A LONGITUDINAL STUDY OF THE USE OF FORMATIVE ASSESSMENT DATA TO IMPROVE ACHIEVEMENT OF MIDDLE SCHOOL STUDENTS

KATHY L. CARROLL

A Longitudinal Study of the Use of Formative Assessment Data to Improve Achievement of Middle School Students.

A Field Study

Submitted to the

Graduate and Research Council of

Austin Peay State University

in

Partial Fulfillment of the Requirements

For the Degree of

Education Specialist

Kathy L. Carroll

May 2008

kathy.carroll@rcstn.net

To the Graduate Council:

I am submitting herewith a Field Study written by Kathy L. Carroll entitled "A Longitudinal Study of the Use of Formative Assessment Data to Improve Achievement of Middle School Students." I have examined the preliminary copy of this paper for form and content and recommend it be accepted in partial fulfillment of the requirements for the Degree of Education Specialist.

Dr. J. Gary Stewart
Major Professor

We have read this Field Study
and recommend its acceptance:

Dr. Tammy Shutt

Dr. Donald Luck

Charles Linder

STATEMENT OF PERMISSION TO USE

In presenting this field study in partial fulfillment of the requirements for an Education Specialist degree at Austin Peay State University, I agree the Library shall make it available to borrowers under the rules of the Library. Brief quotations from this field study are allowable without special permission, provided accurate acknowledgement of the source is made.

Permission for extensive quotation or reproduction of this field study may be granted by my major professor, or in his absence, by the Head of the Inter-library Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or the use of the material in this field study for financial gain shall not be allowed without my written permission.

Signature Kelly K. Cannell

Date May 2, 2008

DEDICATION

This field study is dedicated to my family. To my parents, Benjamin and Clara King, who have always encouraged me to be more than I ever believed I could be, and to my husband, Grant and my daughters, Bethany and Braden, who have always provided me with continual guidance, encouragement, and endless support. I dedicate this work with love to you all.

ACKNOWLEDGEMENTS

I want to express my sincere appreciation and gratitude I feel for the never-ending support I have received throughout this endeavor. It is because of my family, colleagues, and mentors, who played an important role in this process, that this field study has been accomplished. I thank my advisor and committee chair, Dr. J. Gary Stewart, for his continual support and guidance. Dr. Stewart has given his time and expertise to make my work quality work. I am also grateful to my committee members, Dr. Tammy Shutt and Dr. Donald Luck, for their guidance and support as this work has proceeded through each step. Thank you to Dr. Mike Morris for starting me on this journey, for always believing in my ability, for opening my eyes to the deeper world of education, and for making me a better educator. Jennifer Darden, Kathy Neal, and Jana West thank you for your encouragement and enthusiasm. The three of you have listened to endless stressing out sessions, read through many a rough draft copies and have been there offering help every step of the way. I thank my parents, Benjamin and Clara King, for your support and care over the years; you have instilled in me the importance of hard work and education and to use both skills to make anything possible. I owe the greatest debt to my family for the sacrifices they have made during this field study. Thank you to my husband, Grant, for being a loving support in many ways and believing in my ability to accomplish my goals. My husband and daughters have provided me with the utmost patience, time, and support

enabling me to focus on my goals. My daughters, Bethany and Braden, are not only wonderful daughters but the best personal cheerleaders anyone could ask for. I know you all put up with endless hours, days, and weeks thinking the computer and textbooks were actually extensions of my body. I love you both very much and I hope this experience motivates and encourages you throughout life to reach for all of your dreams.

ABSTRACT

KATHY LYNN CARROLL. A Longitudinal Study of the Use of Formative Assessment Data to Improve Achievement of Middle School Students.

This study analyzed and evaluated the TCAP scores of 275 students in grades six through eight in a rural, southern school district. The purpose of this study was to determine if the use of formative assessment data by educators, to drive curriculum and instructional decisions, had any effect on student achievement. ANOVA tests were used to analyze data for statistically significant differences between group means. The study was conducted to test four null hypotheses at the .05 level of confidence. Analyses revealed formative assessment can be beneficial in aiding student growth in reading/language arts and math TCAP scores.

TABLE OF CONTENTS

List of Tables

CHAPTER I

Introduction

Background of the Problem

Enacted in 1965, the *Elementary and Secondary Education Act* (ESEA) funds primary and secondary educators' professional development, instructional materials, resources to support educational programs, and the promotion of parental involvement. The Act was originally authorized through 1970, but has been reauthorized every five years. On January 8, 2002, President George Bush signed the *No Child Left Behind Act* (NCLB), which reauthorized and amended federal education programs designed under the ESEA of 1965. Over the years, the Act has undergone name changes and presidencies. However, the basic intent of the law still stands today; providing resources to help ensure those who are considered disadvantaged have access to a quality public education equal to their non-disadvantaged peers. The major focus of *No Child Left Behind 2001* is to disaggregate achievement data (low socio-economic, gender, special education, and ethnicity) so that schools can identify areas of deficiency and target resources in an attempt to close the achievement gap between disadvantaged students and their peers.

Under NCLB, every student in the United States should be proficient in every subject area by the year 2014. The ultimate goal is to better align daily instruction with state standards and more effective use of disaggregated test data to identify areas of

weakness and change instruction in those areas. States are also required to bring a percentage of students to a proficient or advanced level each year. Students are assessed each year and labeled as advanced, proficient, or non-proficient. A student labeled non-proficient must be provided extra resources to bring them up to proficiency before the next yearly assessment.

NCLB mandates create a situation in which educators are struggling with the identification of effective methods of instruction for those students that continue to be identified as non-proficient. Administrators and teachers have to be creative in finding ways to reach at-risk students. Every instructional tool available to educators and the use of collaboration among educators must be implemented in order for all students to achieve success. Every student should have access to a quality public education and should be given the tools to ensure individual success.

Statement of the Problem

The requirements of *No Child Left Behind* call for all students to be proficient in all assessed subject areas by the year 2014. Educators must look for methods to identify students who are non-proficient and provide a means by which these students will be brought to proficiency levels by the time the state mandated achievement tests are administered.

Purpose of the Study

The purpose of the field study was to determine if the longitudinal use of formative assessment data to drive instructional decisions had an effect on student achievement.

Significance of the Study

The field study determined what effect data-driven instruction, based on the results of formative assessments, had on the achievement scores of six through eight grade students in an effort to meet NCLB mandates of adequate yearly progress (AYP).

Research Questions

Is there a difference in student achievement due to the use of formative assessment data to drive curriculum and instructional decisions?

Is there a difference in student achievement in regard to student socioeconomic status due to the use of formative assessment data to drive curriculum and instructional decisions?

Is there a difference in student achievement in regard to gender due to the use of formative assessment data to drive curriculum and instructional decisions?

Is there a difference in student achievement in regard to ethnicity due to the use of formative assessment data to drive curriculum and instructional decisions?

Hypotheses

There is no statistically significant difference between the achievement scores of students instructed with the use of formative assessment data and those not instructed with the use of formative assessment data.

There is no statistically significant difference in student achievement based on student socioeconomic status due to the use of formative assessment data to drive curriculum and instructional decisions.

There is no statistically significant difference in achievement scores of male and female students instructed with the use of formative assessment data.

There is no statistically significant difference in achievement scores based on ethnicity due to the use of formative assessment data to drive curriculum and instructional decisions.

Limitations

This field study was limited to a population of mainly Caucasian students from a small, rural geographical area. The population was also limited to students from one school.

Delimitations

This field study was delimited to students in grades six through eight in two academic areas – reading/language arts and math. The study involved the collection of TCAP data from years 2004-2007.

Assumptions

TCAP achievement scores reflect the true achievement of a student. All classroom teachers are teaching to their full potential and all students perform to their full potential on the TCAP achievement test.

Definition of Terms

Formative assessment – Black and Wiliam (1998) define assessment broadly to include all activities teachers and students undertake to get information that can be used

diagnostically to alter teaching and learning. Assessment become formative when the information is used to adapt teaching and learning to meet student needs.

Data-driven decision-making (DDDM) – DDDM in education refers to teachers, principals, and administrators systematically collecting and analyzing various types of data, including input, process, outcome, and satisfaction data, to guide a range of decisions to help improve the success of students and schools (Marsh, Pane & Hamilton, 2006).

ThinkLink – Founded by Vanderbilt University, ThinkLink has pioneered a unique approach to formative assessments using a scientifically research-based continuous improvement model that maps diagnostic assessments to each state's high stakes test. ThinkLink Learning's Predictive Assessment Series (90% accuracy) assesses student progress toward meeting state standards for reading/language arts, math, science, and social studies. The tests are preconfigured and analyzed by ThinkLink experts to provide teachers the prediction of mastery, proficiency, and adequate yearly progress (AYP) that is so critical to monitoring student progress toward state mandated goals. The benchmark tests are administered in the fall, winter, and spring. Each test is carefully constructed to mirror and match the state test (ThinkLink Learning, 2006).

TCAP- Tennessee Comprehensive Assessment Program is a timed; multiple choice state-mandated yearly exam that measures skills in reading, language arts, mathematics, science, and social studies of students in grades three-eight. Student results are reported to parents, teachers, and administrators.

CHAPTER II

Review of Literature

What is formative assessment?

No Child Left Behind (NCLB) was passed with the intent to ensure all students, regardless of race, gender, socio-economic status, English proficiency level, or a special education status receive a fair and equal opportunity to acquire a high-quality education through standards, accountability, and assessment. McDonald (2002) defines the standard for "high quality" per NCLB as attaining a minimum level of proficiency on state required academic achievement standards and summative assessments. NCLB sets a goal of closing the achievement gap between students identified as disadvantaged and their peers by the year 2014. Schools and school systems are now faced with the task of not only achieving a hefty goal of all students performing at a proficient or advanced level by 2014 but they must also demonstrate a closing of any achievement gaps. Proof of such is evidenced by attaining adequate yearly progress (AYP) as demonstrated through an incremental rise in achievement levels indicated by annual state assessment results.

