AN ABSTRACT OF THE THESIS OF

_____Kendra R. Gross_____ for the _____Specialist in Education_____

in _____School Psychology_____ presented on _____May 5, 2004_____

Title: _An Examination of Curriculum-Based Measurement as a Predictor of State_____

Assessment Proficiency Levels_____

Abstract approved: _[signature]_____

This study examined the utility of curriculum-based measures (CBM) of reading to

predict proficiency levels on the Kansas Reading Assessments. The CBM reading probes

were administered during the spring semester to third grade students. The Kansas

Reading Assessments were administered to the same students during the spring semester

of Grade 4 ($n = 163$). The utility of CBM mathematics to predict proficiency levels on

the Kansas Mathematics Assessment was also investigated. The CBM mathematics

probes were given in the spring semester to fourth grade students. A year later, these

same students, now in Grade 5, received the Kansas Mathematics Assessment ($n = 236$).

Statistical analyses were performed using Pearson Product Moment correlations as well

as separate Z tests to investigate potential differences in the correlations between groups

such as gender, race, and socio-economic status. Significant correlations were found

between both CBM measures and both Kansas State Assessments; however, no

significant differences in the CBM measures were found between gender, race, and socio-

economic status. The findings of the study indicate CBM reading and mathematics

probes predict the proficiency levels on the Kansas Reading and Mathematics

Assessment, respectively. CBM can be used as a predictive tool to determine students

who are at risk of failing on high-stakes tests. The clinical utility of these CBM measures

is discussed.

AN EXAMINATION OF CURRICULUM-BASED MEASUREMENT AS A

PREDICTOR OF STATE ASSESSMENT PROFICIENCY LEVELS

———————

A Thesis

Presented to

The Department of Psychology and Special Education

EMPORIA STATE UNIVERSITY

———————

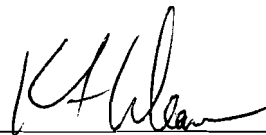In Partial Fulfillment

of the Requirements for the Degree
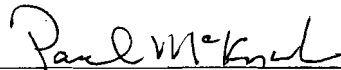
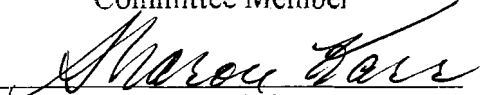Specialist in Education

———————

by

Kendra R. Gross

August 2004

_____

Approved for the Department of
Psychology and Special Education

_____
Committee Member

_____
Committee Member

_____
Committee Chairperson

_____
Approved for the Graduate Council

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee members, Dr. Sharon Karr, Dr. Jim Persinger, and Dr. Paul McKnab for their knowledge and instructional guidance in the writing of this thesis. My deepest thanks go to Dr. Sharon Karr for the time and effort that she has invested in this project, as well as her continued support, guidance, and understanding during my education at Emporia State. I also want to thank Dr. Ken Weaver for his statistical knowledge and help in revising this thesis. I would also like to thank Deborah McVey for her assistance in helping me obtain the data for this thesis and for the helpful advice she has given me throughout the process. My sincere appreciation also goes out to Rhett Dubiel for his constant support, patience, and encouragement during the writing of this thesis. Lastly, I would like to thank my parents for their love and support as I continued my education.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

There are many important and challenging issues facing educators in today's

society. One important issue in the field of special education is the topic of assessment

procedures. The traditional assessment procedures used to evaluate student progress are

being questioned as to whether they are adequately meeting the educational needs of

students.

Standardized tests are commonly being used to determine the levels at which

students are performing. Although this method of assessment is popular among schools, it

also has some criticisms. One of the main criticisms of standardized testing is that there is

a mismatch between the curriculum that is being taught and the test content, which

creates problems with the interpretation of the testing results (Sibley, Biwer, & Hesch,

2001). Another concern is that teachers are feeling compelled to "teach to the test" as a

result of the extreme pressures placed upon school districts for students to perform well

on standardized tests (Sibley et al., 2001).

Problems existing in the area of standardized testing and assessment indicate the

need to examine alternative assessment procedures. One alternative assessment procedure

that continues to gain respect is curriculum-based measurement (CBM). CBM is a

process used to help improve student achievement. An important strength of CBM is that

its main goal is to evaluate student instructional outcomes. This type of assessment has a

more individual focus and measures skills directly, which addresses one of the problems

with traditional assessment procedures. Although traditional assessment can be highly

useful for specific purposes such as screening and evaluation of programs, it is also

problematic in the area of special education and the decision making process. More

assessments now provide evidence that students are struggling with curriculum skills, which is where CBM becomes beneficial (Sibley et al., 2001). The information taken from initial CBM assessments provides objective data that can be incorporated into the development of appropriate intervention plans directly related to a student's curriculum (Fewster & MacMillan, 2002).

