Comparison of Selected Benchmark Testing Methodologies as Predictors of Virginia Standards of Learning Test Scores

Paul Andrew Cox

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Education
In
Educational Leadership and Policy Studies

James R. Craig, Chair
Ann T. Abel
Susan G. Magliaro
N. Wayne Tripp

November 2, 2011

Blacksburg, Virginia

Keywords: Benchmark Assessments, Virginia Standards of Learning, Curriculum Based Measures, Predictors, Student Achievement

Comparison of Selected Benchmark Testing Methodologies as Predictors of Virginia Standards of Learning Test Scores

Paul Andrew Cox

ABSTRACT

This study examined how Math-Curriculum Based Measurement (M-CBM) benchmark assessment scores predict fifth grade math Standards of Learning (SOL) assessment scores. Three school districts participated in the study by providing student data on math benchmark assessments and fifth grade math SOL assessment scores. Scores were organized and analyzed to determine the level of prediction between the two data sets. The results of the study indicated that M-CBM benchmark assessments were good predictors of fifth grade math SOL assessment scores.

A second purpose of the study was to measure the differences in the capacity of three different M-CBM benchmark assessments in predicting fifth grade math SOL assessment scores. The three school districts used M-CBM benchmark assessments that varied in the method of creation and the type (commercial, teacher created, released test item) to develop the assessments. The findings show the M-CBM benchmark assessments created by teachers were better predictors of student scores on the fifth grade math SOL assessment followed by released test items and the commercially developed assessments.

The third goal of the research study was to determine if there were differences in costs associated with three different M-CBM benchmark assessments. The reported costs for each of the three assessment types were analyzed but the gross cost per student did not accurately reflect the total costs involved in creating, operating, and/or maintaining the M-CBM benchmark assessment systems in any of the three districts.

A literature review identified contributions in the field in the areas of assessment, Curriculum Based Measurement (CBM), and predicting student performance. The research design was quantitative and the school district data collected was over the 2009-2010 school year. Recommendation for future research was to focus on the implementation methods and utilization of M-CBM benchmark assessments by teachers to modify instruction. Possible research on how the benchmark data can be utilized or converted into a practical classroom predictor of SOL assessment outcomes was recommended also.

DEDICATION

I dedicate this work to my wife, Valerie, who has always been my biggest fan. Your love, encouragement, and patience were the inspiration that allowed me to complete my degree. Without Valerie, this project and so many things in my life would never have been possible.

I dedicate this work to my sons, Neal and Trent. Of all the times you both waited patiently for me to type one more line, it is now time to play.

I dedicate this work to my daughter, Emma. You give me unlimited strength and drive when you say “Daddy you can do anything.”

ACKNOWLEDGEMENTS

Dr. Jim Craig, thank you for the patience, guidance, encouragement, and belief in me. Thank you.

Dr. Wayne Tripp, a special thanks for serving on this committee and providing excellent guidance throughout the entire program. Thank you.

Dr. Ann Abel, thank you for serving on this committee and so many encouraging words that kept me always moving forward.

Dr. Sue Magliaro, thank you for serving on this committee. A sign of a great educator is one who is there when students need them the most. You were there.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

"Certis rebus certa signa praecurrunt."
(Certain signs precede certain events.)
De Divinatione, Marcus Tullius Cicero, 44 BC

Improving students' mathematics achievement is a national concern and a central goal of the American education system (Clarke & Shinn, 2004; Jordan, Kaplan, Locuniak, & Ramineni, 2007; National Educational Goals Panel, 2002). In 2003, fewer than 33% of fourth graders demonstrated proficient skills on the National Assessment of Educational Progress (NAEP) mathematics test (Manzo & Gallery, 2003; VanDerHeyden & Burns, 2008). The pressure on students to perform well on standardized tests has never been higher and ensuring that students are prepared for high stakes tests is critical (Amrein, Berliner, & Rideau, 2010). Since the introduction of the Virginia Standards of Learning (SOL) assessments in 1998, there has been increasing emphasis on the importance of student proficiency on SOL assessments as an indicator of both student achievement and instructional effectiveness (Stronge, Ward, Tucker, & Hindman, 2008).

Accordingly, there has been a corresponding increase in the interest of formative assessment instruments to inform instruction that evaluate student progress and predict student performance in advance of formal mathematics SOL assessments (Crawford, Tindal, & Stieber, 2001). In contrast to summative evaluation that is retrospective (i.e., data are collected after completion of instruction), formative evaluation involves the collection of data during instruction as a basis for modifying that instruction (Deno & Espin, 1991; Thurber, Shinn, & Smolkowski, 2002). Such assessments offer several advantages for educators who need to monitor and measure student progress within the context of formal SOL assessments. For example, effective formative assessment instruments can provide students and parents with timely feedback about student progress toward SOL assessments and, for teachers, can either validate or indicate problems with instructional methodology (Deno, 1986; Stiggins & DuFour, 2009).

Background: Setting the Context for the Inquiry

On January 8, 2002, the No Child Left Behind Act (NCLB) was signed into law as a means to reauthorize and extend the Elementary and Secondary Education Act of 1965 (NCLB, 2001). Among the main elements of the NCLB was the requirement that states conduct annual standardized testing of third through eighth grade students in math and reading. In addition, an end of the course test in science was required in at least one year between third and twelfth grade (NCLB, 2001). Schools that do not meet the standardized testing goals may then be subject to financial and resource sanctions (NCLB, 2001). As states have implemented standardized testing, a need has arisen a need for predictive methods that allow teachers to assess student performance and adjust instruction before the final standardized test is taken. The measurement procedures must be technically sound, quick and easy to administer and interpret, and must yield useful information about student performance in basic skills (Deno, 1985; Limbke & Stickler, 2007; Shinn, 1989). Formative assessments allow a teacher to optimize instructional techniques and materials to maximize performance and ensure best case outcomes on the standardized test.

With the emphasis on scores, schools and districts have created predictive assessment tools for early identification of student progress (Yeo, 2009). In many cases, an investigation of the predictive performance and costs of district created assessment methodologies has not been accomplished and reported. The thrust of the research study herein reported was to examine the available evidence regarding the effectiveness of formative assessments in predicting mathematic SOL assessment scores.

No Child Left Behind Act of 2001

In April of 1965, President Lyndon B. Johnson signed Public Law 89-10, the Elementary and Secondary Education Act of 1965 (ESEA). ESEA was part of President Johnson's "War on Poverty" initiative and represented one of the most comprehensive government reforms of education in the history of the United States. In later years Brauer (1982) wrote, "When President Lyndon B. Johnson declared metaphorical war on poverty in 1964, he set in motion an important, complex, and controversial phase in the history of reform in the United States…" (p. 89). The goal of ESEA was to provide an equitable educational environment for underprivileged children, and ESEA "…has consistently remained the single largest fiscal source of federal support for educationally vulnerable schoolchildren" (Thomas & Brady, 2005, p. 51).

With reauthorizations every five years, it was not until 2002 and the signing of the NCLB reauthorization that standardized state testing was included as a requirement for federal funding.

The NCLB requires annual testing of public school students in grades 3-12 (NCLB, 2001). Testing results indicate school status toward adequate yearly progress (AYP) and, based on student performance, the level of sanctions/rewards imposed by the Federal Government. Taylor, O'Day, & Floch (2010) reported the following:

> The NCLB is designed to achieve an ambitious goal: All children will be proficient in reading and mathematics by the 2013-2014 school year. A key strategy for achieving such a high goal is accountability. NCLB holds schools and districts accountable for their students' mastery of state academic content standards, as measured by state tests. (p. xix).

To comply with the NCLB and to avoid federal financial and resource sanctions, the Commonwealth of Virginia used the already established state-wide standardized testing assessments known as the Standards of Learning (SOL) assessment program.

Virginia Standards of Learning

The SOLs were officially adopted as the standards of student proficiency in Virginia Public Schools in June 1995 and student testing began in 1998 (VDOE, 2001). At the core of the SOL testing was the establishment of fundamental measures of learning and achievement that are also expected in order to meet the NCLB mandate. Overseen by the Virginia Department of Education (VDOE), SOL testing addresses four areas identified as critical to educational achievement: English, mathematics, science, and history/social science. English and mathematics are assessed annually in Grades 3 through 8, while science and history/social science are tested in Grades 3, 5, and 8. The SOL program has identified the objectives and goals of the specific curricula for each of the aforementioned test subjects. The standards are rigorous and proficiency is measured through the annual SOL assessments (VDOE, 2011).

The actual SOL assessments provide a way to quantitatively determine whether curricular benchmarks have been met through the measurement of student performance. Such testing has significant implications for state school systems and their associated Virginia School Performance Report Cards where a "passing" rate (i.e., 70% of the students score at least 400 points out of a possible 600 points on the test) allows a school to retain its accreditation and, conversely, a failing rate means loss of accreditation status. Loss of accreditation can result in federal sanctions on a range of resources, such as financial assistance. As Freidman (2006) notes, such high stakes have brought significant pressure to bear on both teachers and students to maintain a passing rate on the SOL assessments. Research on the third grade math SOL assessments by Wilkins and Jones (2009) provided a cautionary bookend to high stakes pressure as they reported:

> Unfortunately, based on our findings, a passing score on the tests does not indicate that students have mastered the conceptual understanding and problem solving specified in the mathematics SOL. As a result, the test scores must be interpreted and used with caution. (p. 194)

In spite of research by Wilkins and Jones (2009), the SOL assessments continue to be utilized and school systems have turned to benchmark tests to help prepare their student for such high stakes tests. The practice of benchmark testing has been developed to provide teachers the ability to assess student progress and, in some school districts, predict final SOL assessment scores for each student. Utilizing benchmark testing information, a teacher can interactively adjust curriculum content and instructional practice to help ensure optimal student performance on the final SOL assessment.

Curriculum Based Measurement

Benchmark testing is founded on the concept of Curriculum-Based Measurement (CBM) or CBM benchmark assessments. CBM benchmark assessment is implemented as a series of short tests in specific areas of instruction to provide near real-time feedback on student progress and evaluate progress toward the acquisition of basic skills during instruction (Hintze, Keller, Lutz, Santoro, & Shapiro, 2006). Assessing student progress using CBM benchmark assessment depicts academic growth because it is a classroom-based method that provides repeated samples of student performance (Fuchs & Fuchs, 1999). Immediate feedback allows a teacher to adjust curricular structure and instructional methodology “on the fly” to achieve optimal outcomes.

Theoretically, CBM benchmark assessments can be conducted over very short time intervals, possibly intervals as short as every day or every week (Deno, 2003). However, in a classroom environment, practical application of SOL benchmark testing occurs every two to three months, typically resulting in a total of three benchmark tests per school year. CBM benchmark assessments used three to four times per year as part of a district norming process can be made more effective by (a) providing advanced notice of students who may fail to meet state standards of performance, thereby giving teachers needed time to implement necessary interventions, and (b) supplying evidence of within- and across-cohort student progress (Helwig, Anderson, & Tindal, 2002).

Contemporary CBM-based SOL benchmarking developed from the concept of General Outcomes Measurement in the late 1970s and early 1980s and was originally validated by Deno, Marston, and Mirkin (1982). The early work focused on the use of CBM for student monitoring but has broadened in application to current usage of the techniques in student screening and test benchmarking. Since 1980, CBM has been extensively validated by a number of investigators (Merino & Beckman, 2010). Christ, Silberglitt, Yeo, and Cormier (2010) notes as follows:

> CBM is used to index annual student growth across the primary grades. Procedures and measurement metrics are developed for mathematics, spelling, written expression, and reading. CBM oral reading (R-CBM) rate is the most researched and well established of those available procedures. (p. 448)

Tsuei (2008) provides a similar view supporting the effectiveness of CBM for assessment, stating,

> Longitudinal CBM studies have indicated that incorporating CBM feedbacks into instructional planning enables general educators to provide more effective instructional programs and thus promotes students' achievements. (p. 47)

Based on the extensive validation of CBM techniques and more than 30 years of research supporting the use of CBM to aid in screening and progress monitoring in academic areas (Lembke & Stecker, 2007), it is generally accepted that CBM benchmark testing can be a valid predictor of final mathematics SOL assessment scores.

Statement of the Problem

As public school teachers become subject to increasing accountability, the pressure to guarantee student progress and performance has never been higher. As a result, schools have been working to find accurate assessment tools that can provide progressive glimpse of student progress that can be used to predict future performance on SOL assessments and adjust instruction appropriately (Deno, 2003).

Over 20 years ago, Deno (1986) reported that formative assessment instruments can provide students and parents with timely feedback about student progress toward high stakes tests. He also reported that data from formative assessments could be used by teachers to either validate or indicate problems with instructional methodology. Since that time, several research studies have been conducted that show formative assessments have the potential to be used to predict student scores on state tests (Merino & Beckman, 2010). However, Stecker, Lembke, and Foegen (2008) show that developing reliable, valid tests for students is a challenging process for educators. The problem is the lack of detailed analyses of the types of formative assessment tools available with which educators can make informed decisions on the type and application of assessments to optimize predictive performance for a given instructional environment. Therefore, there is a gap in educational research regarding the types of formative assessments that reliably predict student performance on high stakes tests such as the SOL mathematics assessment.

Significance of the Study

Scholarly Significance

The scholarly significance of the research regarding CBM benchmark testing that was conducted is found in the potential advancement of knowledge in the field of standardized testing and benchmark assessment methodology based on empirical evidence. The research findings reported add to and build upon previous scholarly work such as researchers Savage and Carless (2004) who contended that it is important for schools to assess children's early mathematics abilities in order to predict later performance. In previous studies Christ, Silberglitt, Yeo, and Cormier (2010), Fuchs, Fuchs, and Hamlett (1989), Fuchs, Fuchs, and Stecker (1989), showed that teachers who use formative assessment data are likely to use more specific and measurable goals, rely more substantially on data to guide instruction, and modify instruction more frequently. The data from formative assessments allow school leaders, teachers, and parents to adjust instructional programs to best fit the needs of each student.

Practical Significance

Using assessments that accurately predict student outcomes on the Virginia SOL assessments allows teachers to modify instruction and provide interventions to improve student performance on state achievement tests. It has been demonstrated repeatedly that part of effective instruction is formative evaluation (Fuchs & Fuchs, 1986; Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, Hamlett, & Stecker, 1990; Thurber, Shinn, & Smolkowski, 2002). By formatively evaluating students' mathematics progress, teachers can assess the effectiveness of their instruction within weeks to determine if their instruction is producing the desired results (Deno, 1986, 2003). However, few studies have been published that examine the relationships between formative CBM benchmark assessment and state assessments in math (Hintze, Keller, Lutz, Santoro, & Shapiro, 2006).

Definition of Key Terms

1. Benchmark assessment. A benchmark assessment is a formative assessment given periodically and designed to help the teacher to shape or form instruction. Benchmark assessments prepare students for the summative assessments by showing strengths and weaknesses of skills that can be improved through instruction (Lang, Stanley, & Moore, 2008).
2. Correlation. The measure of the extent to which two variables are linearly related (Myers & Well, 2003).
3. Coefficient of determination or (r^2). A statistical index that is the square of a correlation and that explains how much of the variability of a factor can be caused or explained by its relationship to another factor (Easton & McColl, 2011).
4. Cross-validation. Cross-validations determine " . . . how well a regression equation obtained in one sample performs in another sample from the same population " (Pedhazur, 1997, p. 209).
5. Curriculum-based measure (CBM). An evidence-based system of screening and progress monitoring that teachers use on a frequent basis to screen all students in a school, grade, or class and to assess the effects of instruction on student performance (Lembke & Stecker, 2007).
6. Standards of Learning (SOL). The SOLs describe the Commonwealth of Virginia's expectations for student learning and achievement in grades K-12 in English, mathematics, science, history/social science, technology, the fine arts, foreign language, health and physical education, and driver education (VDOE, 2011).
7. Standards of Learning (SOL) assessment. The end of course assessment used to measure student achievement in English, mathematics, science and history/social science (VDOE, 2011).
8. Step-wise multiple regression. A 'best' regression model is sometimes developed in

stages. A list of several potential explanatory variables is available and is repeatedly searched for variables which should be included in the model. The best explanatory variable is used first, and then the second best, and so on (Pedhazur, 1997).

9. Student achievement. A student's performance on the Mathematics SOL assessment.

Summary

The chapter provides the background that sets the context for the research that was conducted including NCLB, SOL testing, and CBM formative assessment. The statement of the problem follows coupled with brief summaries of the scholarly and practical significance of the study. The chapter concludes with the definitions of key terms.

CHAPTER 2
REVIEW OF THE LITERATURE

Introduction

Chapter 2 contains a discussion of the published scholarly findings relative to mathematical proficiency, formative assessments, summative assessments, predicting student performance, math curriculum based measurement, teacher developed assessments, released test item assessments, and commercially developed assessments. The chapter closes with a statement of purpose, the identification of the research questions and a summary.

Mathematical Proficiency

Proficiency in mathematics is a fundamental part of successfully functioning in today's world. Accordingly, the ability to assess mathematical proficiency in an accurate and feasible manner is a vital part of any school mathematics program (Helwig, Anderson, & Tindal, 2002). Given the importance of both effective mathematics instruction and accurate learning assessments, research that bridges assessment and the predictability of student performance in mathematics has the potential to provide significant advances in instructional techniques and technologies that can increase student achievement (Helwig, Anderson, & Tindal, 2002).