Assessment and accountability are essential components of education. However, as a result of NCLB, accountability is now limited to yearly achievement results of annual state summative assessments which are based on state standards that are often too long and too narrow to be thoroughly covered by educators. These state assessments are

usually composed of a small number of multiple-choice questions intended to assess learning of each state standard within each content area. Once the yearly summative assessment is over and evaluated, the results from these assessments take a lengthy amount of time to be returned and therefore limit the use of the data by educators to improve student success.

School systems and schools trying to meet the demands of NCLB in a positive, effective way sometimes resort to methods they believe will quickly increase student performance (Mizell, 2003). To compound this problem, various educational vendors and educational consultants have come forward with fast fixes to improve low student performance in response to educators' fears and intimidation of NCLB mandates. The demanding task of adapting to the requirements of and meeting the mandates of NCLB requires a reform like approach in which educators submit to a willingness to change, find and develop instructional strategies that lead to specific results, perform frequent assessments, and the willingness to reflect on the results to modify/adapt instruction accordingly. An example of such modification exists in the work of Gary Smith. Smith (2007) was teaching two undergraduate science classes consisting of 169 students at the University of New Mexico during the 2004-2005 school year. He was concerned with how effective his assignments were in promoting learning. Smith felt not all students learned equally from lab exercises and homework where group work was encouraged or required. He used StatView to perform a regression analysis. StatView showed exam scores did not correlate closely with lab scores ($r = 0.277$) or homework scores ($r = 0.365$). The median exam score was 12% points lower than lab scores over a five-year period. Smith decided to replace homework with online weekly assessment quizzes and

small group or pair-problem solving assignments. He strongly used the formative assessment component of descriptive feedback. After analyzing all results, Smith found online quiz scores and exam scores correlated strongly ($r = 0.86$). Those students that used the feedback given and revisited quiz questions were able to score 20% higher. The mean exam grade for those revisiting quiz questions was nine % higher than those that did not revisit missed quiz questions. Smith also found there were very strong correlations between exam grades and scores on infrequent written assignments assigned in class ($r = 0.93$ and 0.99). Smith found through the use of the concepts within formative assessments, he was able to help his students increase their achievement levels on his summative assessments.

Why use formative assessment?

Educators are constantly looking for educational practices to improve instruction and to measure student learning. The use of formative assessment is one such practice. Literature published in recent years promotes formative assessment as a way to improve student achievement. Black and Wiliam (1998) collected 580 research articles and compiled information from 250 of the articles. The goal of the study was to determine if improving formative assessment would raise standards, was there room for improvement, and could formative assessment be improved? Their research findings indicated all three questions could be answered with yes. In addition, Black and Wiliam found formative assessment produced learning gains, with effect sizes between 0.4 and 0.7 and the benefit was greatest for low achievers. Frequent assessment of student learning provides valuable information for educators at the student level, building level, and county level. A one size

fits all approach to educating students no longer works. The use of formative assessment data can be used to drive the instruction to meet individual student needs at the building and county level. Researchers have found formative assessment can have a great impact on student achievement, especially of low-performing students; however, there are too few teachers who know how to effectively use assessment information (Neill, 2006).

The general lack of assessment knowledge on the part of educators is very disheartening considering during the 1990's professional organizations associated with teaching adopted professional competence standards for teachers that included an assessment component. Within this component, teachers are to be trained to choose and develop proper assessment methods; to administer, score, and interpret the results; to link the results to specific decisions; to effectively assign grades; and to communicate student achievement (Stiggins, 2002). Many of these competence standards have been in place for over a decade but are not part of the preparation/pre-service programs for new teachers at many universities. Compounding the problem, many veteran teachers, administrators, and district level supervisors are not literate in assessment. If individuals associated with and responsible for student learning were trained in the use and benefits of formative assessments then schools, districts, and states could raise the achievement results for all students (Shepard, et.al, 2005).

The article by Black, Harrison, Lee, Marshall & Wiliam (2004)"Working Inside the Black Box: Assessment for Learning in the Classroom" details mini-experiments conducted with 19 teachers who used formative assessments to increase student achievement through the use of specific assessment methods such as effective questioning techniques, feedback through grading, peer and self-assessment, and the use

of summative tests administered in a formative manner. The researchers found the use of these assessment strategies led to an average effect size of 0.3 standard deviations demonstrating the impact teachers can have using the assessment methods listed above. These results are comparable to raising a school in the lower quartile of the national performance tables to the above average ranking (Black, Harrison, Lee, Marshall & Wiliam, 2004).

Bell and Cowie (2001) define formative assessment as "the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning" (p. 536). Formative assessments are yet another tool in the toolbox of educators enabling the identification of student understanding of standards taught. This tool also provides educators with the opportunity to make decisions regarding curriculum and instruction before the accountability stage of teaching. Campos and O'Hern (2007) performed a study involving first and fifth grade math students from within one school and within the researcher's own classrooms. Teacher observation, lack of knowledge and skill on students' part, and low test scores demonstrated to the researchers a problem existed. The researchers set out with a goal to examine the ability to increase student empowerment over a four month period of time by using formative assessment strategies. The researchers changed their instructional strategies to include beginning each math lesson with an awareness of the expected learning goals, the use of teacher-made assessments to determine growth toward the learning goals, the use of graphs to track student performance/growth, and the use of portfolios to organize all information gathered allowing students to perform self-assessment and peer-assessment aiding in the identification of student areas of need. The researchers also implemented a

15 minute intervention session each day used to re-teach learning goals to low performing individuals within a small group setting. Campos and O'Hern (2007) also incorporated family involvement with family letters being sent home at the beginning of each unit that described the expected learning goals and necessary vocabulary for the unit. The researchers collected data through teacher made assessments, surveys of students, and the analysis of classroom graphs. After analyzing all collected data, the researchers concluded the attitude toward mathematic instruction remained unchanged with most students exhibiting a positive attitude. However, data collected from student interviews indicated a 50% increase in the self-awareness of the expected learning goals, a 26% increase in the students ability to verbally express their learning goals, a 26% increase in the belief they could monitor their progress toward the learning goals, and a 16.5% increase in the students' awareness of their need for additional help, as well as, how to ask for and receive the help they needed. This study demonstrated the use of formative assessments enable students to take ownership of their own learning.

While formative assessment is a relatively new concept in education, educators are all too familiar with summative assessment. However, it must be understood there is a distinct difference between summative assessment and formative assessment. Instruction that has been guided through the use of formative assessment is different from summative assessments because it is focused on the student and pedagogy (Kaftan, Buck & Haack, 2006). Summative assessment occurs at the end of instruction and measures the knowledge gained. Formative assessments occur throughout the learning process informing the educator on how to better educate the student. It must be thought of as an assessment for learning that is occurring instead of learning that has already occurred. An

important aspect of formative assessments is they provide students with expected learning goals. As students work toward their goals, they are given descriptive feedback on where they stand in their learning process and how to improve or close their personal achievement gap. Descriptive feedback is used to identify problem areas and provide suggestions for improvement, which encourages students to focus on the task and not the right answer (Boston, 2002). Descriptive feedback, an essential component of formative assessment, should be focused on achievement gaps between the goal and current status allowing the student to participate in assessing his/her own learning (self-assessment) and self-exploration in identifying ways to help achieve their learning goals. The effectiveness of feedback is largely impacted by the quality of the feedback versus the quantity (Sadler, 1989).

In a study performed by Lesley Greer (2001), the performance of 217 college accounting students was evaluated to determine if changing the method by which a student is assessed during a course would have an impact on the student achievement of attaining expected learning goals. Greer created two groups in which one group received timely feedback prior to their summative assessment, allowing them an opportunity to get help with misunderstood concepts via the use of a provided tutor. The other group did not get the timely feedback and did not have time to get additional help with misunderstood concepts before the summative assessment. A method of triangulation was used to collect data (Greer, 2001). Quantitative differences in students were identified by assessment marks of students on an assigned module. Qualitative data were collected via the feedback from tutors on their perceptions of overall student performance. Through the use of student surveys, Greer found that 90.7% of the students felt the feedback on the in-

course assessment helped them to understand the problem areas they were having difficulty with and 94.6% of the students felt the in-course assessment helped them to better understand the actual content being taught in the module. Examination results and feedback from students and tutors showed overall student performance was better in the group receiving feedback and particularly better in "weaker" students. Fuchs, et.al, (1997) found both low-achieving students and students identified as learning disabled demonstrated an increase in student achievement with the use of frequent assessment feedback.

The benefits of using formative assessment are many. Effective school research indicates, in an effective school, there is frequent monitoring of student progress. Students are frequently assessed to measure progress and those assessments are used to improve the student performance as well as the curriculum (Lezotte, 2001). Students who are able to describe their learning goals score an average of 27% higher than students who cannot (Marzano, 2005). An additional benefit of formative assessments is the ability of students to build relationships with their teachers and administrators because the students know the educators are genuinely concerned about their learning and are willing to respond to needs of the student(s).

