greatest potential improvement in GCS. In summary, it is evident that although there is a teacher shortage in our classrooms today, with the high accountability standards, the teacher in the classroom makes all the difference in the world in terms of student achievement. In order to reach the goals set forth by the state, we must have quality teachers in the classroom that will improve student learning and ways of determining which teachers have a positive influence on the progress of our students. As Harry Wong (2007) states, “quality teaching is the most critical means by which to improve student achievement or close the achievement gaps.”

In order to hire the best teachers, districts must find and attract quality teachers and to do this, Guilford County Schools has selected to use the Haberman Interview tool that is designed to identify teachers with high potential for success. Since this instrument helps to select teachers with a high potential for success, then what is the relationship between the score on the Haberman instrument and the tool used to determine teacher effectiveness in Guilford County Schools, value-added data? Additionally, does the value-added data of teachers look better based on their years of experience? This research study is designed to answer both of these questions.

## **CHAPTER III**

### **RESEARCH DESIGN AND METHODS**

#### **Purpose of Study**

The purpose of this study was to determine whether middle school mathematics teachers in the Guilford County Schools were producing student growth as determined by value added data. The measurements for the study included two independent variables of the Haberman Interview score and the SAS Educational Value Added Assessment System (EVAAS). This study estimated the impact of “effective” teachers on the students’ scores. Further, this study investigated if there was a relationship between effective teacher scores on the Haberman instrument and teacher value added scores in Guilford County Schools.

#### **Research Questions**

The main question investigated by this study examined whether there was a relationship of the total score of an effective teacher in middle school mathematics on the Haberman instrument and student growth on the End of Grade test scores. Two specific research questions guided the collection of data for this study:

Research Question 1: What was the possible relationship between teacher effect data and years of experience?

Research Question 2: What was the relationship between an effective teacher as measured by the total score on the Haberman instrument and teacher effect scores

as measured by SAS EVAAS, value-added data as noted by the End of Grade tests scores in middle school mathematics?

### **Research Approach**

Achievement scores of all students taught under specified teachers were matched against the value added data of the teachers. Bill Sanders, a Tennessee statistics professor, created a technique called value-added modeling, which uses a complicated statistical system to compute data. The system follows the academic achievements of individual students over several years and was not skewed by such factors as race, income or English proficiency. According to Stewart (2006),

Value-added modeling is a method of measuring student academic progress over time. It uses the annual standardized test scores for individual children, administered at the beginning and the end of the school year, to plot their progress in fundamental academic skills, and applies the results as a measure of the effectiveness of teachers and schools. (p. 1)

Value-added modeling works by attributing growth patterns higher or lower than average to teachers' skills—or the lack of them. When the method is used, many researchers, policymakers and educators are confident it provides strong evidence that teachers are the most important element in student achievement—no matter what the students' racial, class or economic backgrounds. A student who is assigned to a series of good teachers is more likely to achieve than one in a “better” school, in a more prosperous area, assigned several less effective teachers (Stewart, 2006).

This study was similar to the Tennessee Value-Added study (TVAAS) conducted by Bill Sanders. The approach that he used is called mixed-model methodology meaning

that he used quantitative data as well as used a qualitative approach to his data analysis. The model used for this dissertation was solely based on quantitative data gathered from the SAS Institute with the permission of Guilford County Schools. When teachers teach in a certain area for more than two years, value added data can be determined for them. For this study, the value added data that was retrieved from the SAS institute was given to the human resource department. The human resource department could then determine the middle school math teachers that had the value added data. Simultaneous to this gathering of data from the human resource department, middle school math teachers were selected based on their scores on the Haberman interview tool which uses the STAR teacher qualities for highly effective teachers in an urban district as denoted by Martin Haberman. The human resource department also kept the hiring information on all candidates that were hired and not hired. The Haberman scores were kept in a file in the human resource department. Both of these data files (Middle school value added data and the Haberman scores) were transferred to the GCS Accountability office using only social security numbers as identifiers. These two files were matched together from the office of accountability and then the data was dispersed to me. Only 39 teachers emerged from the data base for this study for the following reasons:

- Only 1,947 teachers were recruited using the Haberman instrument; if teachers were hired using another instrument their data was not retrieved;
- From that set, teachers who were not middle school math teachers were extracted;
- From that set, teachers who did not have value added data were extracted;

- If the middle school math teachers did not teach for 140 days of the 180 day school year, they would not have any value added data and therefore had to be removed.

This process left only 39 teachers to be a part of the data set. Additionally, the relationship between teacher effect data and years of experience was also examined. Included in both sets of data, the teachers years of experience were a part of the excel file. The result of this study was to determine if the sole selection criteria for placement of teachers in this urban district is positively correlated to student growth.

The matching of the teacher data received from SAS institute was completed by a member of the HR staff and then transferred to the research and accountability office. Once the office of research and accountability completed the analysis and retrieval of data, de-identified data of teachers and analysis was given to the researcher. At no time was the researcher able to determine which teachers were included in the study or their current assignment.

### **Description of Sample**

Achievement scores of all sixth-, seventh-, and eighth-grade students in the Guilford County School District who participated in the Math End of Grade tests in 2007 and 2008 were matched with teacher records in the Value Added database. The number of participants varied between 10 and 40. Teachers were then sorted into two cohorts based on their years of experience:

Cohort 1 – including teachers from 1 year of experience to 5 years

Cohort 2 – from 5.1 years of experience and above

Students of teachers for whom SAS EVAAS estimates were not available were excluded from the study as well as teachers for whom no SAS data were available. Teachers must have been employed in the Guilford County School system for at least one year in order to have had EVAAS data. Additionally, teachers were sorted within the cohort by their total scores on the Haberman Interview tool. Teachers who scored with a high average were placed in row C; those with a high score were placed appropriately in row B, and those teachers considered to be “Star” teachers were placed in row A. Table 4 shows an explanation of the cohorts:

**Table 4**

*Description of Data Included in the Study*

<b>Haberman Scores (Total)</b>	<b>Cohort 1 (1-5 years of experience)</b>	<b>Cohort 2 (<math>\geq 5.1</math> years of experience)</b>
22-33.5 High Average <b>Group 1</b>	6 <sup>th</sup> – 8 <sup>th</sup> grade Teacher Effect	6 <sup>th</sup> -8 <sup>th</sup> grade Teacher Effect
34-37.5 High <b>Group 2</b>	6 <sup>th</sup> – 8 <sup>th</sup> grade Teacher Effect	6 <sup>th</sup> – 8 <sup>th</sup> grade Teacher Effect
38 + Star <b>Group 3</b>	6 <sup>th</sup> – 8 <sup>th</sup> grade Teacher Effect	6 <sup>th</sup> – 8 <sup>th</sup> grade Teacher Effect

**Group 1** = Has difficulty understanding advice or explaining own ideas and plans. Engages in a variety of procedures with apparent confidence. These individuals start out as “Chargers” or High Average

**Group 2** = Is able to conceptualize about teaching and is sensitive to the purposes of activities, but has difficulty implementing ideas. These individuals start out as “Hesitant” or High.

**Group 3** = Is “with it.” Is able to implement advice; and/or act on plans. These individuals start out as “Stars.”

(Haberman, Stafford, & Dill, 1993).

These data were collected after the 2008 End of Grade scores had been verified and released by the Department of Instruction for the State of North Carolina.

Haberman scores are kept by the Human Resources department within the Guilford County School system. The district keeps a database of all applicants interviewed and their final Haberman score. Teachers who have been hired in the district specifically in middle school mathematics were examined, sorted by their Haberman scores and then by their years of experience by the Office of Research and Accountability of the Guilford County School district. Those teachers were then compared with the EVAAS database to ensure that these teachers had growth data. The mathematics teacher effect data was merged into the Value Added database kept by SAS Institute. The database provided the basis for the district's estimation of district, school and teacher effects. As stated before, teachers for whom no SAS EVAAS data were available were excluded from the study.

### **Methodology and Data Analysis Plan**

This study was similar to the Tennessee Value-Added study (TVAAS) conducted by Bill Sanders, a Tennessee statistics professor, in the 1980's. Bill Sanders looked at each student's performance against his or her own past performance. Sanders used a mixed-model methodology. His method used a complicated statistical system to compute data. He followed the academic achievements of individual students over several years and, according to some of his staff, was not skewed by such factors as race, income or English proficiency. Although his study looked at student progress over a period of time

and attributed growth patterns higher or lower than average to a teacher's skills or lack of them, there was no correlation to any particular hiring instrument for effective teachers.

This study examined whether the teachers hired within the GCS district had positive value added data since the Haberman instrument had been used to hire teachers. Value added data compares students' actual growth with their predicted growth. The difference is called teacher effect.

The teacher effect is a statistical calculation that provides an estimate of a teacher's influence on a students' academic progress. This calculation is a complex process that compares the average performance of each teacher's students to the system performance (Kupermintz, 2003). The calculation of a teacher effect score is a function of the difference between the students observed scores and their expected scores. Students' observed scores are their actual performance scores in scale score points on the End of Grade tests. Students expected scores or predicted scores are a calculation of students past performance on End of Grade tests which are averaged to predict a students' growth if they were taught by an average teacher. To reach a teachers' predicted score, a group of students' individual predicted scores are averaged to reach a teachers' mean predicted score.

If the difference of the mean predicted score and the mean observed score is positive, then the teacher report indicates that a teacher has made significant gains with students based on their end of grade scores. If the difference is negative, then the teacher report will indicate that the teacher has not made any gains with the students taught and is noted as being ineffective.

### ***End of Grade Tests: Mathematics***

When students take the End of Grade tests at the end of the year, they receive developmental scale scores based on their performance on the test. The North Carolina End-of-Grade Test of Mathematics, which is designed to meet the need for higher standards of achievement for North Carolina students, assesses grade-level knowledge and skills established by the North Carolina *Standard Course of Study*. The *Standard Course of Study* identifies the following seven strands for mathematics: numeration, geometry, patterns and pre-algebra, measurement, problem-solving, data analysis and statistics, and computation. The test contains both an applications section and a computation section with a total of 80 multiple-choice questions at each grade level. The two parts of the test are combined to provide one total score. The applications section of the test examines a student's ability to solve problems and explain mathematical processes. Students may use calculators and rulers for this part of the test. The computation portion of the test examines a students' ability to complete computations without technical assistance.