Although schools have always been under pressure to produce positive outcomes for all students, the passage of NCLB increased expectations that schools improve student performance and monitor student growth over time (Lembke & Stecker, 2007). When NCLB was signed into law in 2002, new goals and mandates were placed on schools to ensure students learn (NCLB, 2001). The expectation for schools to produce mathematically proficient students continues to be a national priority (Clark & Shinn, 2004). The NCLB mandate to monitor student growth has provided a wealth of student performance data (VanDerHeyden & Burns, 2005). A study conducted by Clarke & Shinn (2004) indicates that the current performance of United States students may be such that students will not have the necessary skills to meet the changing demands of the workplace. Proficiency in mathematics is a requirement for any technical work environment. In fact, many of the fields projected to have the highest rate of growth in available jobs will be open only to individuals who are proficient in mathematics (Clarke & Shinn, 2004).

A driving force behind recent educational reforms has been the demand for skilled workers who can apply problem-solving skills (Fuchs, Fuchs, & Courey, 2005). Employees are expected to be proficient in mathematics. According to the *Occupational Outlook Handbook* (Clarke & Shinn, 2004; IEL, 2003; USDOL 1997, 2007), individuals with mathematics fluency have greater earning potential. The Bureau of Labor Statistics reported that on average, 28-year-old workers who tested in the top quartile of mathematics skills on the NAEP earn 37% more than those in the lower quartiles (Clarke & Shinn, 2004; USDOL 1997).

Poor achievement in mathematics is a national concern (Jordan, Kaplan, Locuniak, & Ramineni, 2007) and one of the primary goals of instruction for schools is the development of students who are proficient in mathematics (Clarke & Shinn, 2004). Known as “the nation’s report card,” the federally sponsored NAEP is given to some 340,000 students every two years. Results from the 2003 NAEP indicate state and national efforts to raise student achievement are continuing to yield far greater progress in mathematics than in reading (Manzo & Galley, 2003). In mathematics, 31% of fourth grade pupils scored at or above proficient in 2007, compared with 22% in 2000 (Manzo & Galley, 2003). However, according to the results of 2009 NAEP, scores for fourth-grade students have remained the same since 2007 (NCES, 2011). The need for math proficiency as a life skill coupled with a limited level of achievement evokes a need to examine ways to increase student mathematical performance. An illustration of the trend in fourth and eighth grade mathematics scores is shown in Figure 1.

Figure 1. Trend in fourth and eighth grade NAEP mathematics average scores (USDOE, 2011).

Note. Figure 1 is from "National Assessment of Educational Progress," by U.S. Department of Education, Institute of Education Science, National Center for Education Statistics. Retrieved on April 28, 2011 from http://nces.ed.gov/nationsreportcard/pubs/main2009/2010451.asp

Now, perhaps more than ever before, a national premium has been placed on the importance of early identification and intervention of academic skill weaknesses (Hintze & Silberglitt, 2005). The need for early identification and intervention has driven the development of a range of evaluation methodologies to assess student performance which can be broadly classified into summative and formative assessments.

Summative Assessments

Probably the most widely recognized method of evaluating student progress is the practice of administering a test at the end of an instructional period to determine how much of the curriculum was learned by a student (Marshall, 2005). The use of an end of the year assessment is known as summative assessment. Summative assessments are implemented at the school, district, or state level and are primarily formed based on state standards and/or district curricula. Some examples of summative assessment include: end of unit tests, chapter tests, final exams, statewide assessments, and national standardized tests. Most states summative assessments appear in the form of high-stakes tests that determine whether the student has mastered certain skills (Lang, Stanley, & Moore, 2008).

Since summative assessments measure the amount of growth that has occurred after instruction is complete, the information gained from summative assessments is not usually beneficial to the teacher for changing instructional practices to improve learning with current students (Marshall, 2005; Shores & Chester, 2008). However, summative assessments can provide data to guide future instruction. The NCLB incorporates summative assessment by requiring all public schools to administer an annual state-wide standardized test (Merino & Beckman, 2010). While many had hoped that once-a-year tests would provide useful information for curriculum development, educators and others know close monitoring is not occurring (Perie, Marion, & Gong, 2009).

Given that summative tests are administered at the end of the year and scores are sometimes not received until the summer, summative tests are not useful in providing timely instructional feedback during the school year. Unlike end of the year summative assessments, formative assessments can provide more timely feedback to educators and allow instructional practice to be modified through the school year. The need for frequent feedback to educators has encouraged the development of formative assessment instruments.

Formative Assessments

Unlike "end of instruction" summative assessment tests, formative assessments are based on a series of near real-time data collections during the instructional process (Thurber, Shinn, & Smolkowski, 2002). Formative assessments use short time-frame, multiple assessments to track each student’s performance and provide feedback as instruction progresses, can give teachers a series of assessment snapshots that indicate the extent of students’ conceptual understanding of grade-level content, and can serve to fill the assessment void (Helwig, Anderson, & Tindal, 2002). In fact, it has been demonstrated repeatedly (Fuchs & Fuchs, 1986; Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, Hamlett, & Stecker, 1990; Thurber, Shinn, & Smolkowski, 2002) that part of effective intervention is formative assessment. Since children with math difficulties (compared to reading difficulties) are likely to be underserved in early elementary school (Jordan, Kaplan, Locuniak, & Ramineni, 2007), good predictive assessments can provide timely data to teachers and parents. The earlier a learning difficulty is identified, the sooner an intervention plan can be implemented to address the difficulty (McGlinchey & Hixson, 2004).

Teachers who use continuous, formative assessment are seldom surprised by an individual student's performance on an annual achievement test (Marshall, 2005). With increasing attention to accountability and high-stakes assessments, the relationship between CBM benchmark assessments and state-mandated testing programs has been a topic of interest (Hintze & Silberglitt, 2005).

Curriculum Based Measurement (CBM)

CBM formative assessments were first developed in the mid 1980s under the leadership of Stanley Deno at the University of Minnesota. CBM is a standardized procedure conducted repeatedly over time and provides information relevant to making decisions regarding student achievement (Deno, Marston, & Mirkin, 1982; Fuchs & Deno, 1991; Fuchs & Fuchs, 1999; Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Fuchs, Fuchs, Hosp, & Jenkins, 2001; Shinn, Shinn, Hamilton, & Clarke, 2002). CBMs provide teachers with reliable, valid, and efficient procedures for obtaining ongoing performance data with which to evaluate instructional programs (Fuchs & Fuchs, 1999) that represent a type of formative assessment that research has repeatedly shown is an effective component of instructional intervention (Fuchs & Fuchs, 1986; Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, Hamlett, & Stecker, 1990; Thurber, Shinn, & Smolkowski, 2002).

CBM uses the core educational curriculum of a subject area as the basis for test development and is designed primarily as a measurement and evaluation system that school psychologists and teachers can routinely use to monitor individual student progress and instructional effectiveness (Hintze & Silberglitt, 2005). Teachers can use the CBM data to assess the effects of instruction on student performance. In the end, CBMs produce performance indicators that assess current learning in addition to the retention and generalization of previously mastered material (Hintze & Silberglitt, 2005). Stecker, Fuchs, and Fuchs (2005) have published a review of the relevant CBM research literature that provides evidence that the use of CBMs can contribute to increased student achievement.

CBM Development

In an effort to decrease the separation between measurement and instruction–to make data on student achievement more integral to daily teacher decision making–a program of research was undertaken by Deno (1985). The intent of the research was to advance assessment practices to allow teachers to make changes to instruction during instruction. Hence, a new method of assessing students and instructional data emerged; namely, CBM benchmark assessment. During the six year Minnesota study, new procedures were developed that could be implemented by teachers to evaluate student progress and effective instruction. The results showed teachers were more effective using CBM benchmark assessments (Fuchs, Deno, & Mirkin, 1984) and provided a blueprint from which future CBM benchmark assessment techniques could be developed. Deno (1985) found the following four “design characteristics” (p. 221) were necessary for the optimal development of CBM benchmark assessments:

1. *Reliable* and *valid* if the results of their use were to be accepted as evidence regarding student achievement and the basis for making instructional decisions.
2. *Simple and efficient* if teachers were going to use them, or teach others to use them to frequently monitor student achievement.
3. *Easily understood* so that the results could be clearly and correctly communicated to parents, teachers, and students.
4. *Inexpensive* since multiple forms were to be required for repeated measurement” (p. 221).

In a later study, Herman and Baker (2005) added six criteria necessary to improve the validity of CBM benchmark assessments. Those criteria were: “. . . alignment to content standards; enhanced diagnostic value of assessment results through initial item and test structure design; fairness for all students; data showing technical quality; built in utility; and feasibility” (p.49). When CBM benchmark assessments are continuously woven throughout teaching and learning, the results can provide educators with valuable information about instruction and student progress (Stecker, Fuchs, & Fuchs, 2005).

CBM Benefits

Good assessments are meaningful to parents, useful as measures of progress, and sensitive to the effects of instruction (McGlinchey & Hixson, 2004). One of the best methods of formative assessments in academic areas and a method that exemplifies the characteristics of good measure is CBM benchmark assessment (Deno, 1985). Researchers have concluded that CBM benchmark assessment produces performance indicators that assess current learning in addition to the retention and generalization of previously mastered material (Hintze & Silberglitt, 2005). CBM benchmark assessment has been extensively researched and found to be a reliable and valid indicator of student skill level (McGlinchey & Hixson, 2004). Therefore, CBM benchmark assessment offers a promising alternative approach to traditional end of the unit assessments by continuously measuring student achievement of proficiency in basic skills (Deno, 1985, 2003).

However, it should be noted that some research studies have reported negative consequences associated with the use of CBM benchmark assessments. In a 2006 study, researchers found use of CBMs led to increased grade retention, which has repeatedly been shown to be counterproductive in terms of student learning (Neill, 2006). Neill also found a higher rate of teachers teaching to the test when using CBM benchmark assessment during the course of the academic year. Furthermore, educators have been very concerned about the loss of instructional time due to more time that students are in testing sessions; time spent testing could be time spent learning. CBM benchmark assessments add to an already crowded list of mandated tests (McElroy, 2006).

Predicting Student Performance

Predicting how students will perform on statewide tests of achievement is important for a variety of reasons (Helwig, Anderson, & Tindal, 2002). The demand for students skilled in mathematics coupled with low levels of achievement suggests a need to examine ways to increase mathematics achievement (Clarke & Shinn, 2004). Therefore, educators need reliable and valid assessment indicators to measure student progress during the instructional session so the instruction can be appropriately modified in a timely manner to positively affect student learning (Stecker, Lembke, & Foegen, 2008).

Most educators agree that it is important for schools to assess children's early mathematics abilities in order to predict later performance (Savage & Carless, 2004). However, with reliable CBM benchmark assessment tools for tracking student progress throughout a school year, classroom interventions can be used promptly when data from the CBM benchmark assessments show a lack of growth (Fuchs & Fuchs, 1999). Further justification for predicting achievement scores can be found in the school accountability movement that has put a premium on educators' providing evidence of student learning (Erickson, Ysseldyke, Thurlow, & Elliot, 1998; Helwig, Anderson, & Tindal, 2002).

It is important to note that predicting student outcomes on final summative assessments like the Virginia SOL assessments will require a formative assessment approach (Ainsworth, 2007). One such formative assessment approach is CBM based on benchmark testing. CBMs offer reliable data that are valid with respect to widely used indicators of achievement (Deno, 2003, 1991, 1985; Stecker, Lembke, & Foegen, 2008). In fact, Deno verified the use of CBMs in 1985 when he reported CBMs as valid and reliable methods in which teachers can assess student learning throughout the school.

Evidence from previous research indicates that it is possible for CBM benchmark assessments to be strong predictors of scores on large-scale mathematics achievement tests. In 2002, using CBM benchmark assessments to predict success on state standardized assessments, Helwig, Anderson, and Tindal were able to predict with 87% accuracy whether students would meet state mathematics standards (Helwig, Anderson, & Tindal, 2002). Crawford, Tindal, Stieber (2001) found a positive correlation between student CBM benchmark assessments in oral reading fluency and success on the Oregon state assessment (as cited in Merino & Beckman, 2010). In another study, the data from a CBM benchmark assessment were used to predict with a high degree of accuracy (greater than 80%) those students who were likely to pass the Minnesota Comprehensive Assessment (Hintze & Silberglitt, 2005).

Math Curriculum Based Measurement

CBMs were developed to function as “academic thermometers” to monitor students’ growth in basic academic skills domains. In practice, they are most frequently applied to reading, spelling, written expression, and mathematics (Hall & Mengel, 2002). Research into their use in mathematics curricula, known as Math Curriculum-Based Measurement (M-CBM), has been limited as compared to a wide range of studies on the use of CBM in reading, spelling, and written expression that have evaluated improvements in individual learning goals.

In a classroom where M-CBM benchmark assessment techniques are applied, the data from M-CBM benchmark assessments can be used to identify student progress (Hamilton, Halverson, Jackson, Mandinach, Supovitz, & Wayman, 2009). After each of the periodic M-CBM benchmark assessments have been administered, the information is then used to structure "re-teaching" sessions to students. Most of periodic M-CBM benchmark assessments or probes consist of mathematics problems based upon a District or State curriculum (Lembke & Stecker, 2007). The use of M-CBM benchmark assessments provide teachers with an enhancement in monitoring student progress and identifying specific areas of weakness (Hamilton, Halverson, Jackson, Mandinach, Supovitz, & Wayman, 2009).

One long term research and review effort into the use of M-CBM benchmark assessments to improve student achievement was conducted by Stecker, Fuchs, and Fuchs (2005) from 1980 through 2005. As part of their work, Stecker, Fuchs, and Fuchs examined several studies that had previously researched the relationship between M-CBM benchmark assessment application and corresponding student performance. The studies compared student achievement levels in classes where teachers used M-CBM benchmark assessments and made instructional modifications to classes where M-CBM benchmark assessment was not used. Stecker, Fuchs, and Fuchs found that four of five studies provided evidence that the use of M-CBM benchmark assessment techniques improved mathematics academic performance in elementary aged students. In support of their results, data from other researchers suggest that M-CBM benchmark assessments are accurate and reliable (Marston, 1989; Shinn, 1989). Marston (1989) reports correlations in the $r = .90$ range that indicate a strong relationship between M-CBM benchmark assessments and student performance on high stakes achievement tests.

Several studies have investigated the construct validity of M-CBM benchmark assessments. Thurber, Shinn, and Smolkowski (2002) found correlations ranging from $r = .36$ to $r = .63$ between different M-CBM benchmark assessments and high stakes state achievement tests. In the same study, Thurber and colleagues (2002) found a median correlation of $r = .82$ between computations and measures of basic math facts from the student textbook. And, finally, in a 1998 study, Good and Jefferson (1998) found evidence of concurrent validity with an $r = .60$ between two different M-CBM benchmark assessments taken at the same time.

An illustration of how M-CBM benchmark assessments are used for an individual student is shown in Figure 2. A two-minute M-CBM benchmark assessment was administered multiple times over a three month period to monitor student progress and classroom instruction. The M-CBM benchmark provided the data from which instructional decisions were made. When student performance (dotted line) dropped below the goal line (solid line), interventions were made to stimulate student growth and regain a course on the goal line.

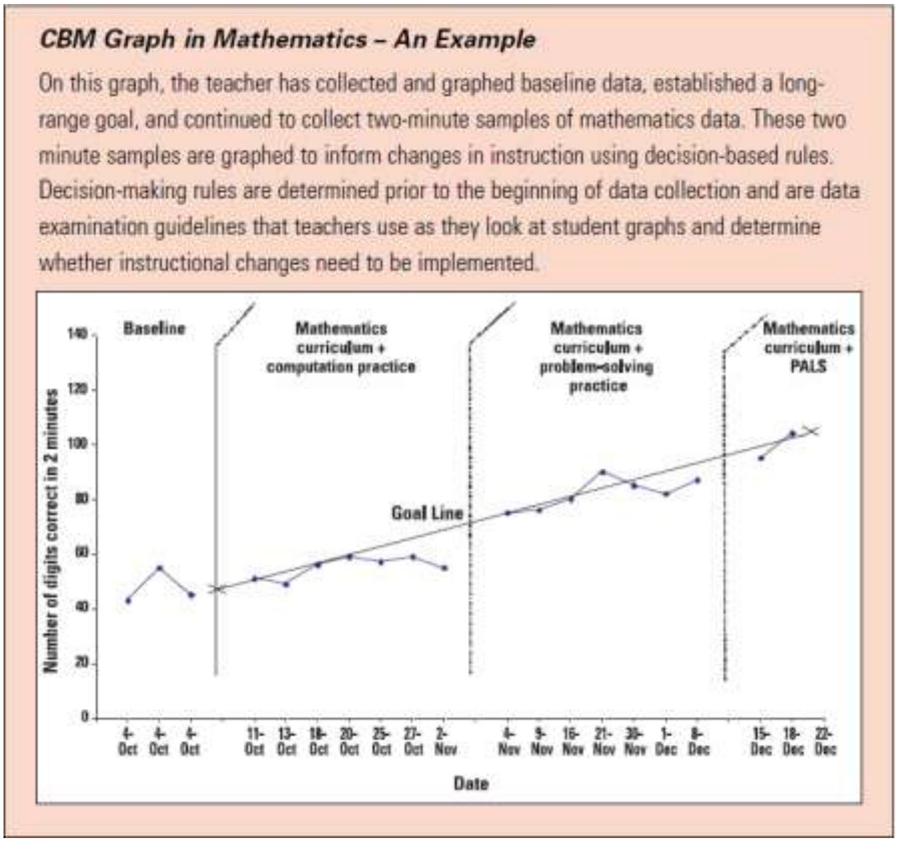

Figure 2. Example CBM graph in mathematics reprinted with permission from author. Lembke, E., & Stecker, P. (2007). *Curriculum-based measurement in mathematics: An evidence-based formative assessment procedure.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

The most important aspect of M-CBM benchmark assessment is how a teacher uses the data (Lembke & Stecker, 2007). Since students are screened frequently to assure they are on track to meet achievement goals, the use of data is of utmost importance. It has been shown that teachers who use such progress monitoring techniques are likely to use more specific and measurable goals, relying more substantially on data to guide instruction, and modify instruction more frequently (Christ, Silberglitt, Yeo, & Cormier, 2010; Fuchs, Fuchs, & Hamlett, 1989; Fuchs, Fuchs, & Stecker, 1989). In 2003, Fuchs, Fuchs, Hosp, & Hamlett, reported that data from CBMs could be classified into three types: *current status*, *growth*, or *dual discrepancy*.