Formative assessment provides students the opportunity to gain ownership of their learning. When students know the expected learning goals, know where they stand in their learning process, and what it will take to get to the goal, there is an increase in student self-confidence. Confidence plays a key role in the development of student goal orientations. Husman, Brem and Duggan (2005) describe two types of student goal orientations. Mastery-oriented learners believe in the value of learning and are only

concerned with their own achievement level. These learners have the ability to learn on their own and they have the confidence to be risk-takers. Performance-oriented learners feel the need to compare themselves to others to determine their level of achievement; they are not as comfortable taking risks. In their study, Husman, et al., (2005) examined the impact formative assessment had in the development of student mastery and performance goal orientations. The researchers tracked 239 students participating in Reading Renaissance in grades three through six for one year. They used surveys occurring in the middle of the first and the third nine-week grading period to gather data. A repeated measures multivariate analysis of variance (MANOVA) was conducted. The mean scores of performance approach orientation decreased over the course of the year for all grades. Mastery scores both increased and decreased depending on grade level but all were above 4.00. Mastery goal orientation did not change, remaining stable around 4.40. Results indicated performance-oriented students thought their teachers were performance-oriented ($r = 0.34$, $p < 0.001$). The more mastery-oriented students believed their teachers to be, the more mastery-oriented the students were ($r = 0.38$, $p < 0.001$). The more performance-oriented the teacher seemed, the more likely the student was to be performance-oriented.

Summative state assessments demonstrate to educators what a student knows on a specific day and under specific conditions. Formative assessments allow educators to monitor and adjust instruction over time to increase student achievement. This allows the educator to work from a "smarter" perspective and not a "harder" perspective. Formative assessments ensure the understanding of a concept and not just the memorization of facts or simple recall of information. Students' not mastering/understanding material during

learning can be intervened upon with re-teaching or varied instructional strategies, which allow the educator to customize curriculum and instruction at the individual level if necessary. "Formative assessment helps support the expectation that all children can learn to high levels and counteracts the cycle in which students attribute poor performance to lack of ability and therefore become discouraged and unwilling to invest in further learning" (Boston, 2002, p. 2).

Why is formative assessment not being used?

Formative assessment programs are deemed by beginning and veteran educators as too intimidating and too time consuming and many educators do not use formative assessment primarily because they lack assessment knowledge or they have not been trained to properly use it. "Thus teachers have to take risks in the belief that such investment of time will yield rewards in the future, while 'delivery' and 'coverage' with poor understanding are pointless and can even be harmful" (Black & Wiliam, 1998). The mandates of NCLB have pushed educators to move away from designing assessment strategies that truly show what a student has learned to a format that is similar to the summative state assessment to which the educator will be held accountable. Better use of formative assessment data requires educators and instructional leaders participate in sustained and ongoing professional development in assessment. Educators entering the field of education should be competent in the use of assessment to make appropriate instructional decisions and increase student achievement. Pre-service teachers should experience exposure to the use of formative assessment already being used in instruction in an observational setting (Shepard, et al., 2005).

Heritage (2007) believes there is too much reliance on summative assessment results and many educators are still operating under the guidelines that if a student can score a passing grade or higher on an assessment, then the student understands the concepts presented. However, sometimes these same students cannot explain what they have learned when probed for more details or are questioned about the application of the knowledge gained. For those teachers that have been teaching pre-NCLB mandates, the move to formative assessments requires change. Change in the amount of time to prepare for class, change in instruction, and most importantly, a change in educational philosophy. "What is needed is a culture of success, backed by a belief that all pupils can achieve. In this regard, formative assessment can be a powerful weapon if it is communicated in the right way" (Black & Wiliam, 1998). Effective schools research asserts that all children can learn and the school is responsible for creating a culture in which students' master the curriculum (Lezotte, 2001).

How can formative assessment be better used?

Student achievement gains that result from some of the most powerful instructional interventions can be met or even surpassed with the use of correctly implemented formative assessment (Shepard, et al., 2005). Schmoker (2003) believes using the goals that they have established, teachers can meet regularly to improve their lessons and assess their progress using another important source: formative assessment data…formative data enable the team to gauge levels of success and to adjust their instructional efforts accordingly (p. 23).

The use and benefit of formative assessments should not be limited to use by individual educators. Formative assessments provide opportunities for teaming within a faculty. Team members experiment with various assessment techniques and then share those ideas with team members. The team can work collaboratively to gain knowledge about formative assessment, create intervention strategies for identified students having difficulty, and together achieve a common goal of increasing student achievement. Formative assessment should be a collaborative effort involving all stakeholders. Parents should be invited to participate in the education of their child. Parents aware of the student's learning goals and the progress the student is making toward those goals have a better understanding that learning involves more than simply passing a test. It is beneficial for parents to see their child be successful and see the ownership their child has taken in his or her own learning. This may require a formative assessment training session for parents at the beginning of a school year. In addition to teacher, administrator, and parent training, students should have an understanding of the components of formative assessment and their individual role in the process. Educators should strive to instill a self-motivating desire for the student to want to learn and understand they are capable of learning and success is within reach. Implemented correctly, formative assessments will increase the student achievement levels on all summative assessments. Educators know there is a need for student achievement scores to increase and based on current research, the most effective way to do this is through the improvement of classroom instruction. Therefore, instructional leaders should be implementing techniques that provide opportunities for this to occur.

With the increase of confidence and a clear focus on expected learning goals, the student becomes a data-driven decision-maker and is able to see personal gains and areas in need of improvement. The student has input on how to close any personal achievement gaps that may exist based on feedback from teachers, self-assessment, or peer assessment. With increased ownership, students are motivated to learn material in order to understand the material and not simply to make a grade. In addition, there is an increase in the internal motivation of the student to continue to learn, to achieve the current learning goal, and for future learning to occur. "One of the strongest motivators is the opportunity to look back and see progress" (Chappuis, 2005, p. 42).

What is data-driven decision-making?

The use of formative assessment data can be used by educators to make data-driven decisions. "It has become a school-reform mantra that is celebrated but widely misunderstood, and is often ignored (despite its hype) or actively feared" (Doyle, 2003, p. 19). Data-driven decision-making (DDDM) in education refers to the collection and analysis of data by educators to make decisions with the intent of increasing the achievement levels of not only students but schools. DDDM took root in the business world but made its way into education in the 1980's in light of the debate regarding measurement-driven instruction (Marsh, Pane & Hamilton, 2006). "The theory of action underlying NCLB requires that educators know how to analyze, interpret, and use data so that they can make informed decisions in all areas of education, ranging from professional development to student learning" (Datnow, Park & Wohlstetter, 2007, p. 10). "Despite the increased interest in DDDM, there has been little research on the

precise effects of using DDDM on student achievement and other important educational outcomes" (Demboskey, Pane, Barney & Christian, 2005, p. 5).

Date-driven decision-making and education

Demboskey, et al., (2005) examined the current practices and factors enabling DDDM through a qualitative case study approach in six school districts and one independent charter school in southwestern Pennsylvania in their study Data Driven Decisionmaking in Southwestern Pennsylvania School Districts. The researchers interviewed teachers, principals, superintendents, district office staff, IU (Intermediate units) representatives, state officials, and consultants (more than 100 individuals total). In addition, 26 superintendents were surveyed. DDDM practices in Southwestern Pennsylvania demonstrated a large variation in persons responsible for analyzing, planning, and using data. Persons responsible varied from teachers, district and building administrators, and district support staff. Sixty-one point five percent of superintendents surveyed strongly agreed all personnel listed above should have responsibility for data analysis and teachers and building administrators need to be able to access, manipulate, and analyze data for themselves. Interviewed teachers reported their use of student achievement data to adjust instruction at three levels – whole class instruction, group instruction, and individualized instruction. All districts examined in the case study were analyzing state achievement data and adjusting curriculum and instructional strategies. Some districts and schools even adopted new curricula and supplemental programs due to their data analyses. One hundred percent of the superintendents surveyed reported experience using data through creating mandated performance reports for the state and

federal government. When asked about data communication with the community, the superintendents responded positively toward their willingness to voluntarily report data to the community allowing for self-reflection by the district and increasing communication with stakeholders via a district website or other web-based sources. The study also demonstrated the need for DDDM to be largely determined by district, state, and federal policies. However, the researchers found several factors identified by participants as enabling the use of DDDM which included policies required at the local, state, and federal level, data sources, technological and financial resources, and cultural development (within the school) toward DDDM. The districts in this case study exhibited a strong interest in DDDM and the instructional leaders within the districts and schools worked hard to begin establishing a culture for the use of DDDM. The districts were identified by the researchers as being in the beginning phase of using DDDM. Many of the educators interviewed saw the value of data but felt they lacked the proper training, technology, and time to work individually or collaboratively to analyze and interpret data and then design intervention strategies. Findings from this study can be applicable to many districts wanting to implement DDDM and include the need for the following:

1. The development of goals and objectives for using data and the establishment of expectations for the use of DDDM within a district;
2. The alignment of instruction, curriculum and assessment with state standards;
3. Focus on the quality of data versus the quantity;
4. Assurance that schools have the necessary tools for effective DDDM;
5. The implementation of assessments designed for diagnostic purposes;

6. The adoption of technological resources such as data systems and software allowing for the use of more frequent, formative assessments;

7. Professional development in DDDM for teachers and administrators;

8. Time for teachers to examine data and to collaborate;

9. Teacher input on key decisions related to DDDM;

10. An increase in the level of communication regarding data with stakeholders (Demboskey, et al., 2005).

What is limiting the use of data-driven decision-making in education?

Many educators desire to implement the use of data-driven decision-making due in part to the accountability mandates of *No Child Left Behind* and therefore, many states, districts, and schools have started to gather a lot of data. It is an overwhelming task to analyze data in schools with limited resources which leads to data that sits on a shelf untouched. Additionally, educators record daily assessment data with no thought of finding a common thread in the data (Cromley, 2000). In such situations, many educators feel they are "data-rich" but "information-poor". "If knowledge is power, then studying the current abilities, skills, attitudes, and learning styles of students empowers educators to adjust the curriculum to achieve whatever goals the school and district have chosen" (Johnson, 2000, p. 1).