CBM is useful because it is taken from the curriculum that students are being exposed to in the classroom. This type of assessment samples the content of what is being taught. Therefore, it is easier to identify the needs of individual children. It is a beneficial component of instruction because it can be used for both progress monitoring and program effectiveness (Fewster & MacMillan, 2002).

In order to study the effectiveness of CBM, it is helpful to examine the relationship between CBM and other tests. One idea is to take specific sections of this assessment process such as mathematics or reading fluency and look at the relationship between scores on these sections of CBM and scores from the state mathematics and reading assessments. The general research question is: Are CBM scores predictive of proficiency levels on state assessments? The current study provided more information about the effectiveness of CBM as an alternative testing procedure.

Understanding the importance of the research question requires examining the issue of assessment for three specific areas: standardized assessment, CBM, and previous research studies. It is important to understand why there is a need to question the effectiveness of traditional standardized testing methods. It is also beneficial to have background knowledge of CBM as a procedure used to improve student achievement in the classroom and identify "at risk" children. Looking at similar studies can also help to

clarify the importance of using CBM scores to predict performance on traditional standardized assessments.

Review of the Literature

While standardized assessment procedures have always been the basis for academic testing in school psychology, professionals have begun to question the adequacy of these tests to measure academic performance. Particular questions have been raised about the value of using these types of tests to assess student progress and the inability of these tests to overlap with the instructional curriculum in the classroom. Another criticism of standardized assessment methods is the fact that the measures were not designed to help develop intervention strategies for students (Shapiro, 1989). While standardized testing can be useful for gaining a normative perspective on the academic abilities of students, it is not effective in providing instructional planning. The items on these kinds of tests are not directly related to the actual classroom activities (Gettinger, as cited in Shapiro & Kratochwill, 1988).

Although there are many problems and criticisms of standardized testing methods, they are still highly valued in education. The process of evaluating students with these methods will most likely continue for decades, and educators need to know how to prepare students for these kinds of tests in order to help them to succeed in school. The state departments of education and local school districts use the results of statewide assessments to make major decisions that can affect students, individual school buildings, and districts. Due to the federal legislation enactment of the No Child Left Behind Act of 2001 (NCLB; Pub.L. No.107-110), which included the most recent set of amendments to The Elementary and Secondary Education Act of 1965 (ESEA; Pub. L. No. 89-750), increased pressure is being placed upon schools to make progress on these high stakes

tests. The alternative method of assessment known as CBM can provide more direct, inexpensive, and less time consuming measures to help identify academic problems and guide interventions (Hartman & Fuller, 1997). For this particular research it is important to be knowledgeable about "high-stakes testing" as well as the Kansas State Assessment tests of reading and mathematics.

High-Stakes Testing

There is a growing public concern about the success of schools in the United States. At the forefront of the educational reform movement, continuous pressure is being placed upon school districts to improve student outcomes. The 2002 re-authorization of No Child Left Behind Act (NCLB) requires states to adopt standards for what students should know and do, and to have a way to monitor progress toward these standards (Braden, 2002). NCLB further requires states to meet "adequate yearly progress" by increasing test scores. Schools that fail to meet these requirements will be faced with a series of sanctions (National Association of School Psychologists, 2002). According to Cizek (as cited in Braden, 2002), results of accountability measures are most often reported in criterion referenced or standards referenced scales. Most states describe student performance relative to proficiency standards. High stakes testing refers to the idea that schools are being held accountable for meeting the state standards. "High stakes tests are tests from which results are used to make significant educational decisions about schools, teachers, administrators, and students" (Amrein & Berliner, 2002, p.7). These tests are also termed "high-stakes" because they can carry serious consequences for students and educators. Schools may be judged according to the average scores of their students. High scores have the possibility of bringing public praise or financial rewards, while low scores may bring public embarrassment and other sanctions. Individual

students with low scores may be held back a grade or denied a high school diploma in some states (American Educational Research Association, 2000).