The *current status* provides a level of performance compared to other students, sometimes referred to as norm score because it is compared to a peer group. The current status score is considered a starting point because no historical information is available (Fuchs, Fuchs, Hosp, & Hamlett, 2003).

The second use of CBM benchmark assessment data is to assess student *growth*. Testing students over time provides points that can be graphed and compared to a goal line. To track interventions, the growth data are used to indicate how much the student has learned. The growth data are very important because they identify which students have the potential to obtain the goal and which students require further intervention.

The third use of CBM benchmark assessment data is for a *dual discrepancy* analysis. The dual discrepancy analysis takes into consideration the current performance level and the amount of growth. In 2003, Fuchs suggested the method as a realistic approach to providing meaningful and useful measurement of student performance (Fuchs, Fuchs, Hosp, & Hamlett, 2003). An example of the dual discrepancy analysis would be comparing a child's growth over time to that of a same-age group (Fuchs, 1995). A student's height is measured and recorded for a predetermined length of time. Data from the one student is then compared to similar group of students. A child who shows a large discrepancy between his or her height and that of a normative comparison group may be considered a candidate for certain types of medical intervention.

In education, if a child is showing a discrepancy between the current level of academic performance and that of same-age peers, then that child may be a candidate for special education services. It should be noted, however, that a low performing child who shows growth rates similar to that of peers in the same classroom would not be a candidate for special education because the child is deriving similar education benefits from that classroom (Fuchs, 1995). Hence, the data gathered shows how much a student has learned over a defined time period and if the student has shown proof of understanding. It should be noted that one problem with M-CBM benchmark assessments is the difficulty in using M-CBM benchmarks in the identification of the particular skill or weaknesses. The use of M-CBM benchmark assessments does not provide a diagnosis for why students are falling behind. The data only show *when* students fall behind, not *why* students are falling behind (Clarke & Shinn, 2004).

Teacher Developed M-CBM Benchmark Assessments

Teacher developed M-CBM benchmark assessments are written or oral assessments that are not commercially produced or standardized (Burke, 1999). Tests are based on questions generated by individual teachers for specific classes. Teacher made tests can consist of a variety of formats, including matching items, fill-in-the-blank items, true-false questions, or essays. The basis for considering such assessments as legitimate evaluation tools arises from the fact that teachers routinely develop custom tests for specific classes as part of their everyday teaching routine. In fact, most classroom assessment involves tests that teachers have constructed themselves (Burke, 1999; Frey & Schmitt, 2007). It is estimated that 54 teacher-made tests are used in a typical classroom per year (Marso & Pigge, 1988). Furthermore, teachers place more weight on their own tests in determining grades and student progress than they do on assessments designed by others or on other data sources (Boothroyd, McMorris, & Pruzek, 1992; Frey & Schmitt, 2007).

Most teachers feel confident their assessments were valid and reliable (Frey & Schmitt, 2007; Oescher & Kirby, 1990; Wise, Lukin, & Roos, 1991). When asked to rate how sure they were that they could make achievement judgments based on their observations, 93% of teachers surveyed said that they were sure or very sure that they could accurately identify a student's level of performance (Deno 1985; Frey & Schmitt, 2007). However, one of the most serious problems with teacher made evaluation is the fact that a primary means of assessment (i.e., the test itself) is often severely flawed or misused (Hills, 1991) and there is little or no evidence of reliability/validity.

Teachers are not particularly good judges of their own ability or knowledge in test construction (Frey & Schmitt, 2007; Oescher & Kirby, 1990) and teachers' own estimates of ability and actual performance (in test construction) have been found to be negatively correlated (Frey & Schmitt, 2007; Marso & Pigge, 1988). Unfortunately, the reliability and validity of teacher created assessments is at best unknown and at worst inadequate (Deno, 1985; Frey & Schmitt, 2007). Stiggins, founder and executive director of the ETS Assessment Training Institute, has conducted extensive research on teacher made tests (Stiggins, Arter, Chappuis, & Chappuis, 2004) and reports that to build a balanced and effective assessment system, four essential conditions must be satisfied: clear learning targets, commitment to standards-based instruction, high quality assessments, and effective communications (Stiggins & Dufour, 2009). The quality conditions named previously must be met to create and produce high quality teacher developed tests (Stiggins, Arter, Chappuis, & Chappuis, 2004).

M-CBM Benchmark Assessments Based on Released Test Items

Since the passage of the NCLB Act, all 50 states have curriculum standards and a means by which to measure student acquisition of the required content (Braden, 2002; Shapiro, Keller, Lutz, Santoro, & Hintze, 2006). In some states, officials have chosen to release to the public sections of the standardized assessment tests from previous years. For example, each spring, the VDOE releases a sample set of the SOL assessments that were administered to Virginia public school students during the previous spring test administration (VDOE, 2011). The test questions are representative of the content and skills included in the Virginia SOL assessments and are provided to assist students and teachers in understanding the format of the tests and the type of questions that they can expect to see on future tests (VDOE, 2010). Teachers who decide to develop their own M-CBM benchmark assessments face several difficult tasks when constructing test items that reflect the learning to be measured, meet the criteria of a well-constructed test item, and provide important feedback to inform instruction (Tabernik, 2010). However, constructing a high quality test can be a complicated process. By using released test items, teachers can greatly reduce the work required to create and implement a series of assessments (Tabernik, 2010).

Tabernik (2010) made recommendations on how the released tests items from Florida public schools should be used by teachers, noting that the released test items should be utilized to examine examples of the state test, review the test length and the difficulty of the questions, consider what students know and do not know, and provide experience taking the test. Tabernik also noted that released tests should not be used to drill students on specific questions for the next year since the questions will be different or identify all the content that might be on future tests. Misusing released tests is explicitly addressed by the North Carolina Department of Public Instruction in the following statement:

> The State strongly cautions local school systems against using the forms (released items) to "teach to the test." While the end-of-grade and end-of-course tests summarize student achievement and serve as an important indicator of aggregate student performance for state and federal reporting, they are not the sole measure of student performance. The use of released tests should be part of an overall plan in school systems for implementing the new state testing program that includes formative, benchmark, and summative assessments. Because the released passages and questions will never again be used on the

state assessments, the use of these questions for test preparation would be misleading. (NCPS, 2011, p. 6)

In sum, current research on the use of released test items has been limited to only broad suggestions rather than specific guidelines regarding how to use release tests and researchers qualify the use of such tests with a cautionary tone about misuse of the resources that such tests provide (Tabernik, 2010; VDOE, 2011).

Commercially Developed M-CBM Benchmark Assessments

The third type of M-CBM benchmark assessment is commercially developed assessments. Because testing companies invest millions of dollars in test item development and field testing for educators to use (Tabernik, 2010), it would appear that schools could purchase reliable and valid M-CBMs that provide a comprehensive, in-district management system that creates, delivers, scores, and reports standards-based district tests using a variety of formats such as paper/pencil, bubble sheet, or computer testing (Pearson, 2010). One of the advantages of the commercially developed M-CBM benchmark assessments is that companies field-test test items on thousands of students to ensure that the items do not discriminate against students, have one correct answer, and reflect an appropriate range of difficulty (Toch, 2006).

While the use of commercially developed M-CBM benchmark assessments appears to be beneficial, such tests tend to be expensive investments and there are concerns about their instructional value (Fuchs & Fuchs, 1999). Eduventures, a Boston based research firm, estimates that states spent over $500 million in the 2005-2006 school year on testing in public schools (Jackson & Bassett, 2005). Although, commercial standardized achievement tests like the SOL assessments provide information about a student’s global standing, the summative score does not provide detailed information about a student’s deficiencies which make the tests inappropriate for early identification of at-risk children (Fuchs & Fuchs, 1999). Commercially distributed achievement tests are not always congruent with curriculum objectives and teachers tend not to value the information obtained from them (Cizek, 1993; Deno, 1985).

Purpose Statement

The purpose of the research study was to determine if M-CBM benchmark assessments predict fifth grade math SOL assessment scores.

Research Questions

The research questions addressed were:

1. Can M-CBM benchmark assessments (i.e., commercial, released test item, and teacher created) be used to predict fifth grade math SOL assessment scores?
2. Are there differences in the capacity of three different M-CBM benchmark assessments (i.e., commercial, released test item, and teacher created) in predicting fifth grade math SOL assessment scores?
3. Are there differences in costs associated with three different M-CBM benchmark assessments (i.e., commercial, released test item, and teacher created)?

Summary

A review of the literature relating to previous research on the development of CBM techniques and procedures and the use of such CBM's for predicting student achievement was presented. Researchers have been investigating for some time assessment techniques which would allow instructional practice to be modified "on the fly" as a curriculum progresses. The literature provides evidence that suggests that formative assessments are part of an effective curricular intervention that can optimize student performance. One type of formative assessment known as CBM benchmark assessment was reviewed and the results from published research suggest that CBM has the potential to be a strong predictor of student scores on summative tests. Thus, further study in the area of M-CBM benchmark assessments could provide valuable data that would advance knowledge within the field.

CHAPTER 3
METHODOLOGY

Introduction

As outlined in previous chapters, the general goal of the study was to determine if there was a predictive relationship between benchmark testing scores and the fifth grade math SOL assessment scores. The specifics of the research methodology employed are detailed below including: the participating school districts and their demographics, the study design, variables and instruments, data collection procedures, and data analysis procedures.

Participating School Districts and their Demographics

The study was conducted using data collected from three contiguous public school districts in a rural area of Southwest Virginia that employed different types of benchmark testing. Using a block of contiguous districts served to limit confounding variability that might have arisen from regional differences among districts. Table 1 compares the number of schools and enrollments in three school districts that participated in the study. As can be seen in Table 1,

Table 1

School District Data

	District A	District B	District C
District Student Enrollment	2355	4996	7486
Number of Elementary Schools	4	6	7
District Fifth Grade Enrollment	187	380	535
Number of Benchmark Tests Given	3	3	3
Frequency of Benchmark Tests	End of 9wks	End of 9wks	End of 9wks
End-of-Course Test	VA SOL	VA SOL	VA SOL
Free and Reduced Lunch Percentage	58%	56%	43%

District enrollment ranged from 2,355 to 7,486 students with the associated number of schools numbering 4 to 7. The fifth grade enrollments varied from 187 for District A to 535 for District

C. 2009-2010 academic year benchmark test scores and final SOL assessment scores were collected for 1,026 fifth grade mathematics students across the three districts. District A purchased commercial M-CBM benchmark assessments for student assessment, District B used teacher created M-CBM benchmark assessments, and District C used assessment benchmark tests created from released test items.

Study Design

The study was a predictive regression, cross-validation study conducted to examine the predictive relationship of M-CBM benchmark assessments to fifth grade math SOL assessment scores.

Instruments and Variables

Instruments

Data were acquired from four instruments: three types of benchmark assessments and the fifth grade math SOL assessment. District A purchased their benchmark tests from a nationally known educational company. District C created their benchmark tests using released test items provided by the Virginia Department of Education. District B used teacher created benchmark tests. All participants took the same fifth grade math SOL assessment.

Criterion Variable

The Virginia SOL assessments, high stakes summative assessments, are a series of tests that have been designed to measure student performance in the +areas of English, math, science, and history/social science. The math assessments are given to grades 3 through 8. Scores on the SOL assessment are scaled scores where a minimum score of 400 is considered passing. Individually, scores are not used toward graduation credit nor do they mandate summer school, remediation, or retention for the student. Although, a sum total of all student scores is used by the VDOE as part of the state accreditation process and by the Federal government to calculate progress toward the NCLB AYP Goals. The VDOE logs the SOL assessment results for all schools in the state of Virginia. Harcourt Testing service created and developed the tests. When students complete the SOL assessment, Harcourt scores the tests and sends the student scores to the VDOE for release to each individual school. Scores used were the fifth grade math SOL assessment administered during the spring of 2010 collected from the three school districts selected to participate in the study.

Student scores on the fifth grade math SOL assessment were used as the criterion variable for two reasons. First, it was assumed that the variability that may have arisen from student unfamiliarity with the SOL assessment process would be greatly reduced because students have already taken SOL assessments in the third and fourth grades. Second, since the fifth grade math SOL assessment has not been updated since 2006, it was assumed that the teachers and testing companies had had time to create M-CBM benchmark assessment's that fairly represented the fundamental elements of the SOL mathematics assessment.

According to the Virginia Standards of Learning Technical Report for the 2008-2009 Administration Cycle, construct validity was confirmed with a school level rank order correlation of 0.76 for the comparison of Virginia fifth grade math SOL assessment and the national percentile ranks of the Stanford 9 results. (The Stanford 9 norm referenced test compares student performance to the performance of a representative sample population of public school students of the same age and grade). The Virginia Standards of Learning Technical Report for the 2008-2009 Administration Cycle reported Cronbach's α to be .88 for the Virginia fifth grade math SOL (VDOE, 2011).

Predictor Variables

The predictor variables were the scores on three different school district formative M-CBM benchmark assessments given during the school year: a commercially prepared assessment was used by District A, a teacher created assessment was used by District B, and a test composed of items released from previous SOL assessments was used by District C. The characteristics of the three bench assessment systems are summarized in Table 2.

Table 2

Description of Selected Benchmark Assessments

	District A	District B	District C
Created By	Commercial	Teachers	Teachers
Type of Questions	Commercial	Teacher made	Used from released tests
Number of Questions	30	20	28
What is Measured	Mathematics	Mathematics	Mathematics
Target Group	Fifth grade	Fifth grade	Fifth grade
Scoring Method	Computer scored	Teacher Scored	Computer scored
Type	Computer delivered	Paper-pencil	Computer delivered
Response Type	Multiple choice	Multiple choice	Multiple choice

Students took the M-CBM benchmark assessments during the 2009-2010 school year during their fifth grade mathematics class. Assessments in School District A and C were scored, recorded, and the percentage correct was calculated by the computer. Teachers in School District B collected and recorded student scores after each assessment was administered. Teachers then recorded the scores as average scores in which the total of correct responses were divided by the number of questions on the M-CBM benchmark assessment. For example, if the M-CBM benchmark assessment for the first nine weeks had 25 questions and the student incorrectly answered five then the score was recorded as 80% (i.e., 20/25 = 80%). No data were collected concerning specific instructional modifications that resulted from the use of benchmark assessment scores.

Commercially prepared M-CBM benchmark assessments. District A purchased a yearly subscription through a nationally known educational publishing company. The web-based subscription allowed educators and administrators employed within District A to create assessments by choosing the number and type of questions from an online data bank. For example, the central office supervisor of elementary schools could create an M-CBM benchmark assessment with 40 questions, in multiple choice format, and select questions by units (e.g. for example adding single digit whole numbers or multiplication of two digit numbers). All questions were in multiple choice format and could be chosen across broad categories as well as specific concepts. Another option for test creation was for the company to create the test for the school district based on individual state educational standards. District A chose the second option and the M-CBM benchmark assessments were created and packaged by the company for the district. Each test consisted of multiple choice questions that were weighted equally. No evidence of the reliability or validity of the commercially prepared M-CBM was provided by District A.

The commercially prepared M-CBM benchmark assessments were given to fifth grade students as untimed assessments at the end of each nine weeks of the 2009-2010 school year for a total of three testing sessions for the year. The fifth grade math teacher took students to the computer lab where the assessments were administered online. At the conclusion of the student's testing session, the average score was stored at the company website for teacher access and analysis.

Teacher created M-CBM benchmark assessments. School District B utilized teacher created tests. The tests were created during a summer workshop in which six fifth grade math teachers attending the workshop collaborated to create three M-CBM benchmark assessments for the following school year. The three teachers created M-CBM benchmark assessments that were paper-and-pencil tests administered and scored by the fifth grade teachers at the end of each nine week period in each elementary school in District B. The percentage correct score was calculated by the teacher and recorded on the assessment. Scores remained with the student's fifth grade teacher. The reliability and validity of the teacher created M-CBM assessments are unknown.

Released item M-CBM benchmark assessments. School District C developed M-CBM assessments based on SOL items released from previous SOL assessments. Using an unspecified procedure, two fifth grade math teachers assembled three tests from released items for the 2009-2010 school year during a summer workshop. The released item M-CBM benchmark assessments were displayed and scored by the computer at the end of each 9 week period in each elementary school in District C. Individual student scores were accessible by the teacher, principal, and elementary supervisor. The released item M-CBM assessments' reliability and validity are unknown.

Data Collection Procedures

Before data were collected, approval to conduct the research was sought from the superintendent of each school district chosen to participate in the study. A sample authorization letter is included as Appendix A. The Institutional Review Board (IRB) approval letter was obtained and is included as Appendix B. Upon receipt of written approval, the relevant test score data was obtained from the Assistant Superintendent for each school district. The test score data included:

1. A copy of the three fifth grade mathematics benchmark tests students took during the 2009-2010 school year.
2. Individual student scores on each of the three fifth grade mathematics benchmark assessments (e.g., Student #1: first benchmark score-92, second benchmark score-85, and third benchmark score-93).
3. Associated individual student scores on the Virginia Department of Education fifth grade math SOL assessment.
4. The cost of developing, implementing, and operating the M-CBM assessment system in each district.