### ***Developmental Scale Scores***

Student performance in mathematics is reported in a variety of contexts, including developmental scale scores and achievement levels. Group and subgroup scores are also reported. Although the mathematics tests were developed as accountability tests, curriculum information is provided on the goal summary reports. In addition, developmental scale scores depict the amount of growth that actually occurs in student achievement over time. These scores can be used to compare the expected rate of growth

to actual student or group performance as the student(s) move from grade 3 to grade 8. Achievement levels show how a student or group of students performed in relation to predetermined grade-level standards set using teacher judgment linked to student performance.

### ***Teacher Effect Data and Teacher Years of Experience***

To address the first research question, the possible relationship between teacher effect data and years of experience, teacher effect data was examined more extensively with the years of experience of the middle school math teachers. Teacher identifiers were used to ascertain teacher effectiveness data and years of experience in the Guilford County School system. Years of experience were categorized as the following: Cohort 1: 1 year - 5 years of experience; and Cohort 2: > 5 years of experience. To understand this analysis, a multiple discriminate analysis was used to see if there was a difference in years of experience and teacher effect data.

For the purposes of this study, Haberman data was retrieved from the human resource office and was sent directly to the research and accountability office electronically from a HR specialist. The Haberman file data (excel spreadsheet) sent included the teachers first and last name, social security number, highest and type of degree earned, years of experience, the Haberman score, the current teaching assignment amongst other data that was not important for this study. The data were received by the Chief of the Accountability department within the Guilford County School system. The accountability office had teacher value added data from the End of Grade test scores in a file as well. This file contained the teacher name, social security number, school name,

their 2007 assignment, 2008 assignment and their value added score. These files were merged into one file from the office of accountability matching social security numbers. After all analyses were conducted by the research department, the file was then de-identified by the office of accountability by deleting the teacher name, social security number and location to only allow the researcher to see the number of teachers with valid data. Teacher identifiers were used so that data could not be linked to a particular teacher. No specific information regarding teacher names or value added data for specific teachers were given to the researcher. The researcher did not see any of the files since they contained personally identifiable information.

### **Haberman Scores**

The *Haberman “Star” Teacher Selection Interview* measures and predicts candidates’ levels of persistence, organization and planning, value of student learning, theory of practice, ability to connect with and teach at-risk students, approach to students, survival in bureaucracy, teacher success, and fallibility. Of these qualities, one that has been found to be the most powerful indicator of success in an urban setting was the “approach to working with high-risk children” (Haberman, 1987, p. 47). Haberman did not share how his data are collected and analyzed, how his conclusions were drawn from the data, nor did he provide an operational definition of “urban school.”

According to Sachs (2004), Haberman’s studies “reported findings as correlates to effective teaching without gathering data from a control group of ineffective teachers leading to the unintentional generalization that these studies had identified attributes

associated with teacher effectiveness” ( p. 180). In fact, Sachs discovered that effective and ineffective teachers responded similarly on the questionnaire.

Teacher total Haberman scores were pulled from the database from Guilford County Schools by one of the Human Resources specialist and placed into one file. The specialist then specifically looked at middle school math teachers who had valid value added data with the GCS system and placed them in another file. Again, both files contained the years of experience of the teachers. These two data files were then sent to the Office of Research and Accountability using only social security numbers. These social security numbers were then changed to numbers (of which had no meaning) for the purposes of this research. It is to be noted that for the purposes of this research, only the total scores of the Haberman instrument was used was not disaggregated by teacher responses to certain questions. Additionally, at no point were any student or teacher names given to the researcher for this study in any part of the process. This would have been extremely difficult since the files were never seen by the researcher. Participant consent was not necessary for this study because no identifiable data was being used.

#### ***Teacher Effect Data with Years of Experience Compared to Haberman Scores***

To determine the relationship between teacher effectiveness and the Haberman scores, the second research question, value-added data as noted by the End of Grade tests scores in middle school mathematics, was examined to see if there is a greater relationship between teachers with higher Haberman scores than that of teachers with a high average of scores. Teacher identifiers (social security numbers) were used to ascertain teacher value added data and years of experience which were the same

identifiers used for ascertaining Haberman scores. Haberman scores were categorized by the following: Total score of 22-33.5 (High Average) – Group 1; total score of 34-37.5 (High) – Group 2; total score of 38+ (Star) – Group 3.

A matching of teacher effect data with years of experience to the Haberman scores of these teachers was done by a member of the Human Resources staff working specifically with the Haberman interview technique. Teachers were then categorized by their years of experience teaching. Teachers were placed into 2 cohorts: Cohort 1 included the teachers who had between 1 and 5 years of experience and Cohort 2 included teachers who were more than 5 years of experience. An analysis was then conducted to determine teacher effect scores by years of experience. Additionally, the Haberman score file was then analyzed with the teacher effect scores. Teachers were once again classified into 3 groups depending on their score on the Haberman interview tool and their teacher effect mean average was analyzed. From this, a univariate analysis of variance or ANOVA was conducted which combined the Haberman scores, the years of experience, and the teacher effect data file. The office of accountability combined all of the data files to determine which teachers could be placed in which cohort and group. Teachers with experience between 1 and 5 years and had an average score on the Haberman interview tool between 22 and 33.5 were placed in Group 1/Cohort 1 and so on. Additionally, a split factorial analysis was conducted to check the validity of the dependent variables (Haberman score and years of experience) to make sure that they were really related to each other as opposed to other indicators. Throughout all of the

retrieval of data and analysis of data, the researcher received no personally identifiable information from the Human Resources Department or from the Office of Accountability.

### **Limitations**

There may be other factors that determine whether a teacher is effective other than the use of the Haberman score. Additionally, this study only looked at a representative group of middle school math teachers with the Guilford County school system. Other teachers who may have taken the Haberman interview to be hired were not included in this study.

There are more than thirty nine middle school math teachers that may have valid value added, but they were not used because they were not hired using the Haberman instrument.

In terms of valid data, value added data comes from students that have valid standardized mathematics test data. These students must have been with the teacher for at least 140 days of instruction. There may have been other students that the teacher may have taught, but those students were not a part of the teacher value added data.

Additionally, students that may have been in Algebra and Geometry at the middle school level were included in this data since they also took the standardized math end of grade test for their grade level.

Since value added data is correlated to the end of grade tests scores, when teachers are given their diagnostic reports, the reports identify patterns or trends of progress among students at three different achievement levels. These achievement levels (low, middle and high) indicate which group of students a teacher showed the most

progress with or the least progress within the class taught. Although the information is worthwhile for determining which group of students the teacher has the most effect on, the achievement levels do not correlate to the levels of the end of grade tests (levels I, II, III, or IV).

It is noted that the sample of teachers used in this study was not compared with a random sample of teachers using the same criteria, so one must be cautious when looking at how this data can be used in a larger sampling of teachers.

This study was very unique to the Guilford County school system and therefore, one cannot conclude that these results could be generalized to other school districts.

### **Validity of Research**

Validity is the essential consideration in the evaluation of the uses and interpretations of any research. There are many variables that may determine whether a teacher is effective and there are many characteristics of an effective teacher which differ based on research studies. When looking at student success, that can also be determined using a multitude of information. There are questions that may arise as to the validity of the data in that there are other variances that may determine the effectiveness of a teacher. Do student characteristics or socio-economic backgrounds account for certain variability? Do other school or community variables play a part with teacher effects? Such questions drive this validity investigation. While we know that there are other characteristics that determine the effectiveness of a teacher, this study will explore the interpretation of the TVAAS findings on teacher effects. Teacher effects differ between-teacher variability in the average test score gains of their students. This variability may

arise for different reasons, some of which directly associated with teacher effectiveness, but others may reflect the context in which teaching occurs or the qualities of the specific group of students being taught. I used the TVAAS value-added information understanding that these alternative explanations when contemplating the consequences, intended and un-intended. These data were not used against a teacher for evaluation purposes. Darling-Hammond (1997) says, "No person should be evaluated for high-stakes decisions based on statistical assumptions rather than on actual information" (p. 255). Policy makers should consider whether minimum requirements for statistical accuracy should be set before information can be employed in personnel and policy decisions.

For the purposes of this study, student demographics, socio-economic background, race, classroom management, relationships between the teacher and the student, and history of former family members who were educators was not used as the determinate in student achievement. This study used the TVAAS value-added data as indication of student success for teachers.

### **Definition of Terms**

The following terms will be found in the remaining chapters of this dissertation. To ensure that an understanding is obtained from the reading, definitions of the terms have been provided below.

1. *Developmental Scale Scores*: a measure to determine students' growth on End of Grade Tests (Sanford, 1996).

2. ***District Gain***: The average gain experienced by the district in that year and subject as compared to the reference gain.
3. ***SAS Educational Value Added Assessment System (EVAAS)***: Process that is very similar in nature to the TVAAS but executed by SAS in Cary, North Carolina.
4. ***Effectiveness***: determined by students experiencing one year's worth of growth within a given year.
5. ***Effect Standard Error***: The standard error of the teacher effect, a measure of uncertainty in the estimated teacher effect.
6. ***End of Grade Tests***: Standardized criterion referenced tests given to students in grades 3-8 in North Carolina assessing grade level knowledge and skills established by the North Carolina Standard Course of Study.
7. ***Mean Student Score***: The average of this teacher's students' observed test scores.
8. ***Mean Score %tile***: The observed percentile rank of the teacher's average student, relative to all students in the district who took the test.
9. ***Mean Predictive Score***: The average of this teacher's students' predictive scores. A student's predictive score is an expected score based on performance on previously taken tests, assuming the student has an average teacher.
10. ***Predictive Score & percentile***: The predictive percentile rank of the teacher's average student, relative to all students in the district who took the test.

11. **Reference Gain:** Number representing what would be one year's growth
12. **Teacher vs. District Average:** The teacher's value-added amount detectably above or below the district average, that is, the magnitude at least two times the standard error. (NDD means no detectable difference from average teacher in the district).
13. **Teacher Effect Data:** The teachers estimated value-added amount, in scale score points, relative to other teachers in the district.
14. **Teacher Gain:** The gain experienced by the students taught by a teacher.
15. **Tennessee Value Added Assessment System:** 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 mixed model methodology which allows all available student data to be included regardless of the degree of missing information (Sanders et al., 1997).
16. **Value Added Data:** These scores refer to the effectiveness of a teacher and a district.
17. **Value-Added:** Value added is a statistical way to analyze test data to determine the influence of teachers, schools, and districts on student learning.