"One specific type of data-based decision-making that shows promise for helping schools dramatically increase student achievement is the use of assessment data to drive instructional improvement" (Protheroe, 2001, Introduction section, ¶ 1). In the study, Using Test Score Data in the Classroom, by Stecher and Hamilton (2006), the researchers

used stratified random sampling to select approximately 100 schools from the states of California, Georgia, and Pennsylvania. The schools represented 25 districts and were elementary and middle schools. Superintendents, principals (approximately 262), and all regular math teachers of grades three, four, five, seven, and eight (approximately 2628) were surveyed. In addition to surveys, the researchers performed case study visits to two schools in each state. Stecher and Hamilton, (2006) wanted to determine how teachers use test results, what factors influence the use of tests, and why data use varies so much. Their findings, in reference to the use of test results, found one-half of math teachers in all three states indicated they use test results weekly to individualize instruction based on the needs of their students. A smaller percentage indicated they use test results weekly to identify curriculum issues. One-quarter of the teachers in all three states increased the frequency of test usage for the purpose of diagnosis and instructional planning from the previous year. Two-thirds of the teachers in all three states indicated they focus more on state standards and look for best practices in teaching due to the accountability tied to the summative state assessments. In some cases, within California and Pennsylvania, teachers reported using test results to identify their own areas of deficiency in content knowledge as opposed to individualizing instruction for students. Principal survey results from all three states indicated over 80% of principals believed their teachers reviewed state test data to improve instruction in the classroom. Principal survey results also indicated about two-thirds found state test results to be helpful in designing school improvement plans, instructional strategies, curriculum issues, and professional development. Stecher and Hamilton, (2006) found factors influencing test data usage included: how the data was reported, how valid the tests appeared to educators, and the

timeliness of the data being given to educators. Approximately 70% of teachers reported they received their reports in a timely manner. Variation was seen in the results of the same question asked of principals. Forty-seven percent of principals in Georgia, 37% of principals in California, and 19% of principals in Pennsylvania felt they were given information in a timely manner regarding their school's performance. Clarity of the test results was agreed to be satisfactory by approximately 80% of the teachers and 83% of the principals from all three states. In terms of validity, Stecher and Hamilton, (2006) found that about one-quarter to one-third of teachers found the test to be aligned with their curriculum. Approximately 55% of the principals surveyed from all three states agreed the state test results were reliable at reflecting their students' achievement level. Support in terms of workshops and professional development was found to be helpful by most educators. However, educators reported pacing guides and data analysis software not to be helpful or were simply lacking in the support offered. In terms of variation in the use of data, Stecher and Hamilton, (2006) found Georgia educators appear to be more receptive to the use of data, perhaps due to their states approach to standards-based accountability. In addition, the results of the study indicate it is difficult to impose a culture of being data-driven. Stecher and Hamilton believe positive outcomes are a result of the better focus of instruction on state standards and to the gaps in student learning. They suggest the use of formative classroom assessment or progress testing over annual state tests be used to improve data-driven decision-making. "Much of the literature on DDDM focuses on the growing importance and use of formative assessments that provide teachers with frequent, ongoing information on how well students are mastering certain content areas" (Dembosky, et al., 2005, p. 6).

On a daily basis, educators gather various forms of data (classroom assessments, formative assessments, behavior notes, classroom projects, state summative assessment scores, and homework). However, many educators do not use the data to drive the decisions they need to make. "So far, despite data-driven decision-making by many vocal proponents, it is equally clear that the message has not yet gotten to the front lines" (Doyle, 2003, p. 20). Mistakenly, state summative assessment data appears to be the most popular and used source of data used to make educational decisions. However, using data that typically arrives after the end of the school year or after a particular group of students move on to another grade level, or perhaps school, makes it very difficult to alter instructional strategies. In the study by Marsh, et al. (2006), more than 80% of superintendents surveyed in California, Pennsylvania, and Georgia felt more frequent assessments, compared with state assessment results, were more useful in their decision-making process and a beneficial tool in the school improvement process as well as guiding instructional decisions. "Effective educators make effective decisions, decisions based on accurate information" (Johnson, 2000, p. 1).

Educators are required to set goals, to identify areas of need in curriculum, and to target instructional strategies. "Yet educators have historically relied less on data to guide their practice than they do on intuition, teaching philosophy, or personal experiences" (Cromley, 2000, p.3). The use of data enables educators to identify students that need extra help in order to reach the proficient status on the state summative assessment and to set goals or differentiate instruction for these students. Marsh, et al., (2006) discovered such students are referred to as "bubble kids". These are students who sit on the fence between proficient and non-proficient. It was discovered more than 75% of principals

surveyed in three states report the encouragement of teachers to specifically focus on these students in an attempt to meet NCLB mandates (Marsh, et al., 2006).

The use of data seems to vary from school to school and district to district. Marsh, et al., (2006) found a large variation of data use within schools that suggested some teachers frequently use data to drive instruction, while others are oblivious to the trend. The largest barrier contributing to the use of data is the lack of training. Many educators do not feel they are trained to properly assess students and then to use the data to alter instruction (Crowley, 2000). Marsh, et al., (2006) found the use of data by educators to make decisions is also influenced by the ability to access data, the viewed validity of the data, the motivation behind the use of data (extrinsic vs. intrinsic), the immediacy in the timeliness of data, the capacity and support of faculty and staff, the pressure to follow pacing guides, the lack of time allotted to gather, analyze and decipher the data, and the culture and leadership of the school and district.

How can the use of data-driven decision- making be increased?

Data is a useful tool for aiding instructional decisions. The use of various forms of data has the potential to improve teaching and thus increase student learning. "The mere presence of data does not automatically imply that usable information is available; educators need support to use these data to the fullest extent" (Wayman, 2005, p. 296). For DDDM to be effective, certain support for educators is needed. Professional development and training must occur in order for educators to understand and have the confidence to accurately interpret and use the data. Bettesworth, (2006) conducted a study to determine if pre-service administrators, from the Initial Administrative Licensure

Program at the University of Oregon, could improve their ability to use data to make instructional decisions by participating in a statistical analysis training session aimed at teaching educators how to analyze and interpret data and make instructional decisions. This study used a nonequivalent comparison group, pre-test, post-test design. A convenience sample drew two groups of educators with similar demographics and desires to further their careers in a leadership capacity within their district. The treatment group had 31 participants and the comparison group had 25 participants. The participants in the treatment group were provided with an overview of DDDM concepts, prior to the statistical analysis training session, which occurred in a three part seminar using computer training modules. In addition to the pre-test and post-test data, the researcher also collected data from surveys, case studies, and probes. An Analysis of Variance (ANOVA) Test was used to compare the overall pretest and posttest scores for the treatment and comparison group and for a matched sample. An independent sample *t*-test revealed no significant difference between groups on pre-test scores $t(32, 25) = .42$, $p > .05$. The results of the ANOVA for the overall pre-post test revealed a significant interaction effect, $p < .0001$. The intervention accounted for 61% of the variance. A follow-up Cohen's *d* indicated a huge effect size of 1.93. The results from surveys, case studies, and probes further supported these findings. Therefore, the researcher concluded training in statistics significantly increases pre-service administrator's ability to use data for decision-making. However, the case studies also revealed the need for participants to practice their skills and experience success in order to transmit their skills in a real-world setting. "Without continuous professional development to build administrators' and teachers' understanding of their data and build their repertoire of improvement strategies,

new accountability systems provide insufficient approaches to improving student achievement" (Bettesworth, 2006, p. 37).

Instructional leaders must support and mentor those attempting to understand and implement the use of DDDM. In two of the three districts studied by Marsh, et al., (2006), 75% of teachers felt instructional leaders helped them make instructional decisions based on analyses of district and state assessment scores. The third district reported 50% of teachers felt this way. The Center on Educational Governance at USC collaborated with New Schools Venture Fund in 2005-2006 to study the relationship that exists between the central office and the school to build capacity in the effective use of data. Datnow, et al., (2007) performed a qualitative case study of two public school districts and two nonprofit charter management organizations, identified as leaders in performance-driven instruction. Various central office staff, principals, and at least five teachers from varying grade levels were interviewed and informal observations were also made. Datnow, et al., (2007) found the key characteristics of performance-driven school systems to be:

1. Establishing a basis for data-driven decision-making through the establishment of student achievement goals and the development of curriculum used through the district;
2. Development of a culture receptive to the use of data and the improvement of curriculum by the creation of expectations and norms throughout the system and school level and the creation of a mutual accountability measure in the school-system relationship;
3. The investment of a user-friendly data management system;

4. The utilization of individuals to mentor data analysis and use at the system and school level;

5. The selection of appropriate data that is diverse in student achievement and instructional practices;

6. The formation of an interim assessment aligned to standards used throughout the system and the accumulation and application of data;

7. The creation of capacity for DDDM at the school level through the implementation of professional development, time for collaboration with peers, support with external sources, and the use of extrinsic motivators (rewards/incentives);

8. The analysis and use of data to increase educator performance by providing tools aiding educators in the communication and implementation of data, the tools needed to monitor progress towards goals, and to employ students in the use of data for self-improvement.

"When all the school's educators examine the efforts and initiatives of the school through the lens of their impact on learning, the structure and culture of the school begin to change in substantive ways" (DuFour, 2002, p. 13). If instructional leaders want to improve the use of DDDM, adequate time must be allowed to collect and analyze data. In addition, time must be allocated for collaboration on the use of the data, as well as, focused on how to analyze data, identify problem areas/trends, and create solutions to identified problems. Implementation of technology and data systems must be implemented to allow access to data and the ability to analyze the data in various ways.