Data Analysis

To answer the first research question, a linear regression cross validation procedure was used to build prediction equations for each school district. Student scores for each school district were randomly split into two groups with one group designated the screening sample and the other group the calibration sample. The screening sample was used to construct a prediction equation using the M-CBM benchmark assessment scores as the predictor variables. Using the regression equation derived from the screening sample to predict fifth grade math SOL score, the calibration sample was used to determine the correlation between observed scores in the calibration sample and the predicted scores. The R^2 from the screening sample was compared to the square of the cross-validity coefficient derived from the calibration sample.

To answer the second research question, the screening sample regression equation developed from the previous research question was used to predict the SOL mathematics assessment scores for all students in each district. The regression equation from the screening sample provided a predicted SOL assessment score for each student. The predicted scores were then correlated with the actual fifth grade math SOL assessment score and a Coefficient of Determination was calculated for each district.

To answer the third research question, the costs associated with creating, operating, and maintaining the M-CBM benchmark assessments from the school districts were requested from the superintendent for each of the three school districts. The information from each school district provided cost data that were directly attributable to the development or purchase of benchmark assessment tests. No data were available relating to other less direct costs of the testing such as computer usage, teaching time and resource usage, supplies, etc. Based on the data provided by the school districts, the cost per student per school district that was directly associated with each district's M-CBM benchmark assessment system was analyzed without the inclusion of such other supplemental costs.

CHAPTER 4
FINDINGS

Introduction

The purpose of the research study was to determine if M-CBM benchmark assessments predict fifth grade math SOL assessment scores. Data were collected and analyzed from three adjacent school districts in Southwest Virginia for the 2009-2010 school year. Students in each district were given three M-CBM benchmark assessments over the course of the school year before taking the SOL mathematics assessment. The findings and analyses are presented in the context of each of the research questions outlined in Chapter 2. School district data is included in Appendix C.

M-CBM Descriptive Data

Each of the three school districts used unique M-CBM benchmark assessments. School District A purchased the M-CBM benchmark assessments from a nationally known testing company to monitor the progress of students in four elementary schools. Students completed the M-CBM benchmark assessments at three regular intervals throughout the school year: around the 45^{th}, the 90^{th}, and the 135^{th} days of school. The mean percentage correct and standard deviations of the three benchmark assessments for the elementary schools in School District A are presented in Table 3. In addition, the mean percentage correct for each benchmark assessment is depicted graphically in Figure 3.

Table 3

Descriptive Data of Assessments (Percent Correct) for School District A Using Commercial M-CBM Benchmark Assessments (BM_C)

District A		School			
		1	2	3	4
Mean (Standard Deviation)	BM_C 1	62.96 (29.00)	63.16 (27.74)	60.74 (26.07)	82.13 (25.00)
	BM_C 2	64.75 (24.00)	66.75 (24.94)	59.28 (22.29)	86.17 (20.18)
	BM_C 3	73.34 (20.02)	66.75 (18.80)	66.75 (21.85)	96.18 (4.69)

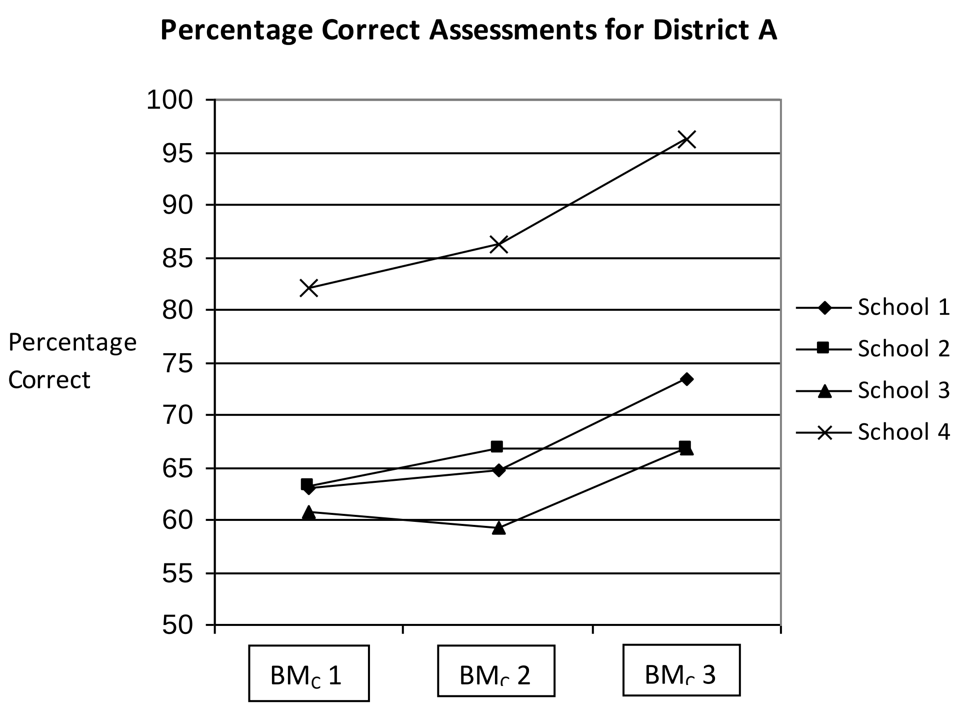

Figure 3. Descriptive data of assessments (percent correct) for school district A using commercial M-CBM benchmark assessments (BM_C).

School District B used teacher created M-CBM benchmark assessments to monitor the progress of students in six elementary schools. Students took M-CBM benchmark assessments at three times throughout the school year: near the 45th, the 90th, and the 135th days of school. The mean percentage correct and standard deviations of the three M-CBM benchmark assessments for the elementary schools in School District B are provided in Table 4. The mean percentage correct for each assessment is shown graphically in Figure 4.

Table 4

Descriptive Data of Assessments (Percentage Correct) for School District B Using Teacher Created M-CBM Benchmark Assessments (BM_T)

District B		School					
		1	2	3	4	5	6
Mean (Standard Deviation)	BM_T 1	68.10 (18.06)	71.00 (17.59)	78.22 (19.46)	81.90 (18.34)	85.10 (11.40)	78.59 (15.33)
	BM_T 2	79.87 (16.70)	82.78 (11.75)	84.71 (10.42)	83.65 (15.37)	84.88 (10.42)	80.35 (16.30)
	BM_T 3	70.83 (17.26)	84.40 (10.31)	85.40 (11.69)	86.53 (12.38)	82.45 (11.57)	80.37 (16.02)

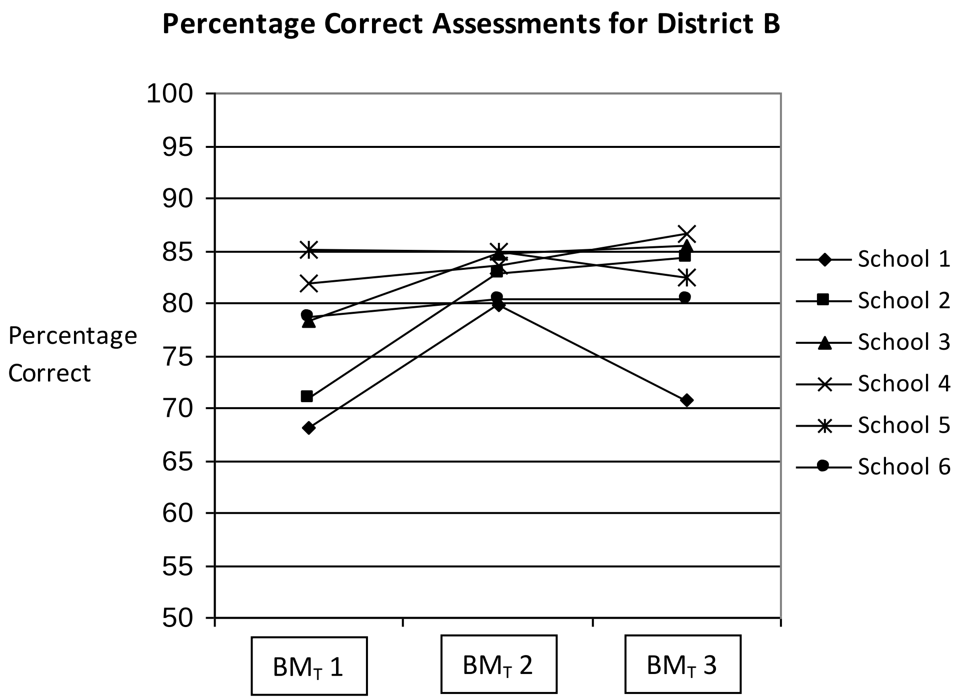

Figure 4. Descriptive data of assessments (percent correct) for school district B using teacher created M-CBM benchmark assessments (BM_T).

School District C used released test items from the VDOE to create M-CBM benchmark assessments to monitor the progress of students. There were seven elementary schools in School District C and students completed the benchmark assessments close to the 45th, the 90th, and the 135th days of school. The mean percentage correct and standard deviations of the three M-CBM benchmark assessments for the elementary schools in School District C are displayed in Table 5 while the percentage correct for each benchmark assessment is exhibited graphically in Figure 5.

Table 5

Descriptive Data of Assessments (Percent Correct) for School District C Using Released Item M-CBM Benchmark Assessments (BM_R)

District C		School						
		1	2	3	4	5	6	7
Mean (Standard Deviation)	BM_R 1	90.95 (9.82)	83.16 (12.44)	75.41 (16.20)	77.28 (16.83)	91.93 (7.22)	79.76 (11.50)	82.51 (14.07)
	BM_R 2	88.38 (9.66)	84.00 (11.38)	83.12 (12.67)	78.41 (17.71)	89.46 (9.00)	86.34 (7.62)	87.67 (11.37)
	BM_R 3	89.39 (11.92)	83.70 (12.82)	73.92 (16.79)	79.91 (15.45)	88.63 (9.87)	86.52 (8.69)	80.97 (14.46)

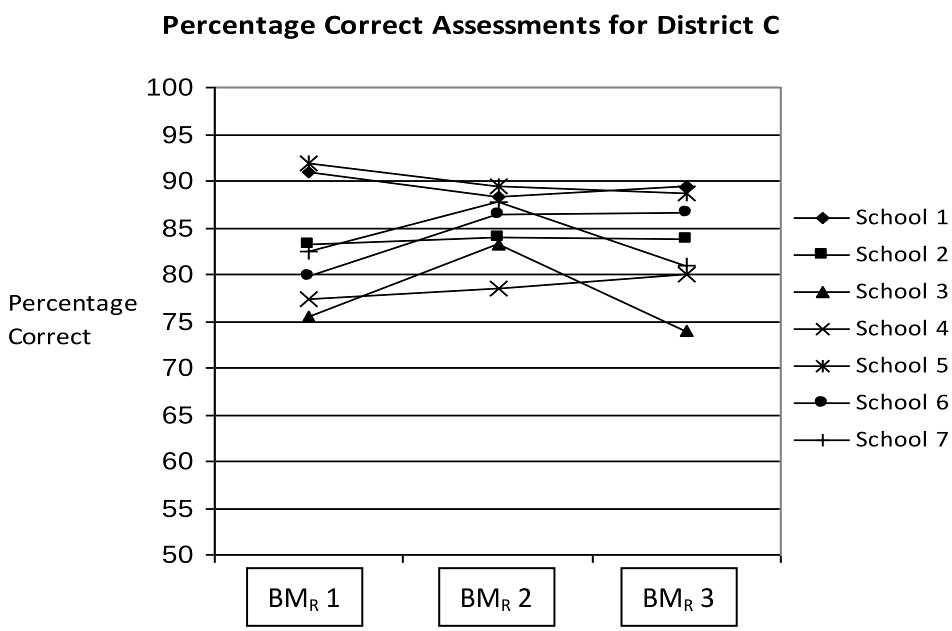

Figure 5. Descriptive data of assessments (percent correct) for school district C using released item M-CBM benchmark assessments (BM_R)

SOL Descriptive Data

The means and the standard deviations of the SOL mathematic assessment scores in mathematics for the students attending the elementary schools of the three participating school districts are presented in Table 6. While the mean SOL mathematics scores for each school district is shown graphically in Figure 6.

Table 6

Mean of SOL Mathematics Assessment Scores for School Districts A, B, and C

District	School						
District A	School						
	1	2	3	4			
Mean	506.11	497.17	464.20	521.94			
(Standard Deviation)	(88.14)	(88.63)	(77.26)	(59.72)			
District B	School						
	1	2	3	4	5	6	
Mean	490.17	492.87	472.00	466.86	488.20	478.46	
(Standard Deviation)	(104.56)	(103.20)	(94.03)	(88.54)	(74.40)	(84.06)	
District C	School						
	1	2	3	4	5	6	7
Mean	509.34	529.91	477.53	489.79	546.51	513.83	534.22
(Standard Deviation)	(72.45)	(57.69)	(73.02)	(83.86)	(53.69)	(59.03)	(60.47)

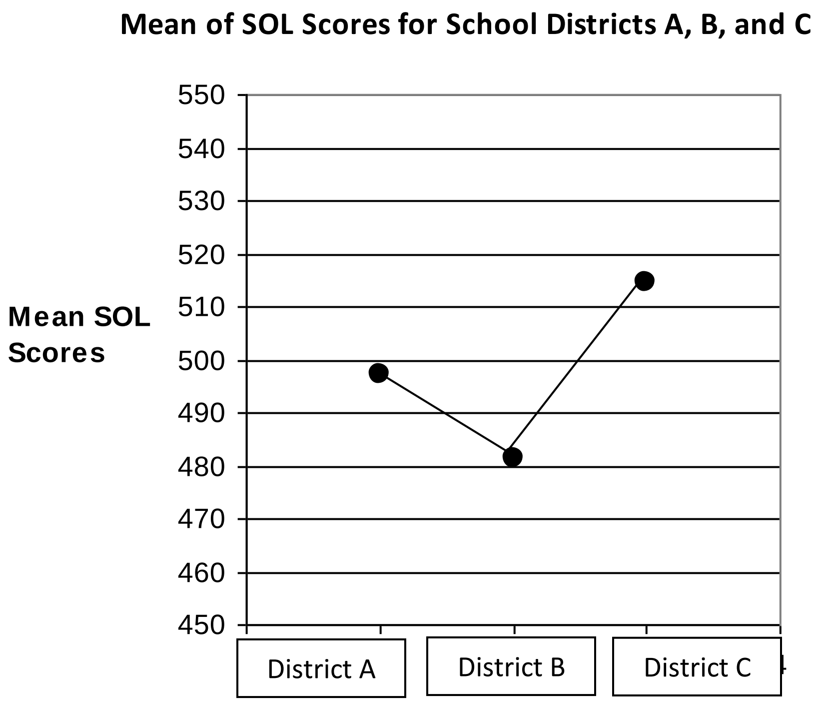

Figure 6. Mean SOL mathematics assessment scores for school districts A, B, and C

Data Analysis for Research Question 1

Research question 1: Do M-CBM benchmark assessments predict fifth grade SOL mathematical achievement?

To answer the first research question, a linear regression cross validation procedure (Pedhazur, 1997) was used to build prediction equations for each school district. Student M-CBM benchmark assessment scores and fifth grade math SOL assessment scores for each school district were randomly split into two groups. For each school district, one group of students was randomly designated to serve as the screening sample with the remaining group of students functioning as the calibration sample. The screening sample was used to construct a prediction equation. The student SOL mathematics assessment scores served as the criterion variable and each district's three M-CBM benchmark assessment scores were used as the predictor variables in three separate analyses. The second sample or calibration sample was used to determine the cross-validity coefficient which is the correlation between observed SOL mathematics assessment scores in the calibration sample and the predicted SOL mathematics assessment scores based on the regression equation derived from the screening sample. For each district, the R^2 from the screening sample was compared to the square of the cross-validity coefficient derived from the calibration sample. If the difference between the two is small, then the screening sample regression equation for a district can be used to predict the SOL mathematics assessment scores for the district.

Table 7 shows the difference between R^2 for the screening sample and the square of the cross-validity coefficient derived from the calibration sample for each of the school district. In each instance, the difference is comparatively small. Thus, the values showed that the screening sample regression equation for a district could be used to predict the SOL mathematics assessment scores for each district.

Table 7

Comparison of R^2 With Square of the Cross Validity Coefficient

	District A	District B	District C
Screening Sample	0.512	0.626	0.427
Calibration Sample	0.408	0.597	0.566

Data Analysis for Question 2

Research question 2: Are there differences in the capacity of three different M-CBM benchmark assessments in predicting fifth grade math SOL assessment scores?

To answer Research Question 2, the screening sample regression equation for each school district was used to predict the SOL mathematics assessment scores for all students in each district. By using the regression equation from the screening sample, a predicted SOL assessment score was produced for each student in each district. The predicted SOL assessment scores were then correlated with the actual SOL assessment scores for each district. In each instance the correlation was squared to produce the Coefficient of Determination for each district; the coefficients are presented in Table 8.

Table 8

Coefficient of Determination between M-CBM Assessments and SOL Mathematics Assessment based on the Screening Sample Analysis

School District	Coefficient of Determination (r^2)
District A	0.453
District B	0.612
District C	0.498

School District B's teacher created M-CBM benchmark assessments were found to produce the highest Coefficient of Determination for predicting SOL assessment scores in mathematics. The R^2s for the M-CBM benchmark assessments that were created from the VDOE released tests in School District C and the School District A purchased M-CBM benchmark assessments were about the same and lower than the R^2 for School District B's teacher created M-CBM benchmark assessments.