High-stakes tests were developed by policy makers with the intention of improving education, however, in some cases there is the potential for serious harm to individual students, teachers, school districts, and the field of education in general. Decisions that affect students' life chances or educational opportunities should not be made solely on the performance of standardized tests (American Educational Research Association, 2002). The results of these tests should be used as indicators of the need for early intervention, evaluation of learning problems, or programmatic changes (National Association of School Psychologists, 2002). A variety of measures of academic achievement, as well as parental input, should be utilized in making these important decisions. The focus of standardized testing is to look at the knowledge and skills that students have acquired and not just their test scores. Accountability testing for low, medium, and high stakes testing is going to continue or increase, and educators need to know how to use and respond to the testing in appropriate ways. School psychologists can use their knowledge of assessment and individual differences to help parents, teachers, and administrators to understand accountability mandates and the consequences of high-stakes testing (Braden, 2002).

Kansas State Assessments

Statewide assessment tests are considered to be high stakes tests because of the increased amount of accountability placed on schools to perform at satisfactory levels. The Kansas State Mathematics, Reading, and Writing Assessments are given to all students in specific grades. The Kansas Mathematics Assessment is given in Grades 4, 7, and 10, and the Kansas Reading Assessment is given in Grades 5, 8, and 11. The Kansas

State Writing Assessment will not be discussed since this research relates to mathematics and reading. The Kansas State Assessments serve the following three main purposes:

1. Provide aggregate state accountability and progress information toward meeting the Kansas Curriculum Standards in the tested areas

2. Provide building and district information to support school improvement evaluation needs as appropriate

3. Report on the performance of students to support instructional planning for individuals and groups as judged appropriate by local educators (Glasnapp, Poggio, & Omar, 2000, p. 2)

Kansas has adopted statewide curricular standards for math and reading which were developed by trained professional teachers. These standards have been aligned with the Kansas Mathematics and Reading Assessments, which are given to all students. The Kansas state assessments are based upon these curriculum standards that were developed specifically for Kansas. The fact that these assessments are based on the Kansas curricular standards sets them apart from other standardized tests such as the Iowa Test of Basic Skills that were not standardized specifically for the Kansas curriculum standards. The Kansas curriculum standards are targeted at higher order outcomes including critical thinking skills, diverse communication skills, problem solving, reasoning, and decision-making skills (Glasnapp et al., 2000).

Using the Kansas Curriculum Standards as a guide, the Kansas state assessments were developed by Kansas educators nominated by their school districts and state professional association leaders. The state assessments were a product of Kansas educators whose development is coordinated by the Center for Educational Testing and Evaluation (CETE) at the University of Kansas and the Kansas State Department of

Education (KSDE; Glasnapp et al., 2000). A series of steps were followed leading to the creation of the assessments. Before the actual test item development began, CETE and the KSDE agreed on the general structure and format for each assessment by content and grade level. Once agreement between these two groups was reached on specifications for an examination (number of items, amount of time allowed for testing, format/layout, and structure and coverage of the questions), the actual questions were developed by a team of four to six experienced, highly regarded Kansas teachers at the grade level for the content area. Teachers were selected based on nominations from local school districts. Teachers were trained on item writing techniques and rules for test construction. The teachers used the Kansas curriculum standards as their sole guide. After the assessments were designed, they went through a series of reviews and editing steps made by the CETE and curriculum standard specialists for each content area and grade level (Glasnapp et al., 2000).

According to the Individuals with Disabilities Education Act 1997, all children with disabilities are included in the state assessment programs with accommodations provided as needed. The assessment participation options in Kansas include general assessment, general assessment with accommodations, modified assessment, and the alternate assessment (Kansas State Department of Education, 2002a). The states are required to report both aggregate scores and student scores by using performance levels. In Kansas, these performance levels include advanced, proficient, satisfactory, basic, and unsatisfactory. At least a certain percentage of students are required in the advanced proficiency level and no more than a certain percentage is allowed in the unsatisfactory level. The idea is to compare the "expected" percentage of students in the middle three levels with the "actual" percentage of students in the same three levels (Kansas State

Department of Education, 2001). According to the Kansas State Department of Education (2002a), in terms of overall achievement, generally over 60% of students performed in the top three proficiency levels in all subject areas.