Protheroe, (2001) suggests an educator within the school building be trained to work with teachers to address grade-level data or subject area data.

> Educators must be encouraged to take on the mindset of researchers, developing an understanding of research and statistics. Effective teaching strategies are typically born from extremely large research studies that do not always correlate with what actually occurs in a typical classroom. Teachers must begin to perform research projects within their own classrooms to find out what works best with the group of students they are currently working with (Johnson, 2000).

Formative assessment enables educators to monitor student learning as it takes place. Data-driven decision-making enables educators to use the data from formative assessments to make effective decisions on implementing intervention strategies for students experiencing difficulty in learning the assigned standards, the identification of problems within curriculum and instruction, and the need for professional development before the state summative assessment occurs. Combined and used correctly, the use of formative assessments and data-driven decision-making has the potential to enable students to successfully achieve the learning expectations set forth in the state standards. This, in turn, increases student achievement scores, school achievement scores, and state achievement scores enabling schools, school districts, and states to meet the mandates of NCLB. Most importantly, their use creates an effective school focused on improving the quality of education being offered to all students. When a school establishes a culture in which teachers are committed to student learning and a focus is made toward intervention instead of remediation, a professional learning community begins to form. DuFour,

(2004) defines a professional learning community as one in which teachers ensure students learn, a culture of collaboration exists, and there is a focus on results. In order for a school to change its culture, the focus must move from teaching to learning. Educators must be willing to work hard and ensure that indeed, no child is left behind.

> President Lincoln began the modern practice of collecting education data at the national level 150 years ago, and with *No Child Left Behind*, data will have to be used, not just collected. It will be used to plot progress (or lack thereof); plan and execute instructional interventions; report results; as well as hold students, teachers, administrators and school systems accountable (Doyle, 2003, p.4).

CHAPTER III

Methodology

Overview

No Child Left Behind requires the use of effective classroom instruction in order to help students meet adequate yearly progress. This field study compared the achievement test results of students instructed with data driven curriculum against students not receiving the same type of instruction. A comparison was also performed to determine if an impact on TCAP achievement scores of one particular student socioeconomic group or one particular ethnic group occurred through the use of formative assessment data to drive instruction and was the effect greater on one particular gender.

Research Design

This field study was a descriptive study of the effect of data-driven instruction on reading/language arts and math TCAP achievement scores of students in grade six through eight.

Participants

Authorization and access was given to collect and analyze the reading/language arts and math TCAP scores of approximately 275 students in grade six through eight.

Student scores, obtained from archival sources, were recorded for a total of four years for the same group of students. No record was made of the students' names. The collection of data included a comparison of one year of TCAP achievement data in which there was no use of data to drive curriculum and instructional decisions and the following three years of TCAP achievement data, in which formative assessment data was used to drive instruction.

Instrument

In this field study, there was an analysis of grade six through eight students' reading/language arts and math TCAP achievement scores instructed with supplemental information provided through the use of formative assessments verses students not instructed using supplemental formative assessment data. The formative assessment tool used to supplement instruction was ThinkLink.

Procedure

Approval was sought and granted from the Austin Peay State University Institutional Review Board and the Robertson County Board of Education to study the TCAP achievement scores of students in grade six through eight. During the 2003-2004 school term, students were not instructed with the use of formative assessment data in reading/language arts and math. The above mentioned scores were compared to the achievement scores of the same students who were then instructed using formative assessment data during the 2004-2005, 2005-2006, and 2006-2007 school terms. The TCAP achievement scores of students tested during the 2004, 2005, 2006, and 2007 testing periods were recorded. A comparison was made between the following years:

2004-2005, 2004-2006, 2004-2007, 2005-2006, 2005-2007, and 2006-2007. Approval was sought and granted from the local board of education to access the records of students enrolled during the 2003-2007 school terms. The names of students were not recorded with any of the information or data analyzed in the study.

Data Analysis Plan

The information gathered from the TCAP achievement scores were compiled and analyzed. Data was entered into a computer using a statistical software application and statistical procedures were performed using SPSS statistical software. An ANOVA was used to test the null hypotheses at the .05 level of confidence for significant differences in the TCAP achievement scores of students instructed with the use of formative assessment data against students instructed with traditional methods of instruction. An analysis was performed to determine if using formative assessments to drive instruction had any impact on the TCAP achievement scores of a particular socioeconomic group or ethnic group. An analysis was also performed to determine if there were significant achievement trends based on gender.

CHAPTER IV

Data and Results

Demographics

This study analyzed the math and reading/language arts TCAP scores of 275 students grades six through eight, in a southern, rural school district. The three grade levels were using one year (2004) in which students were not instructed with the use of formative assessment data compared to three years (2005, 2006, and 2007) when formative assessment data was used by the educator to make curriculum and instructional decisions. The study included 125 male students and 150 female students. Their math and reading/language arts TCAP scores for 2004, 2005, 2006, and 2007 were recorded. The students' socioeconomic status and ethnicity were also recorded. Socioeconomic status was established by whether or not the student qualified for free or reduced lunch during the 2006-2007 school term.

The computer program, SPSS, was used to analyze the data gathered from the TCAP scores for years 2004-2007. Table 4.1, 4.2, 4.3, and 4.4 list the number of students in each group. The dataset contained information including the following variables:

1. Grade – Participants were from one of three grade levels. Neither student nor teacher names were recorded.

2. Subject- Participant test scores were recorded for math and reading/language arts.
3. Gender- Participants were identified by gender.
4. Socioeconomic Status- Participants were identified through their 2007 qualification for free or reduced lunch.
5. TCAP scores (2004) – TCAP scores for 2004 were recorded.
6. TCAP scores (2005) – TCAP scores for 2005 were recorded.
7. TCAP scores (2006) – TCAP scores for 2006 were recorded.
8. TCAP scores (2007) – TCAP scores for 2007 were recorded.

Table 4. 1
Number of students scores analyzed

Grade	Year	*n*
6th	2004	83
	2005	84
	2006	86
	2007	87
7th	2004	89
	2005	94
	2006	95
	2007	95
8th	2004	89
	2005	92
	2006	93
	2007	93

N=275

Table 4. 2
Socioeconomic groups

Grade	Normal	Reduced	Free	Total
6th	62	7	18	87
7th	74	6	15	95
8th	71	4	18	93
Total	207	17	51	275

Table 4. 3
Gender groups

Grade	Male	Female	Total
6th	45	42	87
7th	39	56	95
8th	41	52	93
Total	125	150	275

Table 4. 4
Ethnic groups

Grade	White	Black	Hispanic	Asian	Total
6th	84	1	2	0	87
7th	89	3	3	0	95
8th	87	2	2	2	93
Total	260	6	7	2	275

Hypothesis One:

Null: There is no statistically significant difference between the achievement scores of students instructed with the use of formative assessment data and those not instructed with the use of formative assessment data.

One way repeated measures ANOVA with year as a within-subjects factor indicated year was a significant predictor of reading/language arts scores $F(2.84, 722.2)=51.1$, $MSE=72.142$, Huynh-Feldt $p <.001$. Pairwise comparisons via dependent samples t tests were conducted to compare the reading/language arts scores across years. Results are presented in Table 4.5.

Table 4. 5
Pairwise comparisons of reading scores among the years
Dependent Sample t test results

Year Comparison	*t*	*df*	*p-value*
2004 vs. 2005	-7.27	255	0.000
2004 vs. 2006	-9.52	259	0.000
2004 vs. 2007	-10.27	260	0.000
2005 vs. 2006	-3.20	268	0.002
2005 vs. 2007	-4.51	269	0.000
2006 vs. 2007	-2.04	273	0.042

$p < .05$

One way repeated measures ANOVA with year as a within-subjects factor indicated year was a significant predictor of math scores $F(2.65, 664.7)=54.77$, $MSE=83.7$, Huynh-Feldt $p < .001$. Pairwise comparisons via dependent samples *t* tests were conducted to compare the math scores across years. Results are presented in Table 4.6.

Table 4. 6
Pairwise comparisons of math scores among the years
Dependent Sample t test results

Year Comparison	*t*	*df*	*p-value*
2004 vs. 2005	-5.04	253	0.000
2004 vs. 2006	-10.43	256	0.000
2004 vs. 2007	-9.11	258	0.000
2005 vs. 2006	-6.93	266	0.000
2005 vs. 2007	-6.07	268	0.000
2006 vs. 2007	0.78	271	0.433

$p < .05$

Overall, it was concluded formative assessment did improve both math and reading scores. Both scores increased every year, with an exception of 2007, in which, the scores reached a plateau and did not change noticeably from the previous years.

Formative assessment did improve reading/language arts and math scores across years when we analyze a sample of students without subdividing them into levels of

between-subject factors. Students had significantly lower reading scores before formative assessment began (2004) than any years after (2005, 2006, and 2007). The results also indicated the increase of reading/language arts and math scores occurred every year after formative assessment began except in year 2006-2007.

Factors such as grade, socioeconomic status, gender, ethnicity, and special education status were examined for interaction with the within-subjects factor (year) in a separate factorial design. Descriptive statistics for grade across year can be found in Tables 4.7 and 4.8.