Data Analysis for Question 3

Research question 3: Are there differences in costs associated with the three different M-CBM benchmark assessments?

The cost information associated the implementation and operation of the M-CBM benchmark assessment systems was requested from each school district and the information provided proved to be non-definitive. For example, School District A purchased commercially produced M-CBM benchmark assessments in which the annual subscription fee for three assessments for all fifth grade students was reported to be $2.25 per student. The $2.25 fee included the creation of the three M-CBM benchmark assessments and establishing/maintaining the website that students log in and complete the assessments. The fee also included student accounts where assessments and scoring data are maintained. The fee included online tutorials for educators on how to administer the assessments where scoring for all three assessments is accomplished, where teacher accounts for student data are complied, and where reports that educators can print that detail student percent correct scores can be generated. The total cost of the assessment program for the 187 students tested for the 2009-2010 school year was $420.75 (187 x $2.25).

For School Districts B and C, the costs associated with creating, operating, and maintaining the M-CBM benchmark assessments in both districts were reported as the costs of teacher time devoted to the creation of the assessments. For School District B, six fifth grade math teachers attended a summer workshop to create the three M-CBM benchmark assessments and were each paid $150 to create the three assessments for a total cost of $900 (6 teachers x 150 daily rate of pay). No operational or maintenance costs were reported. Thus, the total cost for creating, operating, and maintaining the teacher created M-CBM benchmark assessment system in School District B was reportedly $900 or $2.37 per student for the 380 students tested in the 2009-2010 school year.

Likewise for School District C, two fifth grade math teachers developed three M-CBM benchmark assessments using released test items from the VDOE released fifth grade math SOL assessments. The two teachers were each paid a daily rate of $160 for one day to assemble the three M-CBM benchmark assessments based on released items for a total cost of $320 (2 teachers X 160 daily rate of pay). School District C also reported no operational or maintenance costs associated with conducting the three M-CBM benchmark assessments. Therefore, the reported cost for the 535 student tested for the 2009-2010 school year was $320 or a cost of $.60 per student.

Based on the information provided by the school districts, the gross cost per student per school district associated with each district’s M-CBM benchmark assessment system is reported in Table 9. However, it should be noted that the gross cost per students does not appear to accurately reflect the total costs involved in creating, operating, and/or maintaining the M-CBM benchmark assessments systems in any of the three districts and any assessment of the cost effectiveness of the three different M-CBM benchmark assessments systems is inappropriate based on the cost evidence acquired.

Table 9

Comparison of Costs Per Student From School District A, B, and C for 2009-2010

School District	Gross cost per student
District A	$2.25
District B	$2.37
District C	$.60

CHAPTER 5
SUMMARY AND LIMITATIONS

Introduction

The chapter begins with a presentation of the research questions addressed in the current study and a short summary of the findings. Next, the significance of the outcomes of the analyses is discussed. The limitations and delimitations are then presented along with suggestions for future research. The chapter closes with conclusions for each of the research questions.

Overview of the Current Study

The research questions addressed in the study were:

1. Do M-CBM benchmark assessments predict fifth grade SOL mathematic assessment scores?
2. Are there differences in the capacity of three different M-CBM benchmark assessments in predicting fifth grade math SOL assessment scores?
3. Are there differences in costs associated with three different M-CBM benchmark assessments?

Summary of the Findings

Using a cross-validation regression procedure provided evidence that three different types of 2009-2010 M-CBM benchmark assessment systems used three different school districts were each significant predictors of student performance on the fifth grade math SOL assessment. For each school district, a regression equation from the screening sample was used to determine the cross-validity coefficient for the calibration sample by applying the regression equation from the screening sample to the calibration sample. The difference between R^2 for the screening sample and the square of the cross-validity coefficient derived from the calibration sample for each of the school district was found to be small. Thus, it was concluded that the screening sample regression equation for a district could be used to predict the fifth grade SOL assessment scores for each district. The findings confirm that M-CBM benchmark assessments predict student scores on the fifth grade mathematics SOL assessment. School District B's teacher created M-CBM benchmark assessments had the highest R^2 and were found to be the most accurate predictors The R^2s were smaller and about equal for M-CBM benchmark assessments created from the VDOE released tests in School District C and the School District a purchased M-CBM benchmark assessments.

Finally, for Research Question 3, it was determined that calculating the costs associated with the three types of M-CBM benchmark assessments was problematic. For example, School District A purchased an annual subscription for each student enrolled in the 5th grade. This cost included the creation, development, delivery, scoring, and reporting of student scores on each of the three M-CBM benchmark assessments. However, School Districts B and C could not provide an inclusive cost estimate for the type of M-CBM benchmark assessments they used. In both instances, teachers were paid a daily rate of pay to create the assessments but the costs of operating and maintaining the assessment systems were not available. Therefore, it is not possible to directly compare costs of creating, operating, and maintaining the three types of M-CBM assessments.

Significance of the Study

In general, the findings were consistent with previous research comparing M-CBM benchmark assessments on skills taught during the entire school year with a statewide math competency tests and SAT scores (Fuchs et al., 2007; Helwig et al., 2002; Magyar et al., 2007; Shapiro et al., 2006). The scholarly significance of the study is found in the advancement of knowledge in the area of predicting fifth grade math SOL assessment scores from M-CBM benchmark assessments and the predictability capacity of different types of M-CBM benchmark assessments. Foegen, Jiban, and Deno (2007) reported that most of the previous studies on M-CBM benchmark assessments focused on limited mathematics skills such as quantity discrimination or identifying the missing number. And of the 41 recent studies on M-CBM benchmark assessments, only one used over 1,000 students in the sample (Foegen, Jiban, & Deno, 2007; Fuchs et al., 2007). The current research also utilized a sample size of over 1,000 students and confirmed the previous research on the capacity of different types of M-CBM benchmark assessments. For educators, the study expands the scholarly knowledge in the area of predicting student performance on the SOL assessment. For educational leaders, the study expands the scholarly knowledge in the area of predicting student outcomes and in the development of M-CBMs that have the most predictability of student scores.

An important secondary finding is that teacher created M-CBM benchmark assessments were found to be better predictors of student fifth grade math SOL assessment scores than assessments created from released test items or commercially developed assessments. Research that shows one type of assessment as a better predictor of students achievement is valuable information for educators and educational leaders who are utilizing teacher created classroom assessments to predict student scores (Burke, 1999; Frey & Schmitt, 2007). As previously mentioned, M-CBM benchmark assessments are formative assessments that are intended to inform the teacher of student progress. The findings suggest that M-CBM benchmark assessments can be used by educators as a precursor for student success on the state assessment. The findings also confirm previous research where M-CBM benchmark assessments were used to predict student performance on high-stakes tests (Hintze & Silberglitt, 2005) and where formative assessments were used by educators to make decisions on future instruction (Lembke & Stecker, 2007). Few studies have been published that examine the relationship between formative CBM benchmark assessment and state assessments in math (Hintze, Keller, Lutz, Santoro, & Shapiro, 2006). The study offers additional findings on the relationship between formative assessments and state assessments.

As noted in previous research, M-CBM benchmark assessments are flexible and can be used to monitor the progress of individual students (Lembke & Stecker, 2007). The findings suggest that all three of the different types of M-CBM benchmark assessments could be used to monitor progress of students toward the final SOL assessment. Another practical application for educators and educational leaders would be the development of new M-CBM benchmark assessments for predicting student success. In particular, the findings suggest that teacher created M-CBM benchmark assessments were found to be the best of the three types of assessments in predicting student outcomes but that all three were predictors of student scores. In future applications, teachers and educational leaders might be able to use a combination of two of the types of development techniques to create even more accurate assessments by creating a hybrid of teacher created and released item test formats.

Limitations

While the results indicated that M-CBM benchmark assessments are good predictors of student fifth grade math SOL assessment scores, there are limitations to the extent that the findings can be generalized. It is recognized that a statistically significant finding of the predictability of M-CBM benchmark assessments in a research study in three Southwestern Virginia school districts does not necessary mean the results are applicable to other school systems. For example, a school in Florida may not find similar results with M-CBM benchmark assessments created to predict the Florida state math assessment. This is probably the most apparent limitation of using a single type of testing tool as a common denominator across a wide range of school systems. Another limitation relates to the fact that the data analysis is simply a "snapshot" of a single school year (i.e., 2009-2010) of testing and that changes which occur in test content and structure each year may limit the generalization of the research results from year to year. Therefore, if the same study was repeated for a different school year, the findings may not be the same due to changes to M-CBM benchmark assessments and/or the fifth grade math SOL assessment.

And finally, even though the Virginia curriculum for fifth grade mathematics is standardized, there can be wide variation in instructional capabilities of the teachers and the way in which the curriculum is implemented. Every teacher has a particular way of teaching and while the SOLs provide a common curriculum across the Commonwealth of Virginia, instruction techniques vary between teachers. The instructional experience for every classroom is different. This variation in student educational experience and teaching environment was not directly addressed by the research data and is a limitation in the reported results. A research effort with a focus on practical classroom application would be able to provide more teacher specific data in this area and, as a result, determine the extent to which teacher actions contribute to the effectiveness of benchmark assessment performance.

Delimitations

Delimitations associated with the study were:

1. The investigation used fifth grade students in three public school districts in Southwestern Virginia.
2. The investigation was limited to fifth grade mathematics.
3. Two of the M-CBM benchmark assessments were used exclusively by the school districts that created them.

4. The study was limited to students who had taken all three M-CBM benchmark assessments during the year and the SOL assessment. Students who had not taken all three M-CBM benchmark assessments and/or the fifth grade math SOL assessment were excluded from analyses.

Suggestions for Future Research

The direction of future research has several possibilities. First, the findings need to be replicated using multiple school districts that use M-CBM benchmark assessments. As previously mentioned, there was one school district for each type of M-CBM benchmark assessment (commercial, teacher created, released test item) used in the study. Each of the three school districts was located in Southwestern Virginia. It is not known how applicable the findings are to other districts in other portions of the state or districts in different states. .

Second, it seems logical that the one of the most important research efforts should focus on identifying and understanding the utilization of M-CBM benchmark assessments by teachers and administrators. Not much is known about if/how teachers and administrators use feedback from M-CBM benchmark assessments to effectively modify instruction and if there are optimal ways to provide that feedback to teachers. So, timely feedback of assessment information would seem to be critical to its use in positively affecting the teaching/learning process, the mechanisms for the accomplishing that feedback are worthy of study (Deno 1985). In addition, the use of M-CBM benchmark assessments as a component of the teacher evaluation system probably should be studied. Teachers who do not currently receive a student growth score from the VDOE may need to create assessments that measures student growth as part of the teacher evaluation system and would benefit from knowing and understanding how M-CBM assessments are created and used.

Third, future research should closely examine the processes by which the various types of M-CBM benchmark assessments are created. For example, the teacher-made tests from School District B were reportedly made from teachers using an undefined procedure and unknown resources. What may be most germane to the effectiveness of teacher-made tests in predicting fifth grade math SOL assessment scores may not lie so much in the fact that teachers created the benchmark assessments but in the processes and resources the teachers used in creating the benchmark assessments. Further, it is not known if issues of reliability and validity were considered and, if considered, how they were addressed.

Lastly, the costs associated with M-CBM benchmark assessments should be thoroughly examined. In the current study it was not possible to accurately determine the actual costs of developing, operating, and maintaining teacher-made or released item M-CBM benchmark assessments. It appears that both the direct and indirect costs are difficult to identify and isolate and may require creating new and different ways to assess the real costs of implementing a benchmark assessment system.

Conclusion

More than 30 years of research supports the use of formative assessments like M-CBM benchmark assessments to aid in the screening and progress monitoring of students (Lembke & Stecker, 2007). M-CBM benchmark assessments can be used to inform instructional techniques with the ultimate goal of increasing student achievement. The findings suggest that M-CBM benchmark assessments, at least as implemented in the school districts studied, are valid predictors of student scores and therefore could provide educators another tool to inform and modify instruction to maximize student achievement.

REFERENCES

Ainsworth, L. (2007). Common formative assessments: The centerpiece of an integrated standards-based assessment system. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp.79-101). Bloomington: Solution Tree.

Amrein, A., Berliner, D., & Rideau, S. (2010). Cheating in the first, second, and third degree: Educators' responses to high-stakes testing. *Education Policy Analysis Archives*, *18*(14), 36.

Boothroyd, R. A., McMorris, R. F., & Pruzek, R. M. (1992). *What do teachers know about measurement and how did they find out?* Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA. (ERIC Document Reproduction Service No. 351309)

Braden, J. (2002). Best practices for school psychologists in educational accountability: High stakes testing and educational reform. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology*, *4*(1), 301-320.

Brauer, C. (1982). Kennedy, Johnson, and the war on poverty. *The Journal of American History*, *69*(1), 98-119.

Burke, K. (1999). *The mindful school: How to assess authentic learning*. (3rd ed.). USA: Skylight training and publishing.

Christ, T., Silberglitt, B., Yeo, S., & Cormier, D. (2010). Curriculum-based measurement of oral reading: An evaluation of growth rates and seasonal effects among students served in general and special education. *School Psychology Review*, *39*(3), 447-462.

Cizek, G. J. (1993). Rethinking psychometricians' beliefs about learning. *Educational Researcher*, *22*(4), 4-9.

Clarke, B., & Shinn, M. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, *33*(2), 234-248.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, *7*(4), 303-323.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children*, *52*, 219-232.

Deno, S. L. (1986). Formative evaluation of individual programs: A new role for school psychologist. *School Psychology Review*, *15*, 358-374.

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education*, *37*(3), 184-192.

Deno, S. L., & Espin, C. A. (1991). Evaluation strategies for preventing and remediating basic skill deficits. In G. Stoner, M. R. Shinn, & H. M. Walker (Eds.), *Interventions for achievement and behaviors problems*. Silver Spring, MD: National Association of School Psychologists, 79-97.

Deno, S. L., Marston, D., & Mirkin, P. (1982). Valid measurement procedures for continuous evaluation of written expression. *Exceptional Children*, *48*, 368-371.

Easton, V. J., & McColl, J. H. (2011). Statistical Glossary, v1.1, Retrieved September 3, 2011, from http://www.stats.gla.ac.uk/steps/glossary/basic_definitions.html

Erickson, R., Ysseldyke, J., Thurlow, M., & Elliott, J. (1998). *Reporting the results of students with disabilities in state and district assessments* (NCEO Policy Directions 8). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Foegen, A., Jiban, C., & Deno, S. L. (2007). Progress monitoring in mathematics: A review of the literature. *The Journal of Special Education, 41,* 121-139.

Frey, B. B., & Schmitt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics*, *18*(3), 402-423.

Friedman, A. (2006). State standards and digital primary sources: A divergence. *Contemporary Issues in Technology & Teacher Education*, *6*(3), 313-327.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children*, *57*, 488-500.

Fuchs, L., Deno, S., & Mirkin, P. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement and student awareness of learning. *American Education Research Journal*, *22*(2), 449-460.

Fuchs, L. S., & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis. *Exceptional Children*, *53*, 199-208.

Fuchs, L. S., & Fuchs, D. (1999). Monitoring student progress toward the development of reading competence: A review of three forms of classroom-based assessments. *School Psychology Review*, *28*(4), 659-671.

Fuchs, L., Fuchs, D., & Courey, S. (2005). Curriculum-based measurement of mathematics competence: From computation to concepts and applications to real life problem solving. *Assessment for Effective Intervention*, *30*(2), 33-46.

Fuchs, L. S., & Fuchs, D., & Hamlett, C. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education*, *10*, 43-52.

Fuchs, L. S., Fuchs, D., Hamlett, C., & Stecker, P. (1990). The role of skills analysis in curriculum-based measurement in math. *School Psychology Review*, *19*, 6-22.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 27-48.

Fuchs, L., Fuchs, D., Hosp, M., & Hamlett, C. (2003). The potential for diagnostic analysis with curriculum-based measurement. *Assessment for Effective Intervention*, *28*(3-4), 13-22.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, *5*(3), 241-258.

Fuchs, L. S., Fuchs, D., & Stecker, P. (1989). Effects of curriculum-based measurement on teachers' instructional planning. *Journal of Learning Disabilities*, *22*(1), 51-59.

Good, R. H., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (p. 61-88). New York: Guilford.

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making* (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hall, T., & Mengel, M. (2002). *Curriculum-based evaluations: Effective classroom practices report*. National Center on Accessing the General Curriculum. U.S. Department of Education, Office of Special Education Programs.

Helwig, R., Anderson, L., & Tindal, G. (2002). Using a concept-grounded, curriculum-based measure in mathematics to predict statewide test scores for middle school students with learning disabilities. *The Journal of Special Education*, *36*(2), 102-112.

Herman, J., & Baker, E. (2005). Making benchmark testing work. *Assessment to Promote Learning, 63*, 48-54.

Hills, J. R. (1991). *Measurement and evaluation in the classroom* (2nd ed.). Columbus, OH: Merrill.

Hintze, J., Keller, M., Lutz, J., Santoro, L., & Shapiro, E. (2006). Curriculum-based measures and performance on state assessment and standardized tests. *Journal of Psychoeducational Assessment, 21*, 19-35.

Hintze, J., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*, 372-386.