In summary, this final analysis of this study will help to determine if the current interview tool in the Guilford County school district is assisting the HR department in finding effective middle school math teachers. These effective teachers are noted as those

that are producing growth with students by what the system determines as a researched based model for determining growth, value added data.

## **CHAPTER IV**

### **RESULTS**

#### **Introduction**

The purpose of this study was to examine the teachers hired within Guilford County in the area of middle school mathematics and to determine whether these teachers produced student growth on standardized achievement tests as determined by value added data. The measurements for the study included two independent variables: the Haberman Interview score and the SAS Educational Value Added Assessment System (EVAAS).

This study determined whether there was a correlation between the total score of a middle school mathematics teacher on the Haberman instrument and student growth on the End of Grade test scores as measured by teacher effect scores. Two specific research questions were examined:

1. Is there a statistically significant relationship between teacher effect data and years of experience?
2. Is there a statistically significant relationship between an effective teacher as measured by the total score on the Haberman instrument and teacher effect scores as measured by SAS EVAAS, value-added data as noted by the End of Grade test scores in middle school mathematics?

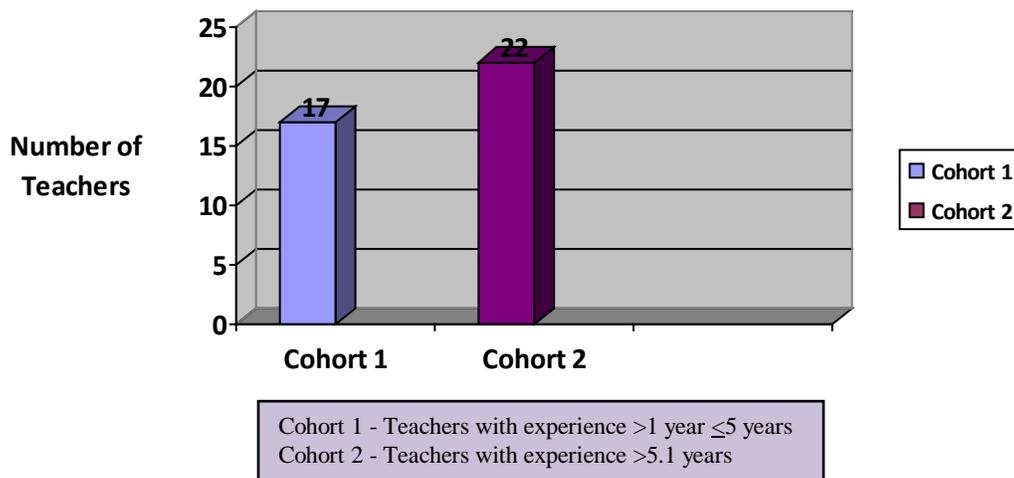
### **Demographic Data**

This section describes the demographics of the teachers who were included in this study. This study examined middle school math teachers who were hired using the Haberman interview instrument and had been teaching math for more than one year in the Guilford County School district. These teachers would have had value added data that was used to determine their effectiveness as a math teacher with the students they taught.

Teachers were placed into two categories regarding their years of experience. Cohort 1 contained teachers with at least one year of experience up to five years. Cohort 2 contained teachers with more than five years of experience. According to Haberman, a teacher's years of experience has a direct correlation to student academic growth (Haberman & Post, 1998). Teachers were also placed into three groups based upon their Haberman score. Haberman group 1 contained teachers who total score was between 22 and 33.5 points. Haberman group 2 were teachers whose total score was between 34 and 37.5 and group 3 were those whose score was 38 and above. The total score used in this study is the total score from the Haberman interview instrument. Haberman stated that only teachers who possess a particular set of attributes and ideology could successfully offer urban students the multicultural curriculum and support that they need (Haberman & Post, 1998). He indicated that a student's academic growth increased at a much faster rate when taught by a teacher who had a higher Haberman score.

Figure 2 shows the total population of the middle school mathematics teachers who had Haberman scores by years of experience for the 2007-08 school years. The

cohort labeled 1 indicates teachers who had between 1 year of experience and 5 years of experience. The cohort labeled as 2 are teachers with greater than 5 years of experience.

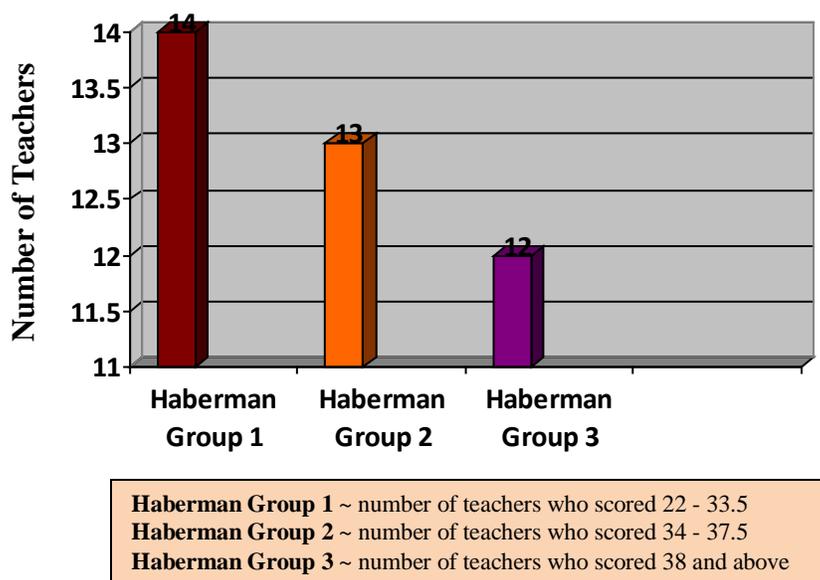


**Figure 2. Number of Sixth- to Eighth-Grade Teachers by Years of Experience**

The data show that 44% of the teachers (17) had between 1 and 5 years of teaching experience and 56% had more than five years of teaching experience (22 teachers).

Figure 3 shows the total population of the 39 mathematics teachers by their Haberman score. The data is arranged so that Cohort 1 reflects teachers whose scores were between 22 and 33.5. Cohort 2 are teachers who had Haberman scores from 34-37.5, and Cohort 3 are teachers whose Haberman scores were 38 or greater. The data show that 36% of the math teachers (Haberman Group 1) can be categorized as chargers (Haberman et al., 1993). These teachers may have difficulty understanding advice or explaining their own ideas and plans. Thirty-three percent of the teachers can be categorized as hesitant (Haberman Group 2) (Haberman et al., 1993). Hesitant teachers

are able to conceptualize about teaching and are sensitive to the purposes of activities, but have difficulty implementing ideas. The remaining 30% of teachers (Haberman Group 3) are considered “stars” (Haberman et al., 1993). Stars are able to implement advice and/or act on plans, according to Haberman.



**Figure 3. Middle School Teacher Haberman Scores Based on Score Levels**

Depending on the category that a teacher may fall into when interviewed, if principals knew the teacher’s final Haberman interview score, it is possible that they will be able to help the teacher in the areas that are weak. For example, if a teacher scores a 25 on the interview and is considered as a “charger,” then the principal may want to work with this teacher on explaining or expressing ideas, since they appear to be confident but have difficulty explaining their own ideas. Likewise, if a teacher falls into the “hesitant” category, the teacher may have great ideas, but the principal may help the teacher to implement them and with fidelity. Often teachers know what to do but putting the talk

into action is sometimes a difficult task. Lastly, if the teacher falls into the “Stars” category then the principal may want to use them as teacher leaders or in other roles that may assist other teachers.

When teachers are selected using the Haberman interview tool and are placed in a particular category based on the final score, principals should receive this information in order to help provide a plan of action to increase teacher effectiveness. Informing the principals of this upon hiring would be beneficial to all.

Based on the data of these three groups, there was not a significant difference in the numbers of teachers within each area based on their scores. With such a small sample of data, the numbers of teachers in each Haberman group only had a difference of one teacher.

### ***Teacher Effect Scores and Years of Experience***

Once teachers were categorized by their years of experience, another analysis was conducted to determine the teacher effect scores by years of experience. As mentioned in Chapter 3, teacher effect is a statistical calculation that provides an estimate of a teacher’s influence on a students’ academic progress. Table 5 shows that teachers with five years or less of experience have a mean teacher effect score of 2.808 and teachers with more than five years of experience have a mean teacher score of 2.903. Ideally, principals would want teachers who have a mean teacher effect score greater than 3 (Holcombe, 2008). The higher the teachers’ score (in scale score points), the more positive the effect. These data indicate that there is no significant difference in the mean average effect scores of teachers when taking into consideration their years of experience.

**Table 5***Teachers' Years of Experience and 2008 Mean Teacher Effect Scores Estimated**Marginal Means*

yexpgp3	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	2.808	.381	2.032	3.584
2	2.903	.325	2.242	3.563

Additionally, when looking at the mean teacher effect scores of teachers with less than 5 years of experience, the teachers mean scores ranged from 2.032 to a high of 3.584 with an average of 2.808. The same model is true for teachers with five or more years of experience. These teachers mean scores ranged from 2.242 to 3.563 with an average of 2.903.

It appears that the number of years of experience a teacher has no correlation to their mean teacher effect score. The mean scores are so close that this particular group of teachers seems to have the same effect on student growth. Since principals have access to teacher value added scores, in addition to reviewing the number of years of experience a teacher has, this data indicates that it might be feasible to look at their teacher effect scores as well. A second year teacher may be having the same effect on student growth that a ten year teacher is having. The use of teacher effect data is helpful in determining the strengths and weaknesses of subject areas and groups of students that a teacher has.

### ***Teacher Haberman Scores and Teacher Effect Data***

Another analysis was conducted of teachers categorized by Haberman scores with their teacher effect scores. Table 6 shows an analysis of teacher Haberman scores in comparison to teacher effect scores. The Haberman teachers were grouped into three categories. Haberman group 1 consisted of teachers that had a total Haberman score between 22 and 33.5 at the end of the interview. Haberman group 2 consisted of teachers who scored at least 34 points but not more than 37.5 points at the end of the interview. Haberman group 3 were teachers that scored 38 or above at the end of the interview. Teachers within Haberman group 3 had the greatest effect on student achievement as indicated by a mean score of 3.125. Teachers in Haberman group 1 had the least impact on student achievement as indicated by a mean score of 2.354. Although the Haberman group 3 mean is slightly above the Haberman Group 2 mean, the difference is not significant (see Figure 4).