Table 4. 7
Descriptive statistics for reading scores

Grade	year	*n*	*M*	*SD*
6th	2004	83	49.58	22.29
	2005	84	54.64	17.11
	2006	86	55.80	20.73
	2007	87	59.33	16.79
7th	2004	89	51.47	18.15
	2005	94	59.15	19.80
	2006	95	60.91	16.60
	2007	95	58.80	15.81
8th	2004	89	52.92	18.00
	2005	92	56.42	17.02
	2006	93	58.73	16.16
	2007	93	61.35	14.91

N= 275

Table 4. 8
Descriptive statistics for math scores

Grade	year	*n*	*M*	*SD*
6th	2004	83	54.73	23.39
	2005	83	57.07	19.93
	2006	86	56.98	19.20
	2007	87	58.34	19.68
7th	2004	87	54.26	20.66
	2005	94	56.96	14.97
	2006	93	66.95	17.30
	2007	94	66.40	14.81
8th	2004	89	53.45	17.51
	2005	92	61.50	17.59
	2006	93	63.78	15.83
	2007	93	61.82	18.48

N=275

Two-way ANOVA with grade (6th, 7th, 8th) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the reading/language arts test. A familywise alpha of .05 was used for all analyses. There was a significant interaction between factors $F(5.68, 715.01)=3.145$, $MSE=71.1$, Huynh-Feldt $p <.006$. The main effect of the within-factor (year) was significant $F(2.84, 715.01)=52.03$, Huynh-Feldt $p <.001$. The main effect of the between-subjects factor (grade), however, was not significant $F(2, 252) =.662$, $MSE=1058.25$, $p =.517$.

Simple effect Welch's ANOVAs were conducted to compare the average reading/language arts score among the grades for each year 2004, 2005, 2006, and 2007. The results supported the conclusion drawn from the main between-subjects effect from factorial design. Grade was not a significant predictor of reading scores for every level of the within-subjects factor (year): 2004 $F(2, 168.72)= .581$, $p=.560$, 2005 $F(2, 177.31)= 1.329$, $p=.267$, 2006 $F(2, 176.03)= 1.652$, $p=.195$, 2007 $F(2, 179.67)= .719$, $p=.489$.

Pairwise comparisons also did not yield any significant results. The results are presented in Table 4.9.

Table 4. 9
Pairwise comparisons of reading scores among grades by year: Games-Howell Multiple Comparisons Results

Year	Grade Comparison	Mean Difference	*S.E.*
2004	6th vs. 7th	-1.89	3.11
	6th vs. 8th	-3.34	3.10
	7th vs. 8th	-1.45	2.71
2005	6th vs. 7th	-4.51	2.77
	6th vs. 8th	-1.78	2.58
	7th vs. 8th	2.73	2.71
2006	6th vs. 7th	-5.10	2.81
	6th vs. 8th	-2.93	2.79
	7th vs. 8th	2.17	2.39
2007	6th vs. 7th	0.53	2.42
	6th vs. 8th	-2.02	2.37
	7th vs. 8th	-2.55	2.24

Simple effect one-way repeated measures ANOVAs were conducted to compare the average reading score across years for each grade: 6th, 7th, and 8th. Year of measurement was a significant predictor of reading scores for all grades: 6th grade, $F(2.74, 213.9)=17.9$, $MSE=78.64$, Huynh-Feldt $p<.001$, 7th grade, $F(2.87, 249.6)=20.51$, $MSE=74.19$, Huynh-Feldt $p<.001$, and 8th grade, $F(2.72, 236.3)=19.8$, $MSE=65.57$, Huynh-Feldt $p<.001$.

Pairwise comparisons were conducted using the dependent samples *t* test and are reported in Table 4.10.

Table 4. 10
Pairwise comparisons of reading scores among the years by grade: Dependent Sample *t* test results

Grade	Year Comparison	*t*	*df*	*p-value*
6th grade	2004 vs. 2005	-3.54*	79	0.001
	2004 vs. 2006	-4.41*	81	0.000
	2004 vs. 2007	-6.11*	82	0.000
	2005 vs. 2006	-1.73	82	0.087
	2005 vs. 2007	-4.06*	83	0.000
	2006 vs. 2007	-2.79*	85	0.006
7th grade	2004 vs. 2005	-5.65*	87	0.000
	2004 vs. 2006	-7.50*	88	0.000
	2004 vs. 2007	-5.23*	88	0.000
	2005 vs. 2006	-1.69	93	0.094
	2005 vs. 2007	0	93	1.000
	2006 vs. 2007	2.16	94	0.033
8th grade	2004 vs. 2005	-3.30*	87	0.001
	2004 vs. 2006	-4.73*	88	0.000
	2004 vs. 2007	-6.44*	88	0.000
	2005 vs. 2006	-2.13	91	0.036
	2005 vs. 2007	-4.78*	91	0.000
	2006 vs. 2007	-2.90*	92	0.005

$p < .05$

Two-way ANOVA with grade (6th, 7th, and 8th) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the math test. A familywise alpha of .05 was used for all analyses. There was a significant interaction between factors $F(5.4, 667.4)=10.85$, $MSE=76.7$, Huynh-Feldt $p < .001$. The main effect of the within-factor (year) was significant $F(2.68, 667.4)=56.3$, Huynh-Feldt $p < .001$. The main effect of the between-subjects factor (grade), however, was not significant $F(2, 249)=.923$, $MSE=1089.9$, $p = .399$.

Simple effect Welch's ANOVAs were conducted to compare the average math score among the grades for each year 2004, 2005, 2006, and 2007. Grade was a significant predictor of 2006 math scores, $F(2, 176.44)= 6.733$, $p = .002$, and 2007 math

scores, $F(2, 175.64) = 5.084$, $p = .007$. However, grade was not a significant predictor of math scores for 2004 math scores $F(2, 166.36) = .092$, $p = .912$, and 2005 math scores $F(2, 171.04) = 2.041$, $p = .133$.

Pairwise comparisons were conducted and indicated the 7th grade had significantly higher math scores in year 2006 and 2007 than 6th or 8th grades. Results are reported in Table 4.11.

Table 4. 11
Pairwise comparisons of math scores among grades by year:
Games-Howell Multiple Comparisons Results

Year	Grade Comparison	Mean Difference	*S.E.*
2004	6th vs. 7th	0.47	3.39
	6th vs. 8th	1.29	3.17
	7th vs. 8th	0.81	2.89
2005	6th vs. 7th	0.11	2.68
	6th vs. 8th	-4.43	2.85
	7th vs. 8th	-4.54	2.40
2006	6th vs. 7th	-9.97*	2.74
	6th vs. 8th	-6.81	2.64
	7th vs. 8th	3.16	2.43
2007	6th vs. 7th	-8.06*	2.60
	6th vs. 8th	-3.47	2.85
	7th vs. 8th	4.59	2.45

*significant at .05 familywise alpha

Simple effect one-way repeated measures ANOVAs were conducted to compare the average math score across years for each grade: 6th, 7th, and 8th. Year of measurement was a significant predictor of math score for 7th grade, $F(2.69, 228.2) = 53.01$, $MSE = 73.75$, Huynh-Feldt $p < .001$, and 8th grade, $F(2.97, 258.3) = 37.48$, $MSE = 50.75$, Huynh-Feldt $p < .001$. However, the 6th grade did not differ on math scores across the years $F(2.42, 186.57) = 1.648$, $MSE = 113.84$, Huynh-Feldt $p < .189$. Pairwise

comparisons were conducted using the dependent samples *t* test and are reported in Table 4.12.

Table 4. 12
Pairwise comparisons of math scores among the years by grade: Dependent Sample *t* test results

Grade	Year Comparison	*t*	*df*	*p-value*
6th grade	2004 vs. 2005	-0.58	78	0.562
	2004 vs. 2006	-1.52	81	0.133
	2004 vs. 2007	-1.99*	82	0.049
	2005 vs. 2006	-0.88	81	0.380
	2005 vs. 2007	-1.7	82	0.094
	2006 vs. 2007	-1.04	85	0.300
7th grade	2004 vs. 2005	-2.11*	86	0.038
	2004 vs. 2006	-9.31*	85	0.000
	2004 vs. 2007	-8.35*	86	0.000
	2005 vs. 2006	-9.25*	92	0.000
	2005 vs. 2007	-9.06*	93	0.000
	2006 vs. 2007	0.5	92	0.617
8th grade	2004 vs. 2005	-7.11*	87	0.000
	2004 vs. 2006	-10.12*	88	0.000
	2004 vs. 2007	-7.25*	88	0.000
	2005 vs. 2006	-2.45*	91	0.016
	2005 vs. 2007	-0.65	91	0.514
	2006 vs. 2007	2	92	0.049

$p < .05$

In determining if there was a statistically significant difference between the achievement scores of students instructed with the use of formative assessment data and those not instructed with the use of formative assessment data, analyses indicated $p < .001$ in both the content areas of reading/language arts and math. Analyses of data indicate the main effect of the within-factor (year) and between-subjects factor (grade) were both significant with $p < .001$, which was less than the alpha of .05. Therefore, the null hypothesis was rejected.

Hypothesis Two:

Null: There is no statistically significant difference in student achievement based on student socioeconomic status due to the use of formative assessment data to drive curriculum and instructional decisions.

Tables 4.13 and 4.14 detail the descriptive statistics for reading/language arts and math scores of students based on their 2006-2007 SES classification. An ANOVA was used to analyze data for students, grouped according to their qualification for normal, reduced, or free lunch, to determine if any variation in TCAP scores occurred as a result of instruction based on the use of formative assessment data.