Institute for Educational Leadership (IEL): National collaborative on workforce and disability. (2003). Math, science, and technology: Essential skills for career success in the 21st century. *Info Brief 7*. Retrieved January 23, 2011, from http://www.ncwdyouth. info/resources_&_Publications/information_Briefs/issue7.html

Jackson, J. M., & Bassett, E. (2005). *The state of the K-12 state assessment market*. Boston: Eduventures.

Jordan, N., Kaplan, D., Locuniak, M., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*, *22*(1), 36-46.

Lang, S., Stanley, T., & Moore, B. (2008). *Short cycle assessment: Improving student achievement through formative assessment*. Larchmont, NY: Eye on Education, Inc.

Lembke, E., & Stecker, P. (2007). *Curriculum-based measurement in mathematics: An evidence-based formative assessment procedure.* Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Magyar, C. I., Pandolfi, V., & Peterson, C. R. (2007). In J. W. Jacobson, J. A. Mulick, & J. Rojahn (Eds.), *Handbook of intellectual and developmental disabilities* (p. 333-351). New York: Springer.

Manzo, K., & Galley, M. (2003). Math climbs, reading flat on ’03 NAEP. *Education Week*, *23*(1), 81.

Marshall, J. (2005). Formative assessment: Mapping the road to success. *The Princeton Review*. Retrieved January 19, 2011, from http://www.dcsclients.com/~tprk12/Research_Formative%20Assessment_White_Paper.pdf

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M.R.Shinn, *Curriculum-based measurement: assessing special children*. New York, NY: The Guilford Press

Marso, R. N., & Pigge, F. L. (1988). *An analysis of teacher-made tests: Testing practices, cognitive demands, and item construction errors*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA. (ERIC Document Reproduction Service No. 298 174)

McElroy, E. J. (2006). Testimony given at improving NCLB: Successes, concerns, and solutions, before the commission on No Child Left Behind, Jack Morton Auditorium, George Washington University, Washington, DC. Retrieved April 29, 2011, from http://www.aspeninstitute.org/site/c.huLWJeMRKpH/b.1543329/k.CEDC/Hearings.htm

McGlinchey, M., & Hixson, M. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review*, *33*, 193-203.

Merino, K., & Beckman, T. (2010). Using reading curriculum-based measurements as predictors for the measure academic progress map standardized test in Nebraska. *International Journal of Psychology: A Bio psychosocial Approach*, *6*, 85-98.

Mosteller, F., & Tukey, J. W. (1968). Data analysis, including statistics. *Handbook of social psychology*, Reading, MA: Addison-Wesley.

Myers, J. L., & Well, A. A. (2003). *Research design and statistical analysis*. Lawrence Erlbaum Associations, Inc. Retrieved August 2, 2011 from EBSCO*host*.

National Center for Educational Statistics. (2009). The nation's report card: Mathematics 2009. Retrieved April 28, 2011 from http://nces.ed.gov/nationsreportcard/pubs/main2009/2010451.asp

National Educational Goal Panel. (2002). *Raising achievement and reducing gaps: reporting progress toward goals for academic achievement in mathematics*. Retrieved January 9, 2011, from http://govinfo.library.unt.edu/negp/page5.htm

NCLB, No Child Left Behind Act of 2001: United States Public Law, Vol. 20, No. 6301, United States Congress, 107th Congress Session.

NCPS, North Carolina Public Schools. (2011). Released tests frequently asked questions. Retrieved January 30, 2011, from http://www.ncpublicschools.org/docs/accountability/testing/releasedforms/faqs81309.pdf

Neill, M. (2006). Preparing teachers to beat the agonies of NCLB. *Education Digest*, *71*(8).

Oescher, J., & Kirby, P. C. (1990). *Assessing teacher-made tests in secondary math and science classrooms.* Paper presented at the Annual Meeting of the National Council on Measurement in Education, Boston, MA, (ERIC Document Reproduction Service No. ED322 169)

Pearson Education, Inc. (2010). *Pearson benchmark successes*, Retrieved February 11, 2011, from http://www.pearson.com/about-us/education/

Pedhazur, E.J. (1997). *Multiple regression in behavioral research: explanation and prediction (3rd ed.)* Fort Worth, TX: Harcourt Brace.

Perie, M., Marion, S., & Gong, B. (2009). Moving toward a comprehensive assessment system: A framework for considering interim assessments. *Educational Measurement: Issues and Practice*, *28*(3), 5-13.

Savage, R., & Carless, S. (2004). Predicting curriculum and test performance at age 7 years from pupil background, baseline skills and phonological awareness at age 5. *British Journal of Educational Psychology*, *74*, 155-171.

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: reading and math performance in Pennsylvania. *Journal of Psycho educational Assessment*, *24*, 19-35.

Shinn, M. (1989). *Curriculum-based measurement: Assessing special children*. New York: Guilford.

Shinn, M. R., Shinn, M. M., Hamilton, C., & Clarke, B. (2002). Using curriculum-based measurement in general classrooms to promote reading success. In M. R. Shinn, H. M. Walker, & G. Stoner (Eds.*), Interventions for academic and behaviors problems II: preventive and remedial approaches* (p.113-142). Bethesda, MD: National Association of School Psychologists.

Shores, C., & Chester, K. (2008). *Using RTI for school improvement: raising every student's achievement scores*. Thousand Oaks, CA: Corwin Press.

SPSS, Inc. (2008). *SPSS 16.0 for Windows Graduate Pack.* Chicago: SPSS, Inc.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42,* 795-819.

Stecker, P. M., Lembke, E., & Foegen, A. (2008). Using progress-monitoring data to improve instructional decision making. *Preventing School Failure, 52*(2), 48-58.

Stiggins, R. J., Arter, A. A., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment for student learning: Doing it right-using it well*. Portland, OR: Assessment Training Institute.

Stiggins, R., & DuFour, R. (2009). Maximizing the power of formative assessments. *Phi Delta Kappan*, 5, 641-644.

Stronge, J., Ward, T., Tucker, P., & Hindman, J. (2008). What is the relationship between teacher quality and student achievement? An exploratory study. *Journal of Personnel Evaluation in Education, 20*(3-4), 165-184.

Tabernik, A. M. (2010). Mathematics released test items. Science & Mathematics Achievement Required For Tomorrow. Retrieved January 29, 2011 from http://www.smartconsortium.org/index.php?l1=19

Taylor, J., O'Day, J., & Floch, K. (2010). *State and local implementation of the* No Child Left Behind Act: Volume *IX – Accountability Under* NCLB: *Final Report.* http://www.ed.gov/about/offices/list/opepd/ppss/reports.html#au), xix.

Mott, D. (2011). *Validity and Reliability Statement*: Tests for Higher Standards. Retrieved February 11, 2011, from http://www.tfhs.net/validity.pdf

Thomas, J., & Brady, K. (2005). The Elementary and Secondary Education Act at 40: Equity, Accountability, and the Evolving Federal Role in Public Education. *American Educational Research Association, Review of Research in Education, 29,* 51-67.

Thurber, R., Shinn, M., & Smolkowski, K. (2002). What is measured in mathematics tests? Construct validity of curriculum-based mathematics measures. *School Psychology Review*, *31*(4), 498-513.

Toch, T. (2006). Turmoil in the testing industry. *Educational Leadership, 64*(3), 53-57.

Tsuei, M. (2008). A web-based curriculum-based measurement system for class-wide ongoing assessment. *Journal of Computer Assisted Learning*, *24*(1), 47-60.

U.S. Department of Education (USDOE). Institute of Education Science. National Center for Education Statistics, National Assessment of Educational Progress (NAEP). Retrieved on April 28, 2011 from http://nces.ed.gov/nationsreportcard/pubs/main2009/2010451.asp

U.S. Department of Labor (USDOL). Bureau of Labor Statistics. (1997). Occupational outlook handbook (OOH), 1996-97 Edition.

U.S. Department of Labor (USDOL). Bureau of Labor Statistics. (2007). *Occupational outlook handbook (OOH), 2006-07 Edition.* Retrieved January 23, 2011, from http://www.bls.gov/oco/

VanDerHeyden, A., & Burns, M. (2005). Using curriculum-based assessment and curriculum-based measurement to guide elementary mathematics instruction: effect on individual and group accountability scores. Assessment *for Effective Intervention*, *30*(3), 15-31.

VanDerHeyden, A., & Burns, M. (2008). Examination of the utility of various measures of mathematics proficiency. *Assessment for Effective Intervention*, *33*, 215-224.

Virginia Department of Education. (2001). *Every child can succeed: A parent's guide to Virginia's Standards of Learning program*. Richmond, VA: VDOE.

Virginia Department of Education. (2011). Testing & Standards of Learning (SOL). Retrieved January 16, 2011 from http://www.doe.virginia.gov/testing/index.shtml

Virginia Department of Education. (2011). Virginia Standards of Learning Assessments: Technical Report 2008-2009 Administration Cycle. Retrieved September 25, 2011 from http://www.doe.virginia.gov/testing/test_administration/technical_reports/sol_technical_report_2008-09_administration_cycle.pdf

Wilkins, J., & Jones, B. (2009). An Examination of a state-level mathematics test: What type of learning does the test actually measure? *Planning & Changing*, *40*(3/4), 194-206.

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education, 42*, 37-42.

Yeo, Seungsoo. (2009). Predicting performance on state achievement tests using curriculum-based measurement in reading: A Multilevel Meta-Analysis. *Remedial and Special Education*, 31(6), 412-422.

APPENDIX A

PERMISSION LETTER TO SUPERINTENDENTS

Superintendent of Schools
District A, B, and C
City/Town, Virginia 24210

Dear Superintendent,

I hope your school year is going well. My name is Andy Cox and I would like to request permission to conduct a research study to determine if math benchmark assessments are predictors of SOL assessment scores. The purpose of the research is to fulfill requirements of the Doctor of Educational Leadership and Policy Studies from Virginia Polytechnic Institute and State University.

The study will explore the relationship between math curriculum based measurement (M-CBM) benchmark assessments can predict fifth grade math SOL assessment scores. Data will be gathered on all fifth grade students who took the benchmark tests and the End-of-Course SOL test during the 2009-2010 school year. The researcher will gather the data from the Assistant Superintendent.

All information will remain confidential. There will be no reference to individual schools, school districts, or students. Student identification numbers will be used to match benchmark scores to SOL scores. Once the matching scores have been made, student identification numbers will be replaced with a random number.

Thank you for your consideration and if you have any questions, please contact me at (276) 608-9555.

Andy Cox, Principal
Abingdon Elementary School

APPENDIX B

INSTITUTIONAL REVIEW BOARD APPROVAL

VirginiaTech

Office of Research Compliance
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24060
540/231-4606 Fax 540/231-0959
e-mail irb@vt.edu
Website: www.irb.vt.edu

MEMORANDUM

DATE: April 1, 2011

TO: Theodore Creighton, Paul Cox

FROM: Virginia Tech Institutional Review Board (FWA00000572, expires October 26, 2013)

PROTOCOL TITLE: Comparison of Selected Benchmark Testing Methodologies as Predictors of Virginia Standards of Learning Test Scores

IRB NUMBER: 11-372

Effective April 1, 2011, the Virginia Tech IRB PAM, Andrea Nash, approved the new protocol for the above-mentioned research protocol.

This approval provides permission to begin the human subject activities outlined in the IRB-approved protocol and supporting documents.

Plans to deviate from the approved protocol and/or supporting documents must be submitted to the IRB as an amendment request and approved by the IRB prior to the implementation of any changes, regardless of how minor, except where necessary to eliminate apparent immediate hazards to the subjects. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

All investigators (listed above) are required to comply with the researcher requirements outlined at http://www.irb.vt.edu/pages/responsibilities.htm (please review before the commencement of your research).

PROTOCOL INFORMATION:
Approved as: **Exempt, under 45 CFR 46.101(b) category(ies) 2**
Protocol Approval Date: **4/1/2011**
Protocol Expiration Date: **NA**
Continuing Review Due Date*: **NA**
*Date a Continuing Review application is due to the IRB office if human subject activities covered under this protocol, including data analysis, are to continue beyond the Protocol Expiration Date.

FEDERALLY FUNDED RESEARCH REQUIREMENTS:
Per federally regulations, 45 CFR 46.103(f), the IRB is required to compare all federally funded grant proposals / work statements to the IRB protocol(s) which cover the human research activities included in the proposal / work statement before funds are released. Note that this requirement does not apply to Exempt and Interim IRB protocols, or grants for which VT is not the primary awardee.

The table on the following page indicates whether grant proposals are related to this IRB protocol, and which of the listed proposals, if any, have been compared to this IRB protocol, if required.

Invent the Future

APPENDIX C

SCHOOL DISTRICT A 2009-2010 ASSESSMENT DATA

District Code	School Code	Student Number	BM 1	BM 2	BM 3	SOL
1	1	1	52.6	42.1	47.4	359
1	1	2	X	21.1	x	415
1	1	3	100	84.2	68.4	515
1	1	4	31.6	68.4	x	498
1	1	5	21.1	42.1	31.6	390
1	1	6	47.4	36.8	x	445
1	1	7	31.6	57.9	x	374
1	1	8	36.8	47.4	79	482
1	1	9	89.5	73.7	79	600
1	1	10	94.7	31.6	52.6	469
1	1	11	100	89.5	94.7	600
1	1	12	31.6	68.4	73.7	407
1	1	13	42.1	94.7	73.7	561
1	1	14	36.8	57.9	73.7	482
1	1	15	89.5	84.2	94.7	600
1	1	16	100	89.5	100	600
1	1	17	84.2	79	94.7	596
1	1	18	79	36.8	73.7	536
1	1	19	84.2	x	89.5	561
1	1	20	100	52.6	84.2	596
1	1	21	42.1	26.3	57.9	515
1	1	22	73.7	100	94.7	600
1	1	23	31.6	42.1	89.5	596
1	1	24	21.1	36.8	52.6	382
1	1	25	94.7	100	84.2	596
1	1	26	89.5	84.2	52.6	596
1	1	27	x	52.6	68.4	561
1	1	28	36.8	26.3	57.9	469
1	1	29	94.7	79	89.5	561
1	1	30	79	94.7	89.5	536
1	1	31	42.1	47.4	52.6	482
1	1	32	21.1	57.9	x	482
1	1	33	84.2	68.4	52.6	498
1	1	34	94.7	89.5	x	482
1	1	35	84.2	73.7	89.5	596
1	1	36	26.3	21.1	63.2	425
1	1	37	89.5	31.6	63.2	457
1	1	38	47.4	100	94.7	596

1	1	39	89.5	79	94.7	600
1	1	40	94.7	100	100	469
1	1	41	31.6	x	84.2	515
1	1	42	21.1	68.4	x	435
1	1	43	79	94.7	x	498
1	1	44	36.8	57.9	42.1	435
1	1	45	94.7	52.6	79	561
1	1	46	100	100	100	600
1	1	47	42.1	47.4	x	265
1	1	48	31.6	42.1	57.9	425
1	1	49	84.2	36.8	73.7	561
1	1	50	89.5	84.2	79	561
1	1	51	36.8	94.7	68.4	536
1	1	52	42.1	100	89.5	596
1	1	53	26.3	36.8	57.9	435
1	1	54	36.8	57.9	10.5	232
1	1	55	94.7	89.5	x	596
1	2	1	36.8	57.9	94.7	294
1	2	2	89.5	100	94.7	600
1	2	3	21.1	26.3	47.4	374
1	2	4	89.5	79	84.2	596
1	2	5	94.7	84.2	89.5	600
1	2	6	31.6	68.4	47.4	536
1	2	7	47.4	21.1	52.6	445
1	2	8	42.1	31.6	x	482
1	2	9	x	68.4	79	536
1	2	10	36.8	42.1	73.7	469
1	2	11	84.2	100	79	596
1	2	12	94.7	x	84.2	536
1	2	13	94.7	89.5	79	596
1	2	14	31.6	52.6	73.7	600
1	2	15	89.5	26.3	63.2	536
1	2	16	100	84.2	94.7	596
1	2	17	79	36.8	52.6	561
1	2	18	94.7	100	89.5	600
1	2	19	42.1	89.5	79	561
1	2	20	89.5	68.4	73.7	600
1	2	21	21.1	47.4	36.8	382
1	2	22	52.6	31.6	84.2	301
1	2	23	36.8	52.6	73.7	515
1	2	24	94.7	79	79	600
1	2	25	89.5	100	100	600
1	2	26	x	89.5	94.7	600

1	2	27	26.3	84.2	79	482
1	2	28	36.8	x	57.9	435
1	2	29	84.2	89.5	94.7	600
1	2	30	94.7	57.9	x	323
1	2	31	31.6	57.9	36.8	415
1	2	32	47.4	31.6	21.1	344
1	2	33	89.5	52.6	79	425
1	2	34	84.2	79	63.2	482
1	2	35	52.6	36.8	47.4	435
1	2	36	100	89.5	68.4	596
1	2	37	89.5	100	68.4	536
1	2	38	79	42.1	47.4	498
1	2	39	42.1	26.3	63.2	498
1	2	40	36.8	79	57.9	498
1	2	41	100	84.2	73.7	515
1	2	42	47.4	94.7	57.9	515
1	2	43	42.1	47.4	63.2	482
1	2	44	31.6	68.4	42.1	425
1	2	45	68.4	84.2	63.2	515
1	2	46	84.2	100	47.4	536
1	2	47	94.7	89.5	63.2	536
1	2	48	52.6	47.4	42.1	398
1	2	49	57.9	x	52.6	398
1	2	50	94.7	100	52.6	600
1	2	51	26.3	31.6	57.9	415
1	2	52	47.4	63.2	36.8	398
1	2	53	21.1	57.9	x	390
1	2	54	36.8	47.4	x	445
1	3	1	52.6	36.8	x	330
1	3	2	31.6	68.4	15.8	457
1	3	3	x	x	0	398
1	3	4	x	57.9	26.3	337
1	3	5	21.1	47.4	31.6	482
1	3	6	94.7	42.1	63.2	515
1	3	7	31.6	79	47.4	515
1	3	8	100	84.2	63.2	515
1	3	9	47.4	36.8	52.6	435
1	3	10	89.5	94.7	57.9	515
1	3	11	36.8	68.4	79	435
1	3	12	42.1	21.1	26.3	316
1	3	13	79	52.6	57.9	482
1	3	14	47.4	68.4	57.9	435
1	3	15	94.7	89.5	73.7	596