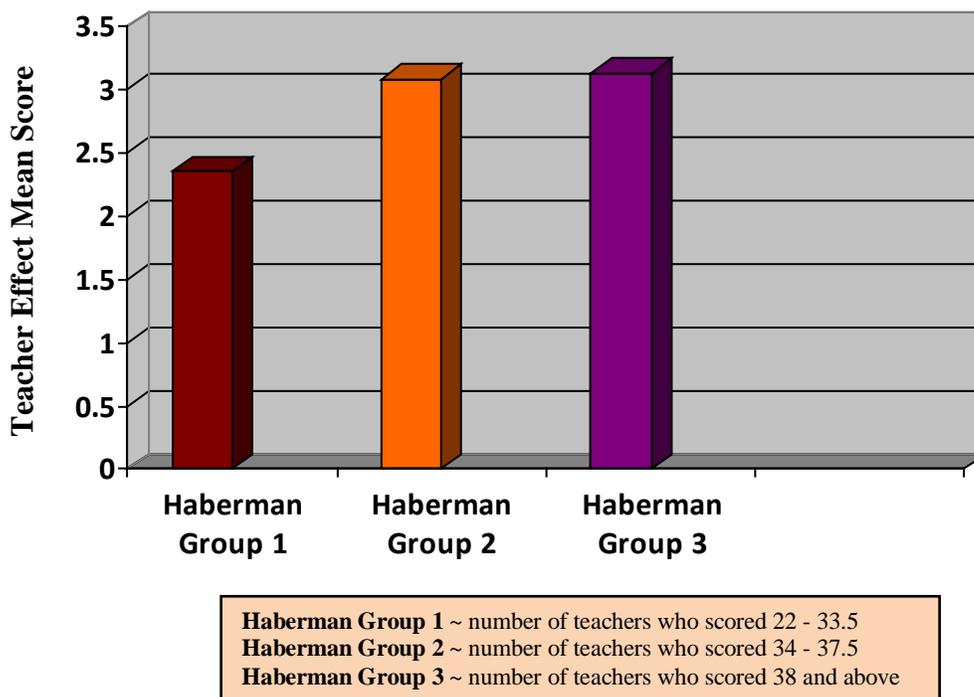
**Table 6**

#### ***Analysis of Teacher Effect Data by Haberman Scores for Sixth- to Eighth-Grade***

##### ***Teachers***

##### **Dependent Variable: Math 08**

habgp3	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	2.354	.407	1.525	3.183
2	3.088	.430	2.212	3.963
3	3.125	.462	2.185	4.065



**Figure 4. Mean of Teacher Effect Data by Haberman Scores**

These data indicate that with this group of 39 teachers, their Haberman scores were in direct correlation to their teacher effect scores; however, the difference is not significant enough to indicate that the higher the Haberman score the more effect a teacher would have on student growth. This implies that principals should consider the Haberman score as one indicator of teacher performance, but not as the sole indicator.

#### ***Teacher Effect Data by Years of Experience and Haberman Scores***

A univariate analysis of variance or ANOVA as noted in Table 7 was conducted to determine the relationship between teacher's years of experience, Haberman score, and the growth that a student experienced. The ANOVA analysis model is a way of examining the association between two variables. It is a way of looking at how one or two variables may affect a single outcome variable, which in this case are student

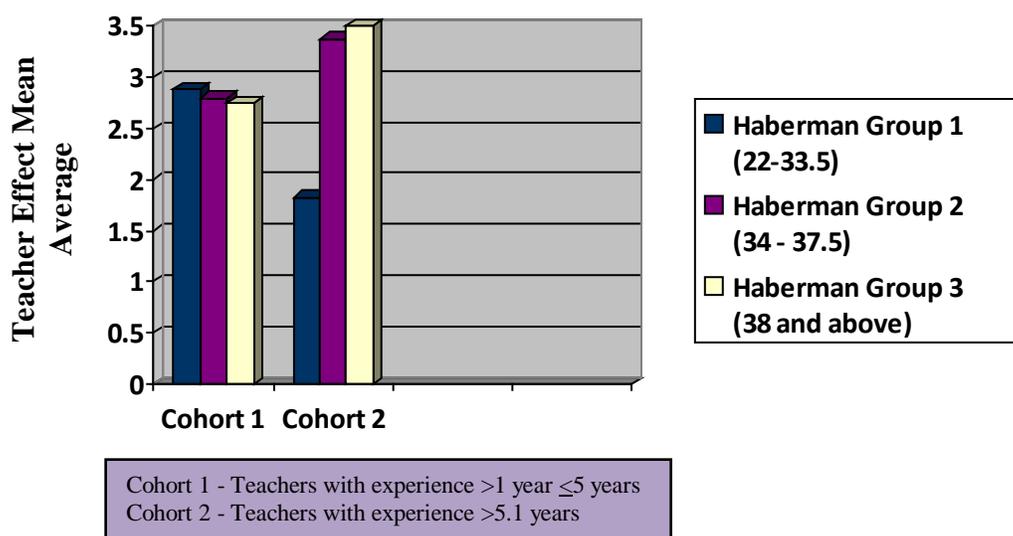
achievement scores. In this study, an analysis was conducted to determine a relationship between two variables (a teacher's years of experience and Haberman scores) and a single outcome variable (teacher effect data). In table 7 below, the analysis began with placing teachers in two cohorts based on years of experience. Cohort 1 denotes teachers that had 1-5 years of experience and cohort 2 contains those teachers with more than 5 years of experience. Haberman states that a teacher is considered average if students mean growth math scores are 3 or below (Holcombe, 2008). From each group of teachers (based on years of experience) a further analysis was conducted and teachers were placed in the three Haberman groups based on their interview scores. From this analysis the teacher effect scores from each of the groups could be determined.

**Table 7**

***Analysis of Teacher Effect Data by Years of Experience and Haberman Scores for Sixth- to Eighth-Grade Teachers***

Years of Experience	Haberman Group	Mean Average
Cohort 1 (Teachers with experience >1 year ≤5 years)	1	2.88
	2	2.80
	3	2.75
Cohort 2 (Teachers with experience >5.1 years)	1	1.83
	2	3.38
	3	3.50

Teachers who have more than five years of experience and an average Haberman score of 34-37.5 points had a teacher effect mean equal to 3.38. A mean average score of 4 or more indicates that students are experiencing above average growth. Teachers who have more than five years of experience and a high Haberman score of 38 or above, (Group 3); have a teacher effect mean equal to 3.50. Cohort 1 has an inverse relationship to the Haberman scores. For example, the higher your Haberman score, the lower your teacher effect. Cohort 2 has a positive correlation in relation to the Haberman scores. The teacher effect scores of teachers who have more than 5 years of experience is positively correlated to the Haberman group (see Figure 5).



**Figure 5. Analysis of Teacher Effect Data by Years of Experience and Haberman Scores for Sixth- to Eighth-Grade Teachers**

When reviewing a teacher effect report as noted in Appendix B, teachers are not only given their teacher effect score, but given a level score as well. This level number serves as an indicator of the teacher's level of effect compared to other teachers in the

district. The higher the level, the more effective the teacher is compared to others in the district. A level 3 indicates that a teacher is as effective as the average teacher who taught the same course in the district (Holcombe, 2008).

Based on the data above, teachers with the higher Haberman scores seems to be the teachers with more than 5 years of experience. However, this data does not take into account whether these teachers were former teachers in other districts. It could be concluded that if teachers have more experience, their Haberman scores are higher. The data clearly shows that teachers with less years of experience have a lower Haberman score. Although the data above shows teachers with more experience may have higher Haberman scores, questions still remain as to whether there is a relationship between a teachers years of experience and their teacher effect scores as well as a correlation between the scores on the Haberman instrument and a teachers' effect scores for middle school mathematics.

The above statistical analyses were used to help answer the research questions below. An additional analysis was conducted to validate the relationship between the dependent variables to ensure accurate results.

### **Research Question 1**

The first null hypothesis was tested to determine:

Is there a statistically significant relationship between teacher effect data and years of experience?

An analysis, as presented in Table 5, was conducted to examine the relationship between teacher effect data and years of experience for the middle school math teachers.

To begin the analysis, teachers in this study were placed into two groups based on the number of years they had been teaching. From that, teacher effect data of each teacher was aligned to the appropriate group and then averaged. After reviewing the averages or mean of the teacher effect scores for each group of teachers, there was no difference in the effect scores when compared to the years of experience. If there was a large difference in the teacher effect mean scores based on years of experience, one could argue that years of experience play a major role in determining how effective a teacher may be when looking at value added data.

### **Research Question 2**

The second null hypothesis was tested to determine:

Is there a statistically significant relationship between an effective teacher as measured by the total score on the Haberman instrument and teacher effect scores as measured by SAS EVAAS, value-added data as noted by the End of Grade test scores in middle school mathematics?

The first analysis as noted in Table 5 examined teacher effect data by teacher's years of experience. The mean scores of the two groups did not have a significant difference (.1). In other words, a teacher's years of experience had no correlation to the teacher effect score.

Table 6 depicts a second analysis which examined the group of teachers based on their Haberman interview scores and their teacher effect scores. Haberman scores were retrieved from an HR representative and sent to the office of research and accountability. These scores were grouped into 3 groups to denote the difference in the scores on the

interview tool. These scores were then matched against the individual teacher's teacher effect scores by the research department and averaged for each group. These findings confirm that the Haberman group 3 (teachers who scored 38 or above) teachers' mean average is greater than those teachers who scored lower on the Haberman interview.