Table 4. 13
Descriptive statistics for reading scores across SES

Grade	year	*n*	*M*	*SD*
Normal	2004	199	54.48	18.67
	2005	206	59.27	17.91
	2006	206	61.66	17.25
	2007	207	62.28	16.09
Reduced	2004	17	36.65	20.67
	2005	16	43.38	18.55
	2006	17	42.47	20.87
	2007	17	47.59	10.43
Free	2004	45	43.13	17.95
	2005	48	50.79	15.48
	2006	51	51.45	14.64
	2007	51	54.00	12.64

N=275

Table 4. 14
Descriptive statistics for math scores across SES

Grade	year	*n*	*M*	*SD*
Normal	2004	197	56.49	19.56
	2005	205	61.24	16.55
	2006	205	65.35	16.80
	2007	206	64.82	17.35
Reduced	2004	17	39.35	23.29
	2005	15	45.73	23.68
	2006	17	49.88	20.49
	2007	17	51.59	18.12
Free	2004	45	49.42	20.86
	2005	49	51.18	16.09
	2006	50	56.26	18.04
	2007	51	55.63	17.66

N=275

Two-way ANOVA with socio-economic status (Normal, Reduced, and Free) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the reading test. A familywise alpha of .05 was used for all analyses. The interaction between factors was not significant $F(5.7, 706.5)=.589$, $MSE=71.94$, Huynh-Feldt $p=.731$. The main effect of the within-factor (year) was significant $F(2.86, 706.5)=21.81$, Huynh-Feldt $p<.001$. The main effect of the between-subjects factor (SES), was also significant $F(2, 256)=13.66$, $MSE=959.73$, $p<.001$.

Two-way ANOVA with socio-economic status (Normal, Reduced, and Free) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the math test. A familywise alpha of .05 was used for all analyses. The interaction between factors was not significant $F(5.3, 663.9)=.685$, $MSE=83.4$, Huynh-Feldt $p=.645$. The main effect of the within-factor (year) was significant $F(2.7, 663.9)=21.3$, Huynh-Feldt $p<.001$. The main effect of the between-subjects factor (SES), was also significant $F(2, 249)=7.89$, $MSE=1032.6$, $p<.001$.

Pairwise comparisons indicated students with Normal SES performed significantly better on the reading/language arts tests and math tests than students with Reduced and Free status. See tables 4.15 and 4.16 for results.

Table 4. 15
Pairwise comparisons of reading scores among SES levels:
Games-Howell Multiple Comparisons Results

SES level comparisons	Mean Difference	*S.E.*
Normal vs Reduced	16.44	4.26
Normal vs Free	9.77	2.14
Reduced vs Free	-6.66	4.49

*significant at .05 familywise alpha

Table 4. 16
Pairwise comparisons of math scores among SES levels:
Games-Howell Multiple Comparisons Results

SES level comparisons	Mean Difference	*S.E.*
Normal vs Reduced	12.61(*)	4.78
Normal vs Free	8.19(*)	2.67
Reduced vs Free	-4.41	5.23

*significant at .05 familywise alpha

Hypothesis Three:

Null: There is no statistically significant difference in achievement scores of male and female students instructed with the use of formative assessment data.

Table 4.17 and 4.18 detail the descriptive statistics for reading/language arts and math scores of students within gender categories. An ANOVA was used to analyze data for students, grouped according to their gender to determine if any variation in TCAP scores occurred as a result of instruction based on the use of formative assessment data.

Table 4. 17
Descriptive statistics for reading scores across Gender

Grade	year	*n*	*M*	*SD*
Male	2004	116	50.78	18.98
	2005	124	54.89	18.11
	2006	125	57.08	18.18
	2007	125	57.16	15.20
Female	2004	145	51.83	19.94
	2005	146	58.46	17.97
	2006	149	59.81	17.66
	2007	150	62.06	16.02

N=275

Table 4. 18
Descriptive statistics for math scores across Gender

Grade	year	*n*	*M*	*SD*
Male	2004	114	53.96	20.01
	2005	123	58.80	18.50
	2006	124	62.81	19.02
	2007	124	62.19	18.66
Female	2004	145	54.27	20.98
	2005	146	58.34	16.81
	2006	148	62.63	16.91
	2007	150	62.37	17.43

N=275

Two-way ANOVA with gender (male, female) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the reading test. The interaction between factors was not significant $F(2.86, 724.4)=2.208$, $MSE=71.3$ Huynh-Feldt $p=.089$. The main effect of the within-factor (year) was significant $F(2.86, 724.4)=49.03$, Huynh-Feldt $p<.001$. The main effect of the between-subjects factor (gender), was not significant $F(1, 253)=1.95$, $MSE=1051.48$, $p=.164$.

Two-way ANOVA with gender (male, female) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict

performance on the math test. A familywise alpha of .05 was used for all analyses. The interaction between factors was not significant $F(2.66, 664.5)=.163$, $MSE=83.68$, Huynh-Feldt $p=.902$. The main effect of the within-factor (year) was significant $F(2.65, 664.5)=54.09$, Huynh-Feldt $p<.001$. The main effect of the between-subjects factor (gender), was not significant $F(1, 250)=.003$, $MSE=1093.59$, $p=.958$.

Analyses of the data revealed the main effect of the between –subjects factor (gender) was not significant ($p=.958$) in the content area of reading/language arts and a $p=.164$ in the content area of math. Both p-values are greater than the alpha of .05; therefore, the null hypothesis is retained. We can conclude gender was not an influential factor in predicting reading/language arts or math scores across years and no effect beyond the effect of the within-subjects factor (year) was discovered.

Hypothesis Four:

Null: There is no statistically significant difference in achievement scores based on ethnicity due to the use of formative assessment data to drive curriculum and instructional decisions.

Table 4. 19

Descriptive statistics for reading scores across Ethnicity

Grade	year	*n*	*M*	*SD*
White	2004	247	51.81	18.95
	2005	256	57.13	17.85
	2006	259	59.08	17.66
	2007	260	60.18	15.79
Black	2004	5	32.80	31.25
	2005	6	48.50	25.16
	2006	6	44.00	24.97
	2007	6	52.50	16.96
Hispanic	2004	7	44.71	25.60
	2005	6	51.17	21.33
	2006	7	51.86	17.88
	2007	7	51.14	14.16
Asian	2004	2	66.00	7.07
	2005	2	59.00	26.87
	2006	2	59.50	21.92
	2007	2	66.50	19.09

N=275

Two-way ANOVA with ethnicity (white, black, Hispanic, and Asian) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the reading test. The interaction between factors was not significant $F(8.61, 720.39)=1.161$, $MSE=71.34$, Huynh-Feldt $p=.319$. The main effect of the within-factor (year) was not significant $F(2.87, 720.39)=3.194$, Huynh-Feldt $p=.025$. The main effect of the between-subjects factor (ethnicity), was not significant $F(3, 251)=1.615$, $MSE=1047.81$, $p=.186$.

Table 4. 20
Descriptive statistics for math scores across Ethnicity

Grade	year	*n*	*M*	*SD*
White	2004	246	54.91	19.90
	2005	255	59.02	17.47
	2006	257	63.23	17.74
	2007	259	62.62	17.98
Black	2004	5	18.80	16.69
	2005	5	46.80	18.54
	2006	6	52.83	17.97
	2007	6	49.83	18.03
Hispanic	2004	6	47.67	26.47
	2005	7	46.43	16.40
	2006	7	52.14	22.04
	2007	7	57.14	14.21
Asian	2004	2	66.00	7.07
	2005	2	70.00	15.56
	2006	2	63.00	0.00
	2007	2	74.50	20.51

N=275

Two-way ANOVA with ethnicity (white, black, Hispanic, and Asian) as a between-subjects factor and year (2004, 2005, 2006, and 2007) as within-subjects factor was used to predict performance on the math test. The interaction between factors was significant $F(8.08, 668.01)=2.042$, $MSE=81.29$, Huynh-Feldt $p=.039$. The main effect of the within-factor (year) was also significant $F(2.69, 668.1)=7.675$, Huynh-Feldt $p<.001$. The main effect of the between-subjects factor (ethnicity), however, was not significant $F(3, 248)=2.216$, $MSE=1073.64$, $p=.087$.

Data analyses indicate the main effect of the within-factor (year) as significant, the between-subjects factor (ethnicity) was not significant in the content area of reading/language arts ($p=.186$) nor the content area of math ($p=.087$). Both p-values are greater than the alpha of .05 and therefore the null hypothesis is retained. Due to extreme scarcity of non-white students and small cell sizes in factorial design, no further analyses

of this hypothesis were possible. The significance of the interaction effect might have been produced solely due to the limited number of observations and must be interpreted with caution.

Summary

The analysis of data in this chapter was used to determine if the use of formative assessment data by educators to make curriculum and instructional decisions had any effect on students' TCAP achievement scores. Data was displayed as it pertained to each of the four null hypotheses.

CHAPTER V

Discussion, Conclusions, and Recommendations

This study analyzed and evaluated four years of reading/language arts and math TCAP scores of 275 students in relationship to the following variable: instruction based on the use of formative assessment data or instruction based on the non-use of formative assessment data. Students' gender, ethnicity, and socioeconomic condition and how they related to the type of instruction were also analyzed. This chapter is a discussion of the field study and research findings. The chapter also includes conclusions and recommendations for further use.

Discussion

The participants in this study were from a southern, rural school. The participants attended the school during the 2003-2004, 2004-2005, 2005 – 2006, and 2006-2007 school terms.

ANOVA tests were used to analyze data for significant differences between group means. Analyses were conducted to test four null hypotheses at the .05 level of confidence each. The data were analyzed with the statistical software program SPSS.

Of the four null hypotheses, two were retained. Although it was determined that year was a significant predictor of success in both reading/language arts and math achievement, neither gender nor ethnicity, when combined with the use of formative

assessment data, could be determined to be a significant factor in predicting achievement on the reading/language arts or math portion of the TCAP test perhaps due to the small cell sizes in the factorial design and/or the absence of a diverse ethnic population within this study.

A statistically significant difference was established regarding the use of formative assessment as a predictor of achievement scores in reading/language arts and math. Specifically, the main effect of the within-factor (year) was found to be significant for both reading/language arts and math. The main effect of the between-subjects factor (grade) was found to be a significant predictor of 2006 and 2007 math scores but not for 2004 or 2005 math scores. In addition, seventh grade was found to have significantly higher math scores in years 2006 and 2007 than sixth or eighth grade. Year of measurement was a significant predictor of math scores for 7th grade and 8th grade. Unfortunately, sixth grade remained statistically similar in math achievement for all years analyzed. Math aptitude for this particular group has been viewed as low by former math teachers.