In recent years the Kansas assessments have been called upon to provide information to contribute to ongoing school accreditation procedures. Results from the reading and mathematics assessments are used to help monitor annual school progress and support Title I monitoring and evaluation requirements. Based on performance, a school may be identified as having achieved the state's "Standard of Excellence." Student classification and school decision points have been decided using typical standard setting approaches. Final cut points are established by the Kansas Department of Education based on reviews of actual score distributions (Glasnapp et al., 2000). The cut scores for each of the state assessment tests are determined based on information from teacher ratings of student performance, student performance on the state assessments tests, and expert judgments of teachers, principals, and curriculum directors. The cut scores are identified to define rules for the classification of students into one of the five performance levels defined by the state. The performance level category cut scores are based on the total percent correct scores.

Using the 5th grade reading assessment as an example, students who have total scores of 93 percent correct or above are considered to be in the "advanced" category. Students with 87-92 % correct scores are considered to be in the "proficient" category. Students with 80-86 % correct scores are in the "satisfactory" category and students with 68-79 % correct scores are in the "basic" category. Lastly, students with a percent correct total score below 68 are considered to be in the "unsatisfactory" category (Glasnapp et al., 2000).

For the 4th grade math assessment, students who have total percent correct scores of 75 percent or above are considered to be in the "advanced" category. Students with percent correct scores of 60-74 are considered to be in the "proficient" category. Students with percent correct scores of 48-59 are considered to be in the "satisfactory" category and students with percent correct scores of 35-47 are considered to be in the "basic" category. Lastly, students with a percent correct total score below 35 are considered to be in the "unsatisfactory" category. The cut scores for tests across grade levels within a content area are the same except for the cut scores needed to be in the "advanced" category in the Grade 10 mathematics assessment. The "advanced" category cut score at Grade 10 is 70 % correct (Glasnapp et al., 2000).

When looking specifically at the Kansas State Mathematics assessment, the test is scored based on two different skills. The first is the knowledge process skill score. The knowledge skill requires students to know and be able to do a set of mathematical concepts, facts, and procedures. The second skill is the application skill score. The application skill requires students to describe how mathematical knowledge can be used and applied in the real world. This skill requires higher level processing skills. The total percent score is determined by adding all knowledge and application items and expressing this score as a percentage of the total number of items combined. According to the year 2000 state mathematics assessment data, the percent correct scores show evidence of high reliability for the intended purpose of the testing. Using the Kuder-Richardson Formula 20 (K-R20), reliability coefficients for the 4th grade mathematics total scores ranged from .86 - .87 (Glasnapp et al., 2000).

When looking exclusively at the Kansas mathematics assessment, there are some important characteristics to note. According to the 2002 Kansas State Mathematics data,

boys continued to perform slightly better than girls. Whites and Asians continued to perform at the highest level, while Blacks and Hispanics performed the least well. Lower socio-economic status (SES) groups were defined as students who were eligible for free or reduced lunches. These students continued to score lower than those not eligible for free or reduced lunches.

The 2002 Kansas State Reading Assessment data indicated girls performed slightly better than boys. Whites and Asians performed at the highest levels, while Blacks and Hispanics performed the least well. However, Hispanics have made increases in the percentages of students in the top three performance levels when compared with the 2001 results. The lower SES groups of students also scored lower on the reading assessment than students of higher SES status (Kansas State Department of Education, 2002b). The Kuder Richardson Formula 20 was used to determine 5th grade reading reliability. The reliability coefficients for Total scores ranged from .91 - .92. Coefficient Alpha yielded reliability coefficients for Total Reading scores from .90 - .92

Another important concept related to technical adequacy is content validity. Content validity is the degree to which an instrument logically appears to measure the intended variable. This type of validity is determined by expert judgment. During test construction, test items are analyzed to minimize bias towards any gender or ethnic group. This process of reviewing test items is known as Differential Item Functioning Analysis (DIF). This process was used with the Kansas State Assessments using two different approaches. The first approach was the review of test items by a panel of expert persons representing impacted groups. The panel of experts conducted logical review of items and looked for bias, insensitivity, and offensiveness. The majority of individuals on the panel were members of the Kansas state trained Equity Council. The second

approach was an empirical examination of item responses. All reading and math items were subjected to a review prior to the production of test forms and booklets (Glasnapp et al., 2000).