1	3	16	94.7	100	89.5	515
1	3	17	42.1	57.9	47.4	359
1	3	18	79	31.6	47.4	482
1	3	19	36.8	47.4	42	425
1	3	20	x	84.2	94.7	600
1	3	21	84.2	36.8	57.9	469
1	3	22	31.6	57.9	52.6	359
1	3	23	57.9	26.3	16	294
1	3	24	52.6	47.4	63	425
1	3	25	26.3	x	57.9	482
1	3	26	89.5	42.1	63.2	515
1	3	27	47.4	36.8	63.2	498
1	3	28	84.2	89.5	57.9	515
1	3	29	100	79	84.2	600
1	3	30	21.1	57.9	15.79	425
1	3	31	63.2	31.6	68.4	398
1	3	32	42.1	68.4	42.1	435
1	3	33	79	94.7	73.7	515
1	3	34	36.8	52.6	42.1	469
1	3	35	94.7	89.5	52.3	561
1	3	36	47.4	79	84.2	515
1	3	37	79	31.6	63.2	457
1	3	38	36.8	73.7	52.6	435
1	3	39	52.6	36.8	63.16	457
1	3	40	100	52.6	89.5	600
1	4	1	89.5	100	100	482
1	4	2	100	89.5	x	596
1	4	3	26.3	94.7	89.5	445
1	4	4	94.7	100	x	600
1	4	5	52.6	x	94.7	596
1	4	6	94.7	26.3	x	561
1	4	7	94.7	89.5	100	482
1	4	8	100	79	84.2	536
1	4	9	x	94.7	89.5	536
1	4	10	x	100	100	469
1	4	11	100	89.5	94.7	536
1	4	12	84.2	100	94.5	600
1	4	13	89.5	84.2	94.7	600
1	4	14	100	100	100	536
1	4	15	100	89.5	100	498
1	4	16	84.2	100	100	536
1	4	17	31.6	79	x	469
1	4	18	47.4	94.7	100	561

1	4	19	100	100	100	498
1	4	20	79	94.7	89.5	596
1	4	21	89.5	84.2	x	536
1	4	22	100	100	94.5	536
1	4	23	94.7	79	x	498
1	4	24	36.8	79	100	515
1	4	25	42.1	21.1	x	435
1	4	26	89.5	100	94.7	482
1	4	27	94.7	79	89.5	515
1	4	28	100	79	100	374
1	4	29	84.2	100	100	536
1	4	30	100	89.5	94.5	536
1	4	31	100	89.5	100	600
1	4	32	94.7	100	100	561
1	4	33	21.1	31.6	x	382
1	4	34	100	94.7	x	600
1	4	35	x	100	x	482
1	4	36	94.7	84.2	x	469

School District B 2009-2010 Assessment Data

District Code	School Code	Student Number	BM 1	BM 2	BM 3	SOL
2	5	1	75	100	89	594
2	5	2	70	73	66	560
2	5	3	95	97	93	560
2	5	4	65	70	40	394
2	5	5	75	87	77	560
2	5	6	60	77	51	600
2	5	7	65	77	66	560
2	5	8	100	97	100	560
2	5	9	75	60	63	410
2	5	10	70	80	77	560
2	5	11	65	73	70	484
2	5	12	90	93	88	560
2	5	13	85	100	96	600
2	5	14	35	43	40	300
2	5	15	80	97	77	560
2	5	16	45	70	55	364
2	5	17	X	80	70	459
2	5	18	75	93	88	600
2	5	19	70	97	85	535
2	5	20	40	83	55	459
2	5	21	60	80	66	293
2	5	22	35	70	51	402
2	5	23	X	40	66	300
2	6	1	75	x	96	600
2	6	2	70	94	84	560
2	6	3	95	82	84	560
2	6	4	65	85	92	364
2	6	5	70	67	92	535
2	6	6	60	91	80	600
2	6	7	60	76	80	535
2	6	8	100	100	96	560
2	6	9	75	x	X	402
2	6	10	75	82	88	560
2	6	11	60	76	76	459
2	6	12	85	91	96	535
2	6	13	85	85	96	600
2	6	14	45	55	72	402
2	6	15	85	79	80	459
2	6	16	50	70	76	293

2	6	17	75	88	68	402
2	6	18	65	x	92	600
2	6	19	75	x	100	600
2	6	20	50	76	80	560
2	6	21	65	70	76	402
2	6	22	30	61	64	293
2	6	23	X	76	72	364
2	6	24	45	88	76	293
2	6	25	80	88	92	600
2	6	26	80	100	92	560
2	6	27	90	100	100	560
2	6	28	100	100	100	600
2	6	29	100	88	76	402
2	6	30	55	82	80	560
2	6	31	65	85	76	459
2	7	1	90	97	X	459
2	7	2	90	88	96	459
2	7	3	60	82	84	402
2	7	4	35	76	68	364
2	7	5	25	73	80	394
2	7	6	60	97	100	594
2	7	7	90	82	96	535
2	7	8	30	55	X	293
2	7	9	45	73	80	300
2	7	10	55	73	68	394
2	7	11	85	88	72	459
2	7	12	85	97	100	560
2	7	13	35	80	76	394
2	7	14	100	88	96	535
2	7	15	100	96	80	594
2	7	16	60	85	80	410
2	7	17	70	76	52	394
2	7	18	80	88	96	459
2	7	19	80	79	64	300
2	7	20	90	88	92	594
2	7	21	95	97	96	560
2	7	22	95	73	80	459
2	7	23	95	82	92	560
2	7	24	100	84	100	560
2	7	25	75	84	76	402
2	7	26	75	55	84	459
2	7	27	80	88	80	459
2	7	28	95	88	X	594

2	7	29	75	80	X	293
2	7	30	80	82	76	410
2	7	31	80	91	100	535
2	7	32	75	73	80	410
2	7	33	100	97	80	560
2	7	34	80	82	76	402
2	7	35	90	100	96	560
2	7	36	80	80	84	410
2	7	37	85	97	X	x
2	7	38	90	84	80	459
2	7	39	100	100	92	600
2	7	40	90	85	100	535
2	7	41	95	100	100	600
2	7	42	85	97	92	594
2	7	43	85	88	88	458
2	7	44	75	80	84	402
2	7	45	80	84	100	594
2	8	1	100	97	100	560
2	8	2	70	55	72	394
2	8	3	100	88	96	600
2	8	4	95	91	80	402
2	8	5	100	91	96	484
2	8	6	55	x	X	364
2	8	7	70	79	84	410
2	8	8	55	49	68	364
2	8	9	95	100	92	560
2	8	10	100	97	96	459
2	8	11	50	61	X	293
2	8	12	90	100	100	600
2	8	13	90	91	96	560
2	8	14	60	64	56	394
2	8	15	100	100	92	560
2	8	16	70	85	72	459
2	8	17	95	97	92	535
2	8	18	65	82	84	410
2	8	19	65	79	84	402
2	8	20	95	82	84	535
2	8	21	100	85	100	459
2	9	1	100	97	88	498
2	9	2	90	79	84	498
2	9	3	90	94	100	600
2	9	4	75	79	76	386
2	9	5	70	79	64	394

2	9	6	80	91	88	515
2	9	7	80	82	80	560
2	9	8	80	100	88	594
2	9	9	45	67	68	357
2	9	10	100	100	100	600
2	9	11	100	85	88	448
2	9	12	85	91	96	484
2	9	13	90	94	92	594
2	9	14	75	76	88	394
2	9	15	60	64	56	438
2	9	16	75	82	76	386
2	9	17	85	94	100	600
2	9	18	80	82	96	560
2	9	19	80	82	72	471
2	9	20	90	100	88	594
2	9	21	90	91	88	448
2	9	22	95	79	68	394
2	9	23	100	85	92	459
2	9	24	100	88	92	498
2	9	25	90	76	84	535
2	9	26	90	94	100	594
2	9	27	85	100	88	484
2	9	28	85	x	96	515
2	9	29	80	88	84	600
2	9	30	80	58	64	379
2	9	31	80	79	80	394
2	9	32	70	91	88	535
2	9	33	55	49	68	394
2	9	34	65	73	64	364
2	9	35	95	82	72	484
2	9	36	x	100	96	498
2	9	37	100	100	88	600
2	9	38	90	88	76	515
2	9	39	100	100	80	600
2	9	40	95	88	100	471
2	9	41	90	82	84	594
2	9	42	85	85	80	459
2	9	43	85	91	88	560
2	9	44	80	88	84	438
2	9	45	75	67	68	438
2	9	46	90	97	88	498
2	9	47	85	94	92	594
2	9	48	90	76	76	394

2	9	49	x	91	84	448
2	9	50	100	91	100	600
2	9	51	80	85	92	535
2	9	52	80	85	72	515
2	9	53	95	76	96	459
2	9	54	100	82	80	560
2	9	55	90	91	100	600
2	9	56	90	79	84	471
2	9	57	90	88	84	438
2	9	58	85	88	80	484
2	9	59	85	94	84	515
2	9	60	90	91	92	594
2	9	61	80	76	64	394
2	9	62	x	88	72	448
2	9	63	75	85	68	379
2	9	64	90	82	100	498
2	9	65	65	67	64	471
2	9	66	100	100	84	594
2	9	67	90	85	80	484
2	9	68	80	73	76	438
2	9	69	85	x	56	419
2	9	70	90	91	84	459
2	9	71	90	94	84	560
2	9	72	85	88	80	535
2	9	73	75	67	68	394
2	9	74	80	82	76	x
2	9	75	75	x	72	459
2	9	76	100	97	100	600
2	9	77	100	91	80	535
2	9	78	100	88	88	594
2	9	79	90	88	92	498
2	9	80	85	85	80	515
2	9	81	85	70	72	419
2	9	82	95	100	100	560
2	9	83	55	61	60	335
2	9	84	75	67	64	394
2	9	85	100	94	88	438
2	9	86	70	85	60	471
2	9	87	85	79	80	438
2	9	88	80	64	68	364
2	9	89	90	88	84	535
2	9	90	90	97	80	515
2	9	91	95	85	100	515

2	9	92	90	79	72	459
2	9	93	100	100	96	594
2	9	94	100	x	96	498
2	9	95	90	x	80	x
2	9	96	90	82	88	459
2	9	97	75	79	68	379
2	9	98	80	82	92	448
2	9	99	85	88	88	484
2	9	100	50	73	64	394
2	9	101	100	94	100	560
2	9	102	90	97	88	594
2	9	103	85	82	92	438
2	9	104	80	85	80	484
2	9	105	75	88	72	379
2	9	106	80	79	76	419
2	10	1	85	91	100	535
2	10	2	100	97	92	594
2	10	3	95	100	96	600
2	10	4	75	70	64	x
2	10	5	50	46	52	335
2	10	6	95	94	100	560
2	10	7	70	76	84	419
2	10	8	95	97	92	600
2	10	9	x	67	76	419
2	10	10	75	82	88	498
2	10	11	90	100	100	600
2	10	12	100	91	96	594
2	10	13	85	97	80	515
2	10	14	95	88	100	535
2	10	15	95	79	100	484
2	10	16	65	49	X	379
2	10	17	85	94	96	471
2	10	18	90	91	88	515
2	10	19	85	100	92	594
2	10	20	65	x	72	386
2	10	21	75	85	84	438
2	10	22	85	94	100	535
2	10	23	100	91	92	438
2	10	24	95	88	100	560
2	10	25	100	97	80	535
2	10	26	100	100	92	594
2	10	27	70	100	88	484
2	10	28	70	64	76	357

2	10	29	85	97	100	535
2	10	30	75	94	84	438
2	10	31	75	88	80	471
2	10	32	45	61	48	335
2	10	33	50	58	60	364
2	10	34	95	79	84	600
2	10	35	85	100	100	535
2	10	36	75	82	68	394
2	10	37	70	67	60	419
2	10	38	95	91	92	535
2	10	39	70	73	88	x
2	10	40	80	94	80	515
2	10	41	65	67	56	438
2	10	42	45	55	48	379
2	10	43	90	100	100	594
2	10	44	85	100	92	560
2	10	45	95	100	96	600
2	10	46	85	85	72	419
2	10	47	100	88	100	594
2	10	48	x	76	72	386
2	10	49	75	70	76	419
2	10	50	65	49	52	357
2	10	51	85	79	88	484
2	10	52	75	64	68	438
2	10	53	80	91	92	498
2	10	54	55	52	64	379
2	10	55	90	85	84	448
2	10	56	95	100	100	600
2	10	57	95	79	96	535
2	10	58	100	85	96	560
2	10	59	100	100	88	594
2	10	60	50	52	56	386
2	10	61	100	100	92	535
2	10	62	100	94	100	594
2	10	63	85	76	72	419
2	10	64	55	x	48	394
2	10	65	75	61	60	419
2	10	66	70	73	76	448
2	10	67	80	97	80	471
2	10	68	45	55	52	379
2	10	69	70	70	60	419
2	10	70	95	94	92	600
2	10	71	85	67	72	419

2	10	72	75	85	80	419
2	10	73	75	76	68	438
2	10	74	80	88	96	560
2	10	75	65	64	56	379
2	10	76	70	82	76	448
2	10	77	65	76	72	419
2	10	78	95	100	100	594
2	10	79	80	79	88	471
2	10	80	65	70	76	438
2	10	81	100	91	96	515
2	10	82	85	100	92	535
2	10	83	100	91	100	594
2	10	84	80	94	84	438
2	10	85	45	52	48	335
2	10	86	80	97	88	600
2	10	87	45	46	52	364
2	10	88	75	82	X	x
2	10	89	45	58	X	379
2	10	90	65	73	72	394
2	10	91	85	94	96	594
2	10	92	85	79	76	419
2	10	93	x	52	64	357
2	10	94	70	85	80	448
2	10	95	80	88	84	560
2	10	96	75	67	68	448
2	10	97	90	100	96	535
2	10	98	80	67	76	438
2	10	99	90	91	96	515
2	10	100	70	61	56	386
2	10	101	85	x	80	419
2	10	102	100	97	88	535
2	10	103	85	100	100	600
2	10	104	100	100	96	594
2	10	105	100	82	100	594
2	10	106	95	79	100	600
2	10	107	50	55	48	379
2	10	108	65	64	60	364
2	10	109	85	94	92	471
2	10	110	85	91	88	498
2	10	111	80	88	88	438
2	10	112	80	76	72	448
2	10	113	75	67	68	448
2	10	114	70	58	64	379

2	10	115	55	49	52	357
2	10	116	85	100	96	560
2	10	117	55	52	56	379
2	10	118	100	97	92	600
2	10	119	85	76	76	448
2	10	120	60	55	X	379
2	10	121	80	85	84	471
2	10	122	75	70	72	438
2	10	123	65	58	60	419
2	10	124	75	67	64	438
2	10	125	90	100	92	600
2	10	126	85	97	88	594
2	10	127	65	79	80	394
2	10	128	55	64	72	419
2	10	129	65	73	60	386
2	10	130	55	52	48	379
2	10	131	95	94	100	594
2	10	132	100	91	80	560
2	10	133	85	97	92	515
2	10	134	75	61	68	419
2	10	135	50	52	56	364
2	10	136	55	46	56	364
2	10	137	75	82	80	419
2	10	138	65	58	56	386
2	10	139	80	100	100	594
2	10	140	75	88	92	484
2	10	141	95	88	96	600
2	10	142	70	64	76	448
2	10	143	45	55	52	386
2	10	144	65	73	64	438
2	10	145	85	100	92	560
2	10	146	90	94	100	535
2	10	147	80	85	88	419
2	10	148	75	82	84	498
2	10	149	75	79	X	x
2	10	150	100	91	100	600
2	10	151	65	x	68	364
2	10	152	70	70	72	419
2	10	153	100	94	100	560
2	10	154	85	100	96	600
2	10	155	90	100	92	594

School District C 2009-2010 Assessment Data

District Code	School Code	Student Number	BM 1	BM 2	BM 3	SOL
3	11	1	x	68	91	560
3	11	2	100	92	87	594
3	11	3	57	88	X	335
3	11	4	71	x	X	x
3	11	5	96	84	100	535
3	11	6	96	88	91	448
3	11	7	93	80	96	448
3	11	8	100	100	100	600
3	11	9	96	88	87	515
3	11	10	96	92	100	600
3	11	11	x	88	70	484
3	11	12	100	100	96	600
3	11	13	96	92	100	459
3	11	14	100	100	100	535
3	11	15	100	96	87	498
3	11	16	93	80	91	560
3	11	17	x	92	96	535
3	11	18	86	68	83	438
3	11	19	x	84	96	560
3	11	20	93	80	83	535
3	11	21	96	88	100	594
3	11	22	68	68	91	448
3	11	23	x	96	78	357
3	11	24	x	76	91	498
3	11	25	96	100	96	448
3	11	26	96	100	91	560
3	11	27	82	88	91	498
3	11	28	93	76	91	535
3	11	29	86	92	91	515
3	11	30	x	96	96	515
3	11	31	93	96	96	600
3	11	32	x	96	96	515
3	11	33	93	92	96	600
3	11	34	x	88	65	459
3	11	35	100	100	91	459
3	11	36	96	96	100	560
3	11	37	93	x	65	386
3	11	38	96	100	91	535
3	11	39	96	92	96	600