To corroborate these findings, another analysis was conducted to examine a teachers' years of experience along with Haberman scores to help determine if these two factors had any relationship to one another. The analysis as noted in table 8 below represents the results of the split factorial analysis that answers both of the research questions. This type of analysis helps to check whether all of the indicators (years of experience and Haberman score) used as a dependent variable are related to each other and not to other indicators that are supposed to measure other dimensions. The degree of freedom (denoted as Df below) shows the number of independent pieces of information that goes into the estimate of each parameter. For example, the Df for the years of experience is 1, which means that there is only 1 piece of information that can go into this parameter. Likewise, the Df for the Years of Experience and the Haberman score is 2, which means there are 2 independent pieces of information that are used in this variable. A mean square (or MS) is some estimate of the variance based on certain sources of variation. A mean square is a particular sum of squares divided by the degrees of freedom. The F value is determined by an analysis of variance that is used to determine whether the expected value of the quantitative variable within several groups differ from each other and is determined once values are placed into the computer. The higher the F value, the more confidence there is in there being a difference in the two means. The

significance shows whether or not the test statistic will reject the null hypothesis when the hypothesis is true.

**Table 8**

*Analysis of Teachers' Years of Experience and Haberman Scores*

Dependent Variable	Df	Mean Square	F Value	Significance
Years of Experience	1	.081	.036	.852
Haberman Score	2	2.424	1.065	.356
Years of Experience and Haberman Score	2	3.117	1.369	.268

In summary, there was not a statistically significant difference between teacher's years of experience and Haberman scores at the .05 level of significance. When considering years of experience, significance was at the .852 level, Haberman score was at the .356 level and the combination of both variables was at the .268 level. In other words, there is no interaction between the years of experience and Haberman scores as denoted by the .268 F value. If there were a strong interaction between the Haberman scores and years of experience, the more experience a teacher has would have an effect on their Haberman score.

All of the analyses above indicated that there is no relationship between a teacher's Haberman score and how effective they are in the classroom. The data indicates that the years of experience a teacher has may play a role in determining how effective they are in the classroom, and may show an increase in the score that the teacher has on

the Haberman interview. However, it does not indicate that there is a correlation between the score on the Haberman interview and the teacher effect score.

There is a phenomenon that exists when hiring teachers and using a prescribed interview tool, that if used appropriately the right teachers are hired for classrooms. Specifically examining middle school math teachers, this study focused on teachers that were hired in the Guilford County school district using the Haberman interview tool. This study was important to the researcher because as high stakes testing determines the success or failure of schools, it is important that the right teachers are in the classrooms. School districts rely on standardized interview tools to help select the right candidates for teaching positions. However, if the interview tool is used as a sole indicator for hiring, school districts might be placing the wrong teachers in the classrooms. To determine whether there was a correlation between a teacher's scores on the interview tool that the Guilford County school district used, research was conducted to determine if middle school math teachers' scores on the Haberman interview had any effect on student growth in the classroom as measured by teacher effect scores. One might expect that if the Haberman score of a teacher were extremely high, that the teacher might produce greater growth for the students that he/she taught. The use of Value Added data has become the process by which school districts determine how effective a teacher is with a group of students. Therefore, value added data or teacher effect scores was used to determine what effect middle school math teachers had on their students. Haberman score information was merged with teacher effect data for the middle school math teachers. The data above indicates that there is not a high correlation between the Haberman interview

scores and teacher effect data. Although teachers may have scored higher on the Haberman based on the number of years of experience they had, this examination of data indicated that teachers with more experience and a high Haberman score had a slightly higher teacher effect mean. Although this mean was slightly higher, the difference was not significant enough to say that there was a correlation between the Haberman score and the teacher effect score.

The following chapter will explain these research findings as it relates to the relationship that a teacher's Haberman score has to student achievement. I will use teacher selection, teacher retention, the Haberman instrument, and teacher effectiveness (value added data) as an indicator of teacher performance to explain this relationship. Additionally, possible recommendations and/or next steps will be discussed.

## **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 determining the relationship that a teacher's Haberman score has to student achievement. The discussion section includes a review of teacher selection, teacher retention, Haberman research as well as teacher effectiveness (value added data) research to inform the work of our principals and district leaders. Further, recommendations or possible next steps are included for educators and policy makers.

#### **Summary of Research Problem**

The purpose of this study was to take an in-depth look at teachers who were hired within the Guilford County School District using the Haberman interview instrument in the area of middle school mathematics and their teaching performance as measured by the sixth through eighth grade End-of-Grade (EOG) tests. This study was to examine whether these teachers produced student growth as determined by value added data. The measurements for the study included two independent variables: the Haberman Interview score and the SAS Educational Value Added Assessment System (EVAAS) data. Additionally, this study investigated the correlation between teacher scores on the Haberman instrument and teacher value added scores in Guilford County Schools. The student data necessary for these analyses were acquired from Guilford County Schools,

Greensboro, North Carolina, which received an analysis by the SAS Institute, Cary, North Carolina. Preparation for this study required the merging of year 2008 value added data using achievement scores from the district with the Haberman scores of middle school math teachers. Descriptive analyses were conducted to demonstrate the relationship between Haberman Scores and teachers' years of experience with the growth students made from that particular teacher. The results of this study use as a foundation the accuracy of the value added assessment formula. While this system is not perfect, researchers who have examined the system agree that the system is much better than the use of simple raw averages to reach conclusions regarding district, school, and teacher effectiveness (Sanders, 2000; Wang et al., 1993).

Basically, when teachers were hired within the GCS system, Haberman interview scores are kept by the district and only the candidates with high scores are placed on the list of possible hires. Once teachers are hired and placed in schools, they are evaluated using value added data which is an indicator of teacher performance. For middle school math teachers in the GCS system that have both pieces of data, a merging of these data was conducted to determine if there was a significant difference in the teacher effect scores of those teachers with higher Haberman scores. Additionally, a teacher's years of experience was also analyzed to see if their years of teaching had an effect on the Haberman score or the teacher effect score.

### **Findings**

Two research questions were explored in this study and the summary of the results are as follows:

*Is there a statistically significant relationship between teacher effect data and years of experience?*

The model examined 39 teachers who ranged in years of experience from 3 years to 19 years. These teachers were placed into two groups—1-5 years of experience and greater than 5.1 years—to see if there was a relationship between teachers in their beginning years of teaching and those who had more experience. When the data was sent to the office of research and accountability, part of the data included for each teacher the number of years of experience that the teacher had. From Table 5, there was not a significant difference between the years of experience and teacher effect data. The data indicated that the more experience a teacher had the higher the student's growth, however the difference was not significant enough to say that a teachers years of experience has a direct correlation to their teacher effect. The mean teacher effect score for Cohort 1 teachers was 2.808 while the mean teacher effect score for Cohort 2 teachers was 2.903. This indicates that since teaching is an art (Marzano, 2003), teachers may perform better as time moves on. Teachers may learn from mistakes they have made in the past and from other teachers as they gain experience from one year to another.

This study only contained 39 teachers which had both sets of data available. To that end, trends should be established with caution. It is good to note that teachers with more years of experience had more growth with the students they taught; however the difference is not significant. This study has implications for principals who believe that the more experience a teacher has, the more effective they are in the classroom. Although this study shows that for the 39 teachers in this study, that implication may be true

without a significant difference, principals must use more than one set of criteria for hiring before placing a teacher in the classroom of which the school will be evaluated on.

*Is there a statistically significant relationship between an effective teacher as measured by the total score on the Haberman instrument and teacher effect scores as measured by SAS EVAAS, value-added data as noted by the End of Grade test scores in middle school mathematics?*

When examining teacher effect scores by the scores teachers made on the Haberman interview, the mean scores of teachers who had the highest Haberman scores was the greatest (see Table 6). Teachers who scores were in the lower Haberman group (group 1) had a lower teacher effect average than the other teachers. However, when looking at the confidence interval of teachers within each of the Haberman groups, the mean scores typically ranged across the same average span.

Additionally, this study examined teachers' years of experience with the three Haberman cohorts as well to determine if there was a difference in the mean teacher effect scores. There was no significant difference between their mean averages of their teacher effect scores regardless of the Haberman group. Although cohort 2 (teachers with more than 5 years of experience) and Haberman group 2 (teachers whose Haberman score was between 34 and 37.5) and Haberman group 3 (teachers whose Haberman score was 38 or higher) had the highest mean teacher effect scores, there was no significant difference between those and the other group.

Although this data appears to imply that the higher a teacher Haberman score, the greater effect the teacher has on student growth, with only 39 teachers in this sample it is

caution should be used when making implications from this data. This data indicates that there is no significant relationship between an effective teacher as measured by the score on the Haberman instrument and the teacher effect scores as measured by SAS EVAAS data. While the mean average of the teacher effect scores increased by one point or so, the difference is not statistically significant enough to see that a teacher's score on the Haberman test has any validity to increasing student proficiency in middle school math. It appears that this may be a function of the years of experience that a teacher has. This indicates that although the Haberman interview tool is used to hire effective teachers as noted by Haberman, this does not provide any indication on the teachers' level of growth performance as indicated by their teacher effect scores. School districts and leaders should use caution when using this tool as the only indicator for hiring teachers. The use of only one tool does not indicate teacher effectiveness; however it can be used as one of the indicators of hiring effective teachers for middle school mathematics.

### **Discussion**

While studying the findings of these research questions, it is indeed a fact that selecting the right teachers begins the process of improving student achievement. The growth in student achievement is not going to happen by just selecting the right teacher using an interview instrument. There are many factors that come into play when the question is asked, "What determines the growth in student achievement?" The answers may vary between, the hiring instruments used, the effectiveness of the teacher, teachers' years of experience, placing teachers with the right group of students, or even having a school culture where teachers learn from one another and or have effective professional

learning communities. Regardless of the answer, this study shows that the hiring instrument or years of experience cannot be used as the sole indicator determining teacher effectiveness. Even after examining the correlation between a teachers years of experience and teacher effect as well as compare the scores on the Haberman to teacher effect, it is imperative that even when teachers are hired something else must be done to improve their teaching which is needed to continue to increase student achievement. The following discussion on teacher selection, the Haberman instrument, and teacher effectiveness (value added data) helps to solidify the importance of using more than one measure for hiring effective teachers to serve and teach our most precious resource, our children.

To begin the discussion it is important to highlight research by Wong (1996) that indicates that by placing a qualified, certified, knowledgeable and effective teacher in the classroom determines the success of our students.

The accountability for school systems and our teachers have placed the teacher selection process at the root of student achievement. It is evident that the selection of teachers is the key to the impact on student growth (Marzano, 2003). The Guilford County School system elected to use several components of the Haberman instrument to select the most appropriate teachers for our county. After taking and passing the on line screener, teachers are brought in and scored on a set of predetermined questions. However, this research suggests that although teachers may score at a specific level, this does not determine their effectiveness.