Formative assessment and SES as predictors of reading/language arts and math achievement revealed the main effect of the within-factor (year) and the main effect of the between subject factor (grade) to be significant. Results revealed year was a significant predictor of reading/language arts and math scores. After the introduction of formative assessment, the scores significantly improved in the content areas of reading/language arts and math on the TCAP achievement test. It was determined scores were significantly lower prior to the use of formative assessment data to make curriculum and instructional decisions. Scores in reading/language arts and math increased every

year except year 2006-2007, where the scores appear to plateau. In addition, those students classified as free or reduced scored significantly lower than students classified as normal. Additional intervention strategies and support systems must be in place for students whose socioeconomic status is classified as free or reduced.

Conclusions

The purpose of this field study was to determine the effect of using formative assessment data to drive curriculum and instructional decisions on the reading/language arts and math TCAP achievement scores of students of in grades six through eight. The following conclusions can be made based on the results of the study:

1. The use of formative assessment data to make curriculum and instructional decisions in the content areas of reading/language arts and math was found to have a statistically significant effect on the TCAP scores of students.
2. The socioeconomic condition of students was found to interact with the within-factor (year), indicating students, with the socioeconomic status of free or reduced, performed higher overall in reading/language arts and math achievement tests but performed significantly lower than those students classified as normal.
3. Gender was not an influential factor in how formative assessment affected reading/language arts or math achievement scores. Formative assessment influenced males and females equally.
4. The significance of the factor ethnicity influencing the effect of formative assessment count not have been determined due to small cell sizes in

factorial design and the extreme scarcity of non-white students in this study.

Recommendations

The following recommendations are proposed based on the findings of this field study:

1. A replication study with a different sample may be beneficial to determine if the use of formative assessment data used to guide educators in making decisions regarding curriculum and instruction has any effect on the TCAP scores of high school students.
2. It may be beneficial to conduct this study with a more ethnically diverse group of students in an urban school district.
3. It may be beneficial to conduct this study with a population of special education students to determine the effect of formative assessment data to instruct these students.

References

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education*, 85(1), 536-553.

Bettesworth, L. (2006) Administrators' use of data to guide decision-making. Ph.D. dissertation, University of Oregon, United States -- Oregon. Retrieved July 3, 2007, from ProQuest Digital Dissertations database. (Publication No. AAT 3224072).

Black, P., & Wiliam, D. (1998). Inside the black box - Raising standards through classroom assessment. *Phi Delta Kappan, 80 (2)*, 139-144.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan, 86(1)*, 8-21.

Boston, Carol (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation, 8(9)*, 1-8.

Campos, J. & O'Hern, J. (2007). How does using formative assessment empower students in their learning? Unpublished master's thesis, Saint Xavier University of Chicago.

Chappuis, J. (2005). Helping students understand assessment. *Association for Supervision and Curriculum Development, 63(3)*, 39-43.

Cromey, A. (2000). Using student assessment data: What can we learn from schools? Oak Brook, IL: North Central Regional Educational Laboratory.

Datnow, A., Park, V., & Wohlstetter, P. (2007). *Achieving with data: How high-performing school systems use data to improve instruction for elementary students*. San Francisco, CA: Institute for the Study of Knowledge Management in Education and NewSchools Venture Fund.

Dembosky, J., Pane, J., Barney, H., Christina, R. (2005). *Data driven decisionmaking in southwestern Pennsylvania school districts*, WR-326-HE/GF, Santa Monica, California: RAND Corporation.

Doyle, D. (2003). Data-driven decision-making: Is it the mantra of the month or does it have staying power. *T.H.E. Journal, 30 (10)*, 19-21.

DuFour, R. (2002). Beyond instructional leadership: The learning-centered principal. *Association for Supervision and Curriculum Development, 59 (8)*, 12-15.

DuFour, R. (2004). Professional learning communities focus on learning for all students. *Educational Leadership*, 61(8), 6-11.

Fuchs, L., Fuchs, D., Karns, K., Hamlett, C., Katzaroff, M., Dutka, S. (1997). Effect of task-focused goals on low-achieving students with and without learning disabilities. *American Educational Research Journal, 34 (3)*, 513-543.

Greer, L. (2001). Does changing the method of assessment of a module improve the performance of a student? *Assessment & Evaluation, 26 (2)*, 127-138.

Heritage, M. (2007). Formative Assessment: What Do Teachers Need to Know and Do? *Phi Delta Kappan, 89(2)*, 140-145.

Husman, J., Brem, S., & Duggan, M., (2005). Student goal orientation and formative assessment. *Academic Exchange Quarterly, 9(3)*, 355-340.

Johnson, J. (2000). Data-driven school improvement [Electronic version]. *Journal of School Improvement*, *1*, Retrieved October 03, 2007, from http://www.ncacasi.org/jsi/2000v1i1/data_driven

Kaftan, J., Buck, G., & Haack, A. (2006, March). Using formative assessments to individualize instruction and promote learning. *Middle School Journal*, 44-49.

Lezotte, L. (2001). *Revolutionary and evolutionary: The effective schools movement*. Okemos, MI: Effective Schools Products, Ltd

Marsh, J., Pane, J. & Hamilton, L. (2006), *Making sense of data-driven decision making in education*, OP-170-EDU, Santa Monica, California: RAND Corporation.

Marzano, R. (2005). *What works in schools* (PowerPoint presentation). Available: http://www.marzanoandassociates.com/documents/latest02.ppt#878,1,What Works in Schools: Translating Research into Action

McDonald, D. (2002, October). No Child Left Behind Act mandates assessment measures. *Momentum*, 33(3), 8-10.

Mizell, H. (2003, January). *Guiding Questions for Middle Grades Reform*. Speech presented at a meeting convened by the Los Angeles Unified School District, Los Angeles, CA.

Neill, M. (2006, April). Preparing teachers to beat the agonies of NCLB. *Education Digest*, 71(8), 8-12. Retrieved July 6, 2007, from Academic Search Premier database.

Protheroe, N. (2001). *Improving teaching and learning with data-based decisions: Asking the right questions and acting on the answers*. Arlington, VA: Educational Research Service.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119-144.

Schmoker, M. (2003). First things first: Demystifying data analysis. *Educational Leadership, 60(5)*, 22-24.

Shepard, L., Hammerness, K., Darling-Hammond, L., & Rust, F., Snowden, J., Gordon, E., Gutierrez, C., Pacheco, A. (2005). *Preparing teachers for a changing world - what teachers should learn and be able to do*. San Francisco: Jossey-Bass.

Smith, G. (2007). How does student performance on formative assessments related to learning assessed by exams? *Journal of College Science Teaching*, *36*, 28-34.

Stecher, B., Hamilton, L. (2006). *Using test-score data in the classroom*, WR-375-EDU, Santa Monica, California: RAND Corporation.

Stiggins, R. (2002, June). Assessment Crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83(10)*, 758-765.

ThinkLink Learning. Retrieved November 24, 2006, from About ThinkLink Learning Website: http://thinklinklearning.com/about.php

Wayman, J. (2005). Involving teachers in data-driven decision making; Using computer data systems to support teacher inquiry and reflection. *Journal of Education for Students Placed at Risk, 10(3)*, 295-308.

Appendices

Appendix A

Robertson County School Board Approval

Robertson County Schools
Daniel P. Whitlow, Director
2121 Woodland Street
P.O. Box 130
Springfield, Tennessee 37172
Phone: (615) 384-5588 Fax: (615) 384-9749

Alice Thomas, MSA
Assistant Director

Danny Weeks, Ed.D
Assistant Director

July 27, 2007

Mrs. Kathy Carroll
C/O Coopertown Middle School
Springfield, TN 37172

Mrs. Carroll:

Please accept this letter as permission to conduct a research project for your field study for Austin Peay State University. You have permission to access TCAP achievement data for Coopertown Middle School students in grades six through eight for the purpose of examining the use of formative assessments to make data-driven curriculum decisions and the effect of those decisions on student achievement.

According to your request, your research will involve searching records, and neither directly contacting nor interviewing students. Please remember, any student-identified material should be considered confidential and may not be disclosed.

We look forward to receiving the results of your study.

If my office can be of assistance, please do not hesitate to contact me.

Respectfully,

Dr. Danny L. Weeks

Appendix B

Austin Peay State University

Institutional Review Board Approval

Austin Peay
State University

College of Graduate Studies

July 31, 2007

Kathy L. Carroll
134 Ridgeview Trace
Hendersonville, TN 37075

Dear Ms. Carroll:

Thank you for submitting your research project "A Longitudinal Study of the Use of Formative Assessment Data to Improve Achievement of Middle School Students" to the Austin Peay State University Institutional Review Board. Your project has been reviewed according to our expedited approval process. This type of study qualifies for expedited review under FDA and NIH (Office for Protection from Research Risks) regulations.

Congratulations! This is to confirm that I have approved your application through one calendar year. Please make sure you follow the research procedures you described in your application when you conduct your research. This study is subject to continuing review on or before July 31, 2008, unless closed before that date. Changes classified as modifications to your research project must be promptly reported and approved by the IRB.

Thank you for your cooperation with the IRB human subjects review process. Best wishes for a successful study.

Sincerely,

Charles A. Pinder

Charles A. Pinder, Ph.D.
Dean and Chair of IRB

www.apsu.edu

P.O. Box 4458 • Clarksville, TN 37044 • P: (931) 221-7414 • F: (931) 221-7641