3	11	40	x	96	83	419
3	11	41	x	68	83	471
3	11	42	86	76	78	419
3	11	43	96	96	100	594
3	11	44	79	92	91	448
3	11	45	86	68	78	471
3	11	46	57	x	x	x
3	11	47	89	92	100	535
3	11	48	96	92	91	600
3	11	49	100	96	87	515
3	11	50	86	88	91	594
3	11	51	96	88	100	535
3	11	52	89	84	61	379
3	11	53	89	76	74	560
3	11	54	100	100	96	600
3	11	55	86	88	100	448
3	11	56	93	92	100	515
3	11	57	79	64	91	394
3	11	58	x	x	x	484
3	11	59	86	88	100	535
3	11	60	100	96	96	600
3	11	61	82	84	96	419
3	11	62	96	92	78	515
3	11	63	96	96	83	364
3	11	64	89	80	96	560
3	11	65	100	96	100	594
3	11	66	x	x	30	335
3	11	67	100	100	91	535
3	11	68	96	76	78	515
3	11	69	79	92	91	459
3	11	70	96	x	91	498
3	12	1	x	80	61	498
3	12	2	96	92	78	484
3	12	3	100	92	87	594
3	12	4	100	96	87	560
3	12	5	82	88	78	600
3	12	6	89	88	96	535
3	12	7	x	88	96	515
3	12	8	100	84	96	515
3	12	9	96	92	96	560
3	12	10	93	96	96	535
3	12	11	96	92	78	448
3	12	12	100	x	100	594

3	12	13	82	84	87	459
3	12	14	96	80	87	594
3	12	15	50	64	65	535
3	12	16	89	76	83	600
3	12	17	79	88	87	515
3	12	18	96	92	96	600
3	12	19	79	92	96	594
3	12	20	96	100	78	594
3	12	21	89	100	87	600
3	12	22	86	84	57	560
3	12	23	x	x	39	342
3	12	24	71	88	91	498
3	12	25	89	72	78	535
3	12	26	64	48	78	vaap
3	12	27	82	72	83	vaap
3	12	28	82	60	78	vaap
3	12	29	79	72	74	498
3	12	30	57	72	74	459
3	12	31	79	84	70	498
3	12	32	68	84	83	419
3	12	33	64	x	100	471
3	12	34	86	80	74	498
3	12	35	61	68	x	vaap
3	12	36	79	96	78	515
3	12	37	79	96	91	560
3	12	38	89	76	96	535
3	12	39	82	x	100	594
3	12	40	64	80	65	459
3	12	41	93	92	91	600
3	12	42	75	72	83	498
3	12	43	75	80	96	560
3	12	44	93	96	83	560
3	12	45	x	88	87	600
3	12	46	79	x	96	515
3	12	47	89	92	74	498
3	12	48	86	96	100	515
3	13	1	54	56	70	357
3	13	2	86	88	70	471
3	13	3	89	92	74	471
3	13	4	86	88	78	484
3	13	5	79	92	87	600
3	13	6	79	88	70	471
3	13	7	50	x	x	moved

3	13	8	89	84	70	364
3	13	9	86	88	78	459
3	13	10	x	80	87	600
3	13	11	x	x	x	moved
3	13	12	x	100	x	419
3	13	13	86	84	65	438
3	13	14	x	60	39	394
3	13	15	68	84	x	448
3	13	16	x	x	74	438
3	13	17	89	100	91	535
3	13	18	82	92	87	564
3	13	19	82	88	70	515
3	13	20	100	92	83	600
3	13	21	82	84	87	484
3	13	22	75	76	57	448
3	13	23	86	48	43	386
3	13	24	82	92	83	535
3	13	25	x	92	30	428
3	13	26	39	x	x	vaap
3	13	27	50	72	70	459
3	13	28	32	x	x	vaap
3	13	29	93	96	96	600
3	13	30	82	80	70	535
3	13	31	50	x	x	vaap
3	13	32	x	96	74	459
3	13	33	75	72	x	402
3	13	34	x	52	74	438
3	13	35	79	84	87	448
3	13	36	x	96	91	594
3	13	37	75	84	91	484
3	13	38	x	100	91	600
3	13	39	79	72	70	471
3	13	40	x	80	x	498
3	13	41	x	x	x	vaap
3	13	42	43	56	52	293
3	13	43	36	84	65	410
3	13	44	93	92	83	535
3	13	45	x	x	91	535
3	13	46	68	x	70	419
3	13	47	82	92	70	535
3	13	48	x	x	91	515
3	13	49	71	80	65	535
3	13	50	86	92	74	438

3	13	51	89	96	78	535
3	13	52	x	96	x	448
3	13	53	x	68	26	vaap
3	13	54	89	92	78	594
3	13	55	x	92	65	438
3	13	56	79	72	61	448
3	13	57	93	96	87	515
3	13	58	x	92	83	438
3	13	59	79	72	78	419
3	13	60	54	68	70	419
3	13	61	64	84	74	459
3	13	62	96	84	78	484
3	13	63	64	56	70	357
3	13	64	x	80	74	448
3	13	65	x	x	x	394
3	13	66	64	68	26	438
3	13	67	86	84	96	560
3	13	68	x	92	96	560
3	13	69	68	84	70	471
3	13	70	96	96	96	594
3	13	71	82	88	83	459
3	13	72	x	76	39	342
3	13	73	x	92	91	594
3	13	74	x	80	78	459
3	13	75	79	88	43	438
3	13	76	82	80	91	471
3	13	77	75	68	74	428
3	13	78	x	84	74	428
3	13	79	x	96	87	600
3	13	80	x	96	83	594
3	13	81	57	52	48	342
3	13	82	93	100	96	535
3	13	83	x	88	x	moved
3	13	84	71	80	87	471
3	14	1	54	76	x	357
3	14	2	82	92	70	428
3	14	3	x	32	x	vgla
3	14	4	x	x	x	moved
3	14	5	64	48	x	vgla
3	14	6	93	80	83	515
3	14	7	36	44	61	307
3	14	8	71	92	87	471
3	14	9	x	80	100	560

3	14	10	x	72	87	535
3	14	11	x	56	61	410
3	14	12	89	92	70	379
3	14	13	82	80	83	419
3	14	14	50	56	x	vgla
3	14	15	54	48	52	293
3	14	16	68	64	61	428
3	14	17	50	44	x	vgla
3	14	18	61	64	x	402
3	14	19	71	48	91	394
3	14	20	x	96	96	600
3	14	21	75	76	78	471
3	14	22	68	56	87	342
3	14	23	100	96	96	594
3	14	24	82	72	87	471
3	14	25	100	96	96	600
3	14	26	x	96	78	498
3	14	27	89	96	96	560
3	14	28	96	96	100	600
3	14	29	93	92	87	535
3	14	30	57	64	83	357
3	14	31	96	88	65	419
3	14	32	x	88	83	moved
3	14	33	96	92	96	560
3	14	34	71	88	83	515
3	14	35	x	96	91	459
3	14	36	82	96	78	448
3	14	37	82	92	78	471
3	14	38	x	96	91	515
3	14	39	93	96	87	515
3	14	40	100	88	100	535
3	14	41	75	52	48	314
3	14	42	50	48	x	vgla
3	14	43	x	44	x	vgla
3	14	44	86	60	74	515
3	14	45	57	40	57	307
3	14	46	x	80	52	438
3	14	47	89	72	83	484
3	14	48	86	68	74	535
3	14	49	25	64	x	vgla
3	14	50	71	68	39	402
3	14	51	50	40	52	484
3	14	52	86	80	74	471

3	14	53	71	80	74	402
3	14	54	82	84	78	515
3	14	55	46	44	39	328
3	14	56	86	96	48	459
3	14	57	79	36	x	vgla
3	14	58	79	76	70	428
3	14	59	82	68	74	428
3	14	60	93	92	87	415
3	14	61	68	76	57	410
3	14	62	57	76	57	535
3	14	63	61	72	74	438
3	14	64	86	100	96	600
3	14	65	82	92	100	600
3	14	66	79	76	83	410
3	14	67	82	80	83	498
3	14	68	89	x	83	594
3	14	69	x	92	74	498
3	14	70	89	96	96	600
3	14	71	x	96	83	560
3	14	72	89	76	61	471
3	14	73	86	88	83	515
3	14	74	96	96	100	600
3	14	75	61	72	57	394
3	14	76	61	76	70	moved
3	14	77	96	88	100	498
3	14	78	86	100	96	594
3	14	79	100	88	96	600
3	14	80	82	84	74	515
3	14	81	57	80	78	471
3	14	82	86	100	87	594
3	14	83	89	84	96	498
3	14	84	93	96	96	594
3	14	85	82	84	87	498
3	14	86	89	84	96	594
3	14	87	86	88	78	535
3	14	88	89	100	100	600
3	14	89	86	92	61	484
3	14	90	x	88	91	600
3	14	91	71	88	74	484
3	14	92	71	72	91	484
3	14	93	100	92	100	600
3	14	94	93	96	96	600
3	14	95	93	96	100	560

3	14	96	86	92	74	515
3	14	97	68	72	78	515
3	14	98	100	100	100	600
3	14	99	86	92	87	600
3	14	100	x	92	74	600
3	14	101	75	84	74	515
3	14	102	x	56	57	410
3	14	103	43	56	x	364
3	14	104	39	64	78	419
3	14	105	71	88	91	488
3	15	1	93	76	78	448
3	15	2	93	80	87	600
3	15	3	100	100	100	600
3	15	4	93	92	91	515
3	15	5	100	92	91	600
3	15	6	93	96	91	498
3	15	7	93	84	87	515
3	15	8	89	92	96	594
3	15	9	82	84	87	560
3	15	10	89	88	96	560
3	15	11	100	92	87	535
3	15	12	96	80	78	471
3	15	13	86	80	91	515
3	15	14	93	76	83	594
3	15	15	86	84	96	594
3	15	16	100	100	96	600
3	15	17	96	92	87	535
3	15	18	96	88	96	498
3	15	19	93	96	100	594
3	15	20	93	96	100	600
3	15	21	100	100	96	594
3	15	22	x	100	96	535
3	15	23	100	96	87	560
3	15	24	86	56	74	515
3	15	25	93	96	91	600
3	15	26	93	92	87	560
3	15	27	96	96	87	600
3	15	28	86	96	87	600
3	15	29	93	100	91	515
3	15	30	93	92	74	594
3	15	31	86	80	87	484
3	15	32	89	88	83	448
3	15	33	93	96	87	594

3	15	34	100	96	96	600
3	15	35	89	96	100	600
3	15	36	96	92	91	600
3	15	37	93	96	96	560
3	15	38	89	88	91	560
3	15	39	x	x	x	vgla
3	15	40	100	88	96	560
3	15	41	82	80	74	459
3	15	42	96	84	91	600
3	15	43	100	100	87	594
3	15	44	x	84	83	484
3	15	45	100	96	100	535
3	15	46	75	88	83	535
3	15	47	100	80	91	560
3	15	48	100	100	91	600
3	15	49	96	100	96	600
3	15	50	96	88	96	535
3	15	51	89	92	96	594
3	15	52	x	x	x	vgla
3	15	53	68	76	x	428
3	15	54	75	80	78	459
3	15	55	89	68	65	515
3	15	56	68	96	65	419
3	15	57	96	100	96	600
3	15	58	89	92	96	594
3	15	59	96	96	100	594
3	15	60	x	72	43	484
3	15	61	100	96	91	515
3	15	62	100	92	91	600
3	15	63	93	92	96	484
3	15	64	93	92	96	535
3	15	65	86	68	87	498
3	15	66	100	96	87	484
3	15	67	89	96	83	535
3	15	68	93	92	83	600
3	15	69	89	92	87	600
3	15	70	86	80	96	484
3	15	71	96	92	91	600
3	15	72	86	96	100	535
3	15	73	86	84	70	438
3	16	1	71	76	87	594
3	16	2	71	92	91	484
3	16	3	x	76	x	moved

3	16	4	93	88	91	560
3	16	5	93	96	100	515
3	16	6	75	92	83	515
3	16	7	64	88	83	594
3	16	8	71	76	87	471
3	16	9	89	96	100	535
3	16	10	96	92	100	535
3	16	11	79	88	78	535
3	16	12	89	88	83	484
3	16	13	82	92	87	594
3	16	14	96	88	100	600
3	16	15	75	92	74	459
3	16	16	96	100	100	560
3	16	17	82	88	100	600
3	16	18	71	88	83	459
3	16	19	82	72	74	459
3	16	20	89	76	74	438
3	16	21	71	88	74	498
3	16	22	x	88	78	419
3	16	23	64	80	83	471
3	16	24	x	x	83	515
3	16	25	79	72	74	410
3	16	26	x	84	87	535
3	16	27	96	100	91	600
3	16	28	79	76	78	484
3	16	29	86	92	91	515
3	16	30	96	92	96	600
3	16	31	86	100	91	594
3	16	32	75	92	70	459
3	16	33	96	84	91	471
3	16	34	79	88	87	459
3	16	35	68	80	96	535
3	16	36	79	88	91	484
3	16	37	71	92	87	448
3	16	38	75	72	70	535
3	16	39	79	84	91	535
3	16	40	50	80	74	498
3	16	41	61	80	91	484
3	16	42	93	92	100	594
3	16	43	82	92	87	594
3	16	44	71	80	83	459
3	16	45	x	x	83	535
3	16	46	54	68	87	438

3	16	47	71	84	74	moved
3	16	48	75	92	83	402
3	16	49	86	84	91	515
3	16	50	89	88	91	600
3	16	51	68	84	87	386
3	16	52	x	92	78	560
3	16	53	93	96	100	535
3	16	54	79	84	96	560
3	16	55	93	84	83	515
3	17	1	46	52	30	x
3	17	2	86	88	83	535
3	17	3	46	52	78	x
3	17	4	93	96	96	594
3	17	5	89	96	91	515
3	17	6	79	88	91	438
3	17	7	89	100	96	600
3	17	8	71	96	70	484
3	17	9	100	100	96	600
3	17	10	68	x	x	x
3	17	11	96	100	91	x
3	17	12	96	92	91	x
3	17	13	100	100	100	600
3	17	14	75	96	x	x
3	17	15	93	96	87	600
3	17	16	x	76	74	438
3	17	17	71	88	91	x
3	17	18	79	84	61	438
3	17	19	96	96	91	600
3	17	20	x	52	x	x
3	17	21	96	100	100	594
3	17	22	71	76	65	498
3	17	23	64	76	48	419
3	17	24	100	100	87	600
3	17	25	93	84	74	515
3	17	26	86	80	x	594
3	17	27	93	96	100	591
3	17	28	x	100	74	x
3	17	29	96	96	100	600
3	17	30	x	84	74	515
3	17	31	93	84	x	600
3	17	32	89	88	100	535
3	17	33	96	92	87	594
3	17	34	96	84	83	600

3	17	35	96	96	91	x
3	17	36	x	88	96	560
3	17	37	96	100	96	560
3	17	38	x	72	78	484
3	17	39	93	100	78	419
3	17	40	64	88	83	535
3	17	41	79	88	78	498
3	17	42	79	92	74	594
3	17	43	89	92	96	600
3	17	44	82	92	83	600
3	17	45	64	80	87	448
3	17	46	100	100	83	600
3	17	47	89	100	91	594
3	17	48	89	88	78	515
3	17	49	89	80	83	600
3	17	50	86	92	83	x
3	17	51	93	92	83	515
3	17	52	75	60	x	471
3	17	53	79	84	83	535
3	17	54	82	88	78	515
3	17	55	x	x	x	x
3	17	56	71	72	70	560
3	17	57	79	92	61	515
3	17	58	68	88	83	594
3	17	59	82	64	57	448
3	17	60	100	100	96	600
3	17	61	86	96	83	594
3	17	62	89	92	96	600
3	17	63	64	80	74	410
3	17	64	64	88	83	498
3	17	65	64	60	57	386
3	17	66	93	80	83	498
3	17	67	100	96	100	600
3	17	68	79	92	78	459
3	17	69	100	92	91	560
3	17	70	96	92	91	515
3	17	71	100	100	96	600
3	17	72	100	96	65	535
3	17	73	46	x	x	x
3	17	74	61	x	x	vaap
3	17	75	75	88	87	594
3	17	76	79	88	83	498
3	17	77	86	76	78	535

3	17	78	50	x	52	x
3	17	79	79	76	74	471
3	17	80	57	92	30	vaap
3	17	81	x	80	61	459
3	17	82	82	84	x	515
3	17	83	75	92	61	471
3	17	84	93	88	78	459
3	17	85	96	100	100	600
3	17	86	82	84	x	x
3	17	87	96	96	78	498
3	17	88	75	84	87	560
3	17	89	75	96	65	594
3	17	90	100	92	87	x
3	17	91	61	96	x	448
3	17	92	75	76	x	x
3	17	93	93	100	87	535
3	17	94	89	88	78	535
3	17	95		96	87	600
3	17	96	79	72	61	498
3	17	97	57	x	x	vaap
3	17	98	82	80	x	515
3	17	99	89	100	96	560
3	17	100	61	68	74	484
3	17	101	93	88	78	594
3	17	102	82	96	87	471