When only scores from a standardized screener are used in hiring, this may not give an accurate picture of the teacher candidate. As indicated by the data from this study, using data from only the Haberman screener did not give an accurate picture of these candidates. Even though teachers who had higher Haberman scores had a higher mean teacher effect score, the difference in the mean scores was not significant enough to say that this will always happen. Wise (1987) noted that there were many criteria that a school system must use on the selection of a teacher: determine if the candidate has cognitive competence of successful teaching, has the ability to work with diverse students and parents, and will use their skills and knowledge to best fit the needs of the school and the system. Given these findings it is crucial that the administrators or the Human Resource Department in the district use more than just this one measure to determine teacher hiring and placement in the schools. As schools begin to interview candidates, some trust must be given to the principals to determine if a candidate is the right fit for their school. The possibility of using more than one method of hiring for individual administrators may prove to be successful. Schools should determine the best method for hiring teachers for their school, however some systematic process needs to be in place before just allowing this to happen.

Haberman (1998) believed that the selection process for teachers is significant in the retention and effectiveness in the classroom. His focus on urban districts and children in poverty helped him to develop the interview tool that he believes selects the most effective teacher to stand before our children. Although his focus was on children in poverty, the Guilford County School system used this interview tool as a primary

determinant in hiring effective teachers district wide. Looking at the findings of the data in this study, there is no significant relationship between teachers hired using the Haberman interview tool and their scores in middle school math. Although the teachers with the higher Haberman scores had the highest mean scores, this data also indicates that these teachers possibly had more years of experience as well. Additionally, these findings indicate that the more experience a teacher has, the more likely they will have an impact on student growth, although that correlation is not significant. In order for school districts to retain highly qualified and effective teachers, the selection process is very important. Not only is the selection important, the retention of teachers seems to have an impact on the amount of growth that a child makes.

One would conclude that regardless of the score on the Haberman interview, the effect of a teacher on the growth of students may be determined by many other factors other than one simple interview score. Additionally, if the use of one score is the primary factor in determining whether a teacher is hired or not, a school system would need to use other data points to determine whether to hire a teacher candidate.

As school districts continue to deal with No Child Left Behind, the accountability for teaching and increasing student proficiency for all students will remain. In order to determine whether we are hiring the most effective teachers, school districts must use a fair and objective way of determining the effect a teacher has on student achievement. SAS or any other value added data assessment will help to determine this effectiveness (Doran and Fleishman, 2005). While the use of value added data (VAD) is not perfect and has some challenges, it provides a fairer accountability measure for schools and

teachers. Value added data focus on student growth and therefore every teacher is measured using the same criteria regardless of demographics, economic status of children, or levels of proficiency. It is based on the assumption that the teacher is the most important influence on student achievement. Value added assessments steer people away from focusing on the percent or number of students who reach an arbitrary standard. Instead it focuses at improvements (Allen et al., 2004).

In order to improve teacher accountability and improved teaching, teachers must be given their value added data and understand how it is determined, which hopefully will give them that intrinsic motivation to be better. Since Value Added Data (VAD) is given to principals, the assumption is that principals use this information to place the most appropriate teachers with specific students. The value added teacher report (see Appendix B) shows upon which group of students a teacher has the most effect, whether positive or negative. This information can be used during a teacher conference to help the teacher with strategies to improve in the area which they are weakest and to assign the teacher to the group of students with whom they have the highest gain. Additionally, if a teacher shows negative growth with all groups of students, this is an indication that a different placement may be needed. Another assumption is that principals use these data to create teams of teachers that support one another in professional learning communities. Teachers must learn to work with one another to create the best possible learning environment for students, understanding that the purpose of school is learning, not teaching. Therefore, teachers who want to know, understand, and do whatever it takes to ensure that students learn are really the most effective ones. This belief causes teachers to

approach their work differently than those who believe that “the locus of control over their efforts and their students’ learning resides outside themselves” (Haberman, 1987, p. 24). Haberman contends that the strategies and efforts that results from this factor are often the difference between the success and failure of a student.

In this study, as in the studies conducted by Wise (1987), the teacher is the most significant contributor to student achievement. This study indicates that the use of value added data can help to determine the effectiveness of a teacher. However, when placed in context with Haberman scores and years of experience, it was difficult to determine which variable was an indicator of the teacher performance when dismissing all other indicators that could have had an impact on the value added data as well. As indicated by the data, teachers who have more than five years of experience and an average Haberman score of 34-37.5 points had a teacher effect mean equal to 3.38. Teachers who have more than five years of experience and a high Haberman score of 38 or above (Group 3) have a teacher effect mean equal to 3.50. This cohort had a positive correlation in relation to the Haberman scores. Conversely, teachers with less than five years of experience have an inverse relationship to the Haberman scores. For example, the higher your Haberman score, the lower your teacher effect. The teacher effect scores of teachers who have more than five years of experience are positively correlated to the higher Haberman scores. The researcher notes here that although the mean average of the value added data increased within Haberman groups for teachers with more than five years of experience, it is difficult to really determine whether the impact on student achievement was due to the years of experience or to the selection of them by the use of the Haberman tool.

In sum, when comparing the Haberman scores of middle school math teachers in the Guilford County school district with their teacher effect scores, the analysis indicates that there is more to finding indicators of teacher success other than the use of a hiring instrument. Thirty-nine middle school math teachers' Haberman scores were examined with their teacher effect scores to see if there was in fact a correlation between the two. Additionally, teachers' years of experience was analyzed to see if this in fact had an effect on the teacher effect scores. After comparing all of these indicators, this study has determined that the sole use of the score on the Haberman instrument is not an indicator of teacher effectiveness for middle school math teachers. As one would conclude, a teacher's years of experience and many other factors play a role in how effective a teacher can be.

Once teachers are hired within a district, the retention of those teachers is highly important as well. As the research indicates, there are many factors that contribute to teacher retention such as supplemental pay for hard to staff areas like math, working conditions, and teachers who felt effective with their students (NC Educational Research Data Center, 2006). To determine how effective a teacher is with certain groups of students, value-added data can be used. The use of value-added data is a step towards improving the teaching and learning in the classroom if used correctly. Additionally, value-added assessment is one technique that helps to identify effective and ineffective teachers and schools (Doran & Fleischman, 2005). If value added data is used by principals and districts to inform instruction, it is possible that teachers will improve and student growth will be on the rise. What must we do to ensure that principals and district

leaders do not use one sole measure for the hiring of effective teachers and use value added data to improve instruction? The following conclusions will answer this question.

### **Conclusions**

It is evident that effective teachers must be hired and retained in all school systems in order to meet the demands of NCLB. The results of this study do not provide any significant findings. However it does suggest that when comparing teacher years of experience and value added data, teachers with more years of experience had a higher mean average than the teachers with fewer years which could be a result of many factors. When a correlation was made with the value added data and Haberman scores added to the years of experience, there was no significant difference on student academic achievement growth between the groups of teachers with a high Haberman score and the group of teachers with a low Haberman score. The mean average showed no significant increase, which suggests that teachers hired based on a Haberman score may not be producing as much student growth as other teachers. While the use of the Haberman score may be an indicator of teacher performance, it cannot be used solely to determine student growth or to place teachers in our schools.

Speaking from the administrative perspective, it is apparent that other means of hiring teachers must be used by our Human Resource department rather than just using one Haberman score. For example, human resource departments may want to consider many interview instruments or may allow principals to hire teachers for their own schools using questions created by the principal. School districts that only allow one instrument

to be used as universal screener may have turned the most effective teachers away because of their inability to provide a prescribed answer.

Although this study indicated that there is no significant correlation between a teacher's Haberman interview score and teacher effect data, it does not indicate the other factors that may have contributed to the teacher effect data such as ineffective leadership, ineffective professional learning communities, and the lack of professional development opportunities or even teacher retention within the same school. Additionally, with a small sample of 39 teachers in this particular study, it is difficult to determine or assume if all of the teachers hired in the district using the Haberman interview instrument have a positive impact on student achievement. Of the 1947 teachers hired within the GCS system using the Haberman instrument in 2007-2008 only 39 teachers emerged from the data that were middle school math teachers and had EVAAS data. Since this data set is very small, caution should be exercised when making implications from this data. A larger number of teachers in the data set would have conducted a better statistical analysis. Finally, Haberman indicated that his focus was mainly on urban school districts and hiring teachers for children in poverty (Haberman & Post, 1998). Due to the fact that GCS does not utilize the Haberman interview for hiring teachers for highly impacted schools, but for all schools, this study utilized the data for math teachers in all middle schools within the district.

I would argue that the Haberman interview tool is a great instrument to be used as one indicator of hiring effective teachers. However, many other factors need to be in place when making teacher selections. The Haberman instrument was designed to be

used for highly impacted schools. To use this instrument to hire for teachers in all schools within the district may not be the best method. Some schools within the district have a higher population of children in poverty than other schools and this tool was designed to help hire teachers in high poverty schools, not districts. Therefore, districts should use instruments for which they are effectively designed for. There is no reason to only use one instrument within a school district if there are schools within the district that have different populations of students. It is possible to have several instruments to be used for the selection of teachers for different schools.

Secondly, when hiring teachers with experience, the use of their value-added scores could prove to be beneficial along with their score on a particular interview tool. It is evident that no one sole indicator can be a factor in selecting an effective teacher, but the use of different factors may prove to be beneficial. For teachers who are new to teaching, their content knowledge and their belief about how students learn and should be taught may be factors in the decision to hire the teacher. For teachers with experience, the use of their value-added scores along with their interview score will help to determine the best placement for the teacher. The use of the value-added score will aid in knowing where a teacher's specific development needs to occur; whether with a certain level of students or with a certain subject. Additionally through teacher observations, assignment of mentors and other methods, principals could communicate with teachers and help them with their weaknesses. This development for a teacher would help in a teachers' delivery of instruction and maybe content knowledge which could possibly help to increase

student growth. This positive relationship could be the answer to an increase in teacher retention as well.

### **Recommendations for Future Study**

This study of teacher Haberman scores and value added data indicates that there is no statistically significant relationship between the score on the Haberman instrument and the teacher effect scores as measured by SAS EVAAS data for middle school students in math. Further research may be appropriate in the following areas: First, there have only been two and a half years that the Haberman interview process has been used to recruit teachers in this particular district. The real impact on student achievement takes a longer period of time. A follow up study using an interview tool and its correlation on student achievement could be done. Secondly, if school districts have changed the interview tool used, a study could be conducted comparing the tools to one another comparing teachers hired using one interview tool against the teachers hired using the other tool with value added data as a common link. Additionally, further research comparing teachers hired using an instrument to those who were hired not by an instrument should be conducted. This could be used as a random sample of other teachers to make a comparison study. Of the 1947 teachers recruited using the Haberman model, when merged with the value added data specifically for middle school math teachers, only 39 teachers emerged from this data. When analyzing data that only includes a small portion of recruited teachers, further research could be conducted after a longer period of time or expanding the data sort beyond middle school math teachers. For

further research on this particular study, a continued study on these 39 teachers could be conducted to determine if they continue to have positive value added data.

According to Haberman (Haberman & Post, 1998), his focus was mainly on urban school districts and children in poverty. Since Haberman interview instrument is to be utilized for hiring teachers for schools of poverty, a further study could be conducted to look at teachers who only teach in highly impacted schools of poverty to determine if there is a significant difference in their value added data.

Finally, any school system could utilize this study to dialogue about the use of interview tools and how they match with the goals of the district. Some districts may choose to pilot different tools for hiring effective teachers at the district level or may leave the hiring of potential teachers to the school building administrators themselves. The results of further studies could provide Boards of Education with direction on how to find effective teachers to place in the front of our classrooms that produce growth with all students especially in these economic times and with high accountability.

### **Recommendations for Policy Makers**

It is imperative that policy makers create policies that will ensure that teachers of quality are hired within school districts, provided ways of improving in areas that may be ineffective, and held accountable for what occurs in their classrooms. When using teacher selection instruments to hire teachers, it is crucial that these instruments are used in the manner in which they were designed. To do this, frequent audits and reviews of the hiring processes must take place within districts. Additionally, when districts have value added data that can be used to determine a teachers' effectiveness, it is imperative that this data

is used appropriately and monitored by the building level administrators. We all know that principals have a lot on their plates, but the key to improving student achievement is by analyzing teacher data as well as student data. Review procedures need to be placed into policy that governs school districts on the use of such data.

### **Recommendations for School Administrators**

Based on this research study, value added data is an important tool in recognizing the strengths and weaknesses of teachers. It is noted that if access to this data is upon your hands, carefully studying and reviewing the data and using it to place teachers appropriately may be of benefit to student achievement and may very well decrease the teacher turnover rate in the school. Discussing the VAD with teachers may prove to be beneficial as well. Teachers may not understand which students they are most effective with and least effective with unless they are shown data to prove it. Once teachers have this information before them, strategies to build on the ineffective areas is key to improving teacher quality which will improve student achievement.

In summary, this research study clearly suggests that having an effective teacher in the classroom is the key to improved student achievement. There are many different views on what effectiveness may be, how to determine effectiveness and even how to retain effective teachers. With the high accountability and the economic times, finding effective teachers to fill classrooms is still an issue. Since school districts have vacant classrooms, the urge to place just anyone in the classroom is high. School districts will have to take the time to find the most appropriate people to stand before our children. Whether scripted interview tools are used or not, school leaders must find creative ways

of selecting the right people and retaining them. The challenge begins with universities preparing potential teachers effectively for public education and for school leaders to adequately maintain effective teachers and improve teaching and learning. The task is before all educational leaders.

## BIBLIOGRAPHY

- Allen, R., Galinat, J., Leedham, C., & Leverich, J. (2004). Research Briefs: Value Added Assessment. WEAC Research Work Group: Brief #2.
- Ansell, S. E., & McCabe, M. (2003, January 9). Off target. *Education Week: Quality counts 2003*, XXII(17), 57-58.
- Ballou, D. (2002, Summer). Sizing up test scores. *Education next*. Retrieved October 9, 2008, from <http://www.educationnext.org/20022/16.html>.
- Boardman, A.E., Darling-Hammond, L., & Mullin, S.P. (1982). A framework for the analysis of teachers' demand and supply. *Economics of Education Review*, 2(2).
- Bredeson, P. V. The Effects of the Letters of Recommendation on the Teacher Selection Interview Process. Ph.D. Dissertation, University of Wisconsin, 1982.
- Center for Public Education. (2005), Teacher quality and student achievement research review.
- Cronbach, L. J., & Snow, R.E. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. Irvington, New York
- Darling-Hammond, L. (1997). Toward what end: The evaluation of student learning for the improvement of teaching. In J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid measure?* (pp. 248-263). Thousand Oaks, CA: Corwin Press, Inc.

- Darlington, R. B. (1997). The Tennessee value-added assessment system: A challenge to familiar assessment methods. In J. Millman (Ed.), *Grading teachers, grading schools*. Thousand Oaks, CA: Sage.
- Denzin, N. K., & Lincoln, Y. S. (2000). *Handbook of qualitative research* (2<sup>nd</sup> ed.). Thousand Oaks, CA: Sage.
- Doran, H. C., & Cohen, J. (2005). The confounding effect of linking bias on gains estimated from value-added models. In R. Lissitz (Ed.), *Value-added models in education: Theory and applications*. Maple Grove, MN: JAM Press.
- Doran, H. C., & Fleishman, S. (2005). Research matters/challenges of value-added assessment. *Educational Leadership*, 63(3), 85-87.
- Evergreen Freedom Foundation. (2004). *Value added assessment: School directors handbook*. Retrieved October 7, 2008, from <http://www.effwa.org/pdfs/Value-Added.pdf>.
- Gage, N. L. (1978). *The scientific basis of the art of teaching*. New York: Teachers College Press.
- Gallup Organization. (2002). *TeacherInsight system*. Princeton, NJ: Author. Retrieved October 12, 2008, from <http://education.gallup.com>.
- Goe, L., Bell, C., & Little, O. (2008). *Approaches to evaluating teacher effectiveness: A research synthesis*. Washington, DC: National Comprehensive Center for Teacher Quality.
- Gootman, E. (2009). Next test: Value of \$125,000-a-year teachers [Electronic version]. *New York Times*. Retrieved August 28, 2009, from

<http://www.nytimes.com/2009/06/05/education/05charter.html?scp=1&sq=&st=nyt>.

Guilford County Schools. (2008). GCS Haberman Interview Process. Retrieved June 12, 2008 from [http://www.guilford.k12.nc.us/boe/2008/6\\_10/haberman\\_process.pdf](http://www.guilford.k12.nc.us/boe/2008/6_10/haberman_process.pdf).

Haberman, M. (1987). *Recruiting and selecting teachers for urban schools*. (ERIC/CUE Urban Diversity Series No. 95). U. S. Department of Education, Office of Educational Research and Improvement. [ERIC Document No. ED 292 942].

Haberman, M. (1995b). *A brief review of the history and development of the urban teacher selection interview*. Unpublished manuscript, University of Wisconsin at Milwaukee.

Haberman, M., & Post, L. (1998). Teachers for multicultural school: The power of selection. *Theory into Practice*, 37(2), 96-104.

Haberman, M., Stafford, D., & Dill, V. (1993). *Star teacher selection interview training manual*. Houston: The Haberman Educational Foundation, Inc.

Haberman Educational Foundation. (2003). *Star teacher on-line pre-screener*. Retrieved April 14, 2008, from <http://www.altcert.org/teacher/dimensions.asp>.

Hanushek, E. A., Kain, J. F., O'Brien, D. M., & Rivkin, S. G. (2005). *The market for teacher quality*. Cambridge, MA: National Bureau of Economic Research.

Retrieved August 16, 2009, from [www.nctq.org/nctq/research/1112806467874.pdf](http://www.nctq.org/nctq/research/1112806467874.pdf).

- Hester, L. S. (2006). *Coming to grips with the achievement divide and the distribution of effective teachers*. Published doctoral dissertation, University of North Carolina at Greensboro.
- Holcombe, A. (2008). *Understanding individual teacher value added data*. Powerpoint presentation during training for Guilford County Schools principals.
- Ingersoll, R., & Rossi, R. (1995). *Which types of schools have the highest teacher turnover?* Issue Brief (Report No. IB-5-95: NCES-95-778). Washington, DC: National Center for Educational Statistics. [ERIC Document No. ED 388 663].
- Ingersoll, R. M. (1997). Teacher turnover and teacher quality: The recurring myth of teacher shortages. *Teachers College Record*, 99(1), 41-44.
- Kemp, L., & Hall, A. H. (1992). *Impact of effective teaching research on student achievement and teacher performance: Equity and access implications for quality education*. Jackson, MS: Jackson State University. (ERIC Document Reproduction Service No. ED 348 360).
- King, J. A. (1981). "Beyond classroom walls: Indirect measures of teacher competence. In J. Millman (Ed.), *Handbook of teacher evaluation*. Beverly Hills, CA: Sage Publications.
- Kirby, S. N., & Grissmer, D. W. (1993). *Teacher attrition: Theory, evidence and suggested policy options*. Washington DC: U. S. Department of Education (ERIC Document Reproduction Service No. ED 364 533).

- Kupermintz, H. (2003, Fall). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee value added assessment system. *Educational Evaluation and Policy Analysis*, 25(3), 287-298.
- Marzano, R. J. (2003). *What works in schools*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Meyer, R. H. (1997). Value-added indicators of school performance: A primer. *Economics of Education Review*, 16(3), 283-301.
- Miller, M. D., Brownell, M. T., & Smith, S. W. (1999). Factors that predict teachers staying in, leaving, or transferring from the special education classroom. *Exceptional Children*, 65(2), 201-218.
- Muller, G. (1981). Teacher selection: A response. *Educational Leadership*, 31, 494.
- National Conference of State Legislators. (2008). Retrieved October 7, 2008, from <http://www.ncsl.org/programs/educ/edu.htm>.
- North Carolina Center for Public Policy Research. (n. d.). Retrieved October 10, 2008, from <http://www.nccppr.org/Teachershortage.pdf>.
- North Carolina Department of Public Instruction. (2009). Teacher turnover report. Retrieved September 7, 2009, from <http://www.ncpublicschools.org/docs/stateboard/meetings/2009/09/tcp/09tcp07.pdf>.
- Norton, M. S. (1999). Teacher retention: Reducing costly teacher turnover. *Contemporary Education*, 70, 52-55.

- Pearson Education, Inc. (2004). Value-added Assessment Systems. Retrieved October 2, 2008, from <http://pearsonassess.com/NR/rdonlyres/2C54ADBA-9959-4F82-82B3-8E65FB822088/0/ValueAdded.pdf>.
- Rebora, A. (2003, February). The retention imperative. Retrieved October 10, 2008, from Education Week Marketplace jobs Online Web Site:  
[http://www.edweek.org/jobs/jobstory.cfm?slug=07keep\\_tr.h02](http://www.edweek.org/jobs/jobstory.cfm?slug=07keep_tr.h02).
- Rice, J. K. (2003). *Teacher quality, understanding the effectiveness of teacher attributes*. Washington, DC: Economic Policy Institute.
- Rosenshine, B. V. (1979). Content time and direct instruction. In P. L. Peterson & H. J. Walberg (Eds.), *Research on teaching*. Berkley, CA: McCutchan.
- Sacks, P. (1999). *Standardized minds*. Cambridge, MA: Perseus Books.
- Sachs, S. K. (2004). Evaluation of teacher attributes as predictors of success in urban schools. *Journal of Teacher Education, 55*(2), 177-187.
- 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. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education, 14*, 329-339.
- Sanders, W. L. (2000). *Value-added assessment from student achievement data*. Cary, NC: Create National Evaluation Institute.
- Smith, J. M. (1980). Teaching selection procedures. *The Clearing House, 53*, 312-314.

- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stevenson, Jr., Z., Dantley, S. J., & Holcomb, Z. J. (1999). Factors influencing the retention of mathematics and science teachers in urban systemic initiative school districts: Administrative perspectives. *Journal of Negro Education*, 68(3), 442-450.
- Stewart, B. E. (2006). *Value added modeling: The challenge of measuring educational outcomes*. New York: Carnegie Corporation of New York.
- Stronge, J. H. (2002). *Qualities of effective teachers*. Alexandria, VA: Association for Supervision and Curriculum.
- Teacher Magazine. (2002, November). The average teacher. *Teacher magazine*, 14(3), 30-31.
- Thomas, R. M., & Brubaker, D. L. (2000). *Theses and dissertations: A guide to planning, research, and writing*. Westport, CT: Bergin & Garvey.
- The Urban Teacher Collaborative. (1996). *The Urban teacher challenge: A report on teacher Recruitment and demand in selected Great City Schools*. Belmont, MA: Segun C. Eubanks. New York.
- Ventures for Excellence, Inc. (2004). Retrieved from <http://www.venturesforexcellence.com>.
- Wang, M. C., Haertel, G., & Walberg, H. J. (1993). Toward a knowledge base of school learning. *Review of Educational Research*, 73, 249-294.
- Walsh, K. (2007). *If wishes were horses: The reality behind teacher quality findings*. Washington, DC: National Council on Teacher Quality. Retrieved on August 16,

2009 from [http://www.nctq.org/p/publications/docs/wishes\\_horses\\_20080316034426.pdf](http://www.nctq.org/p/publications/docs/wishes_horses_20080316034426.pdf).

Weaver, T. (1979). In search of quality: The need for talent in teaching. *Phi Delta Kappan*, 61(1), 29-31, 46.

*What Matters Most*. (1996). New York: National Commission on Teaching & America's Future. [Online]. Available: <http://www.ncpublicschools.org/docs/recruitment/surveys/turnover/2005-06turnoverreport.pdf>.

Wiersma, W., & Jurs, S. (2005). *Research methods in education* (8<sup>th</sup> ed.). New York: Allyn & Bacon.

Wise, A. E., Darling-Hammond, L., Berry, B., Berliner, D. C., Haller, E., Praskac, A., & Schlechty, P. C. (1987). *Effective teacher selection: Recruitment to retention*. Santa Monica, CA: Rand Corporation.

Wong, H. (2007, March 16). *The single greatest effect on student achievement is the effectiveness of the teacher*. Presentation to the North Carolina Principal's Executive Program.

Young, I. P. (1984). The effects of interpersonal performance style in simulated teacher selection interviews. *Journal of Research and Development in Education*, 17, 43-51.

**Appendix A****IRB Approval**

August 18, 2008

Jocelyn B. Becoats  
7098 Toscana Trace  
Summerfield, NC 27358

Re: 080902

Dear Ms. Becoats:

The Research Review Committee has concluded that your proposal *Determining the Correlations of Effective Middle School Math Teachers and Math Student Achievement* meets the requirements of state legislation and the current research policy of the Guilford County Schools. It is suggested that the higher end of your suggested sample size be used so that the chance for sampling error is minimized. The committee understands from your proposal that the identities of individuals and schools will be confidential at all stages of the project.

We encourage researchers to provide direct feedback to the school community where the research has been conducted. This could involve an open session for parents, teachers, and administrators; a summary of the research that is accessible to all members of the school community; or other accessible forms of direct feedback to the schools. In addition, you must provide the Research Review Committee with copies of any publications and presentations.

We hope that the project is successful in helping to achieve your goals. Please feel free to contact me at 336-370-2346 if you have any questions.

Sincerely,

Carolyn Gilbert  
Co-Chair, Research Review Committee

**STRIVING. ACHIEVING. EXCELLING.**

712 North Eugene Street Greensboro, NC 27401 P 336.370.8100

## Appendix B

### Sample Value Added Report

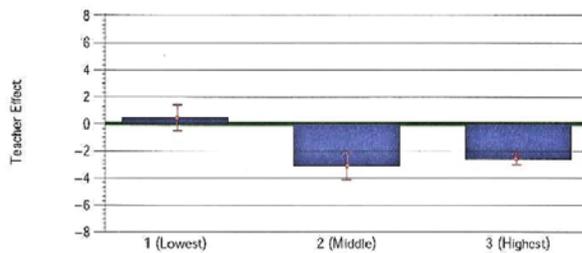
#### SAS® EVAAS® Teacher Value-Added Report for 2008 Guilford County Schools

School:  
Teacher:  
Subject: End of Grade Math, Grade 8

Year	N	Mean Student Score	Mean Score %tile	Mean Pred. Score	Pred. Score %tile	Teacher Effect	Effect Std Error	Teacher vs Comparison Avg
2008	86	361.4	51	363.4	59	-1.9	0.4	Below

Estimates are from multivariate, longitudinal analyses using all available test data for each student (up to 5 years). The analyses were completed via SAS® EVAAS® methodology and software, which is available through SAS Institute Inc. EVAAS, SAS, and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. © indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2008 SAS Institute Inc., Cary, NC, USA. All Rights Reserved.

#### SAS® EVAAS® Teacher Diagnostic Report for 2008



Year		Observed minus Predicted Score by Predicted Score Subgroup		
		1 (Lowest)	2 (Middle)	3 (Highest)
2008	Mean	0.4	-3.1	-2.6
	Std. Error	0.9	1.0	0.4
	Nr of Students	22	24	40
	% of Students	25.6	27.9	46.5

#### Interpreting the Teacher Value-Added Report

Use this report to evaluate how well a teacher facilitates student progress. The **Teacher Value-Added Report** compares each teacher to the average teacher in the participating districts. This comparison indicates how a teacher influences student progress in the given subject.

#### Teacher vs. Comparison Average

The **Teacher vs. Comparison Average** column shows whether there is a statistical difference in the progress rate for this teacher compared to the progress rate for the average teacher in the participating districts. Comparisons are made based on two standard errors.

- **Above** means that students taught by this teacher made significantly more progress than students taught by the average teacher.
- **Below** means that students taught by this teacher made significantly less progress than students taught by the average teacher.
- **NDD** means that the progress made by students taught by this teacher was Not Detectably Different from the progress made by students taught by the average teacher.

#### Teacher Effect & Standard Error

The **Teacher Effect** is a conservative estimate of a teacher's influence on students' academic progress. The number in this column is a function of the difference between the students' observed scores and their expected scores. An average teacher would have an estimate of 0.0. The **Effect Standard Error** provides the basis for establishing a confidence interval around the Teacher Effect value.

#### Mean Predicted Score and Predicted Score Percentile

A student's predicted score is an expected score, based on his or her performance on previous tests, assuming the student is taught by the average teacher in the participating districts. The value in the **Mean Predicted Score** column is the average of the predicted scores for all students taught by this teacher. The **Predicted Score %tile** shows where this score falls in the distribution that includes all students' predicted scores in the participating districts. Similarly, the **Mean Student Score** is the average of the students' observed scores, and the **Mean Score %tile** shows where this score falls in the distribution that includes all students' observed scores in the participating districts.

Comparing the **Mean Score Percentile** and the **Predicted Score Percentile** provides a simple indication of the amount of progress students made this year. When students are taught by an average teacher, the Mean Score Percentile will be very close to the Predicted Score Percentile. The Mean Score Percentile will be lower than the Predicted Score Percentile when the students are taught by a less effective teacher, and higher when they are taught by a more effective teacher.

#### Interpreting the Teacher Diagnostic Report

Use this report to identify patterns or trends of progress among students at different achievement levels. This report is intended for diagnostic purposes only.

The chart offers a visual representation of student progress for different prior achievement levels and can be used to identify patterns in progress among students. The **green** line is the Reference line, or the amount of progress students must make to keep up with their peers. **Blue** bars show the progress of students in the current school year, while **yellow** bars (if present) show the progress of students in previous school year(s). Bars above the green line indicate that students in the subgroup made better than average progress. Bars below the line indicate that students made less than average progress. No bar is presented for subgroups with fewer than five students.

As part of the analyses, each student is assigned to a **Predicted Score Subgroup** based on his or her performance on previous tests, assuming the student is taught by an average teacher. Since students are assigned based on where they fall relative to all students in the participating districts, some teachers may find that certain subgroups are more heavily populated than others.

In the table immediately below the chart, the **Mean** of the difference between students' observed test performance and their predicted performance appears for each Predicted Score Subgroup, along with the **Standard Error** associated with the Mean. The Standard Error allows you to establish a confidence band around the Mean (the red interval on the chart).

- A large positive mean indicates that students within a group made more progress than expected.
- A large negative mean indicates that students within a group made less progress than expected.
- A mean of approximately 0.0 indicates that a group is progressing at an average rate.

When the means among groups vary markedly, teachers may want to explore ways to improve the instruction for the students making less progress.