When taken by students, test items can function differently for gender and ethnic groups. A reason for the DIF bias is that often items can favor a specific group due to gender or cultural experiences. Other reasons why a test item may have differential functioning could be due to curriculum differences or multidimensional test construction. According to Glasnapp et al. (2000) at the Center for Educational Testing and Evaluation, there is consistency in identifying DIF items on the Kansas State Assessments in mathematics and reading; however, it is not 100 % agreement. A few items in any analysis might identify DIF items in one sample whereas in another sample they may not be identified. This concept is important in Kansas because ethnic groups identified in the DIF analysis are congregated in certain school districts and then are compared to random samples of white students across the state (Glasnapp et al., 2000). Further details regarding the technical characteristics of the Kansas Assessments in Mathematics and Reading are available in the extensive technical manual (Glasnapp et al., 2000).

Both the reading and mathematics state assessment tests provide options for modified assessments. An assessment modification is "a change in the procedure for assessment which will in some way change what is being measured" (McVey & Wright, 2003). The modified assessments are available in mathematics, reading, writing, science, and social studies. The eligibility criteria for the modified assessments include the following: any student with an IEP or 504 plan, evidence to support that the student is functioning at or below the 4th percentile on any standardized test, and if a writing or

social studies assessment is not available, then a reading assessment can be used. Whether a student will take a modified assessment is determined by his or her IEP team.

The modified mathematics assessment format includes multiple choice and single correct responses. Time guidelines are also given which include four 45 minute assessment periods (McVey & Wright, 2003). Modified indicators are a statement of knowledge or skills that a student demonstrates in order to meet the benchmark. A benchmark is a specific statement of what a student should know and be able to do at specific times during his/her schooling. For example, modifications in math content, application skills, or a combination of both, were made for 4th grade general education indicators. These modified indicators are the basis for the assessment with modifications. The modified form for the mathematics assessment is available to students in Grades 4, 7, and 10. The students who are eligible to receive these modifications take the modified mathematics assessment.

Students may take one of the two modified reading assessments available. The first is half written, half oral and has a reading level between Grade 1 and Grade 4. It includes a written section which is the same as the general assessments and an oral section which is composed of phonics and decoding. The second is the pre-reading assessment, which is composed of pre-reading skills and includes all oral responses. There are nine pre-reading subtests and one comprehension passage subtest. The modified reading assessment is available for students in Grades 5, 8, and 11 (McVey & Wright, 2003). Students eligible for the modified reading assessment may take either the modified reading assessment, the pre-reading assessment, or the alternate assessment (Kansas State Department of Education, 2001).

Other students who have goals that are vastly different from the majority of the student population may participate in the alternate assessment. The alternate assessment is used for the students with mild/moderate disabilities who have different goals than the general population. The alternate assessment eligibility criteria include students who have an active IEP and are receiving services under the Individuals with Disabilities Education Act (IDEA). In order to be eligible for the alternate assessment

"The students demonstrated cognitive abilities and adaptive behavior require substantial adjustments to the general curriculum. The students' learning objectives and expected outcomes focus on functional application, as illustrated in the benchmarks, indicators, and examples in the extended standards" (McVey & Wright, 2003, p.6).

Students taking the alternate assessment do not take any general state assessments, assessments with accommodations, or assessments with modifications. The student must also have scored at or below the 4th percentile on a nationally or locally normed assessment. The alternate assessment is given at age levels instead of grade levels. It is given at 10, 13, and 16 years of age and is one assessment that covers that age and ability ranges of students who meet the eligibility criteria. A student's IEP team determines whether he or she will take the alternate assessment and selects 15 target indicators from the extended curricular standards on which the child will be assessed. The target indicators are composed of emerging skills which are skills at the student's current instructional level, and maintenance skills which are skills that have been demonstrated and are continually practiced to retain performance. The IEP team must choose at least nine emerging skills (McVey & Wright, 2003).

Curriculum-Based Measurement

Curriculum-based measurement (CBM) typically refers to a standardized set of procedures utilized to measure student performance in the areas of reading, mathematics, and written expression (Shinn, 1998, as cited in Howell, Kurns, & Antil, 2002). These procedures include direct observation and recording of student progress in the local curriculum in order to make informed decisions about instructional outcomes.

CBM is one of many different approaches to a curriculum-based assessment technique. CBM is a set of standardized measures that are used to monitor student progress in the basic academic areas of mathematics, reading, spelling, and written expression. According to Hintze and Shapiro (1997), "CBM uses the general education curriculum 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" (p. 351). CBM is based on the idea that how students perform on tests should indicate their level of understanding of the curriculum used in school.

There are many specific advantages that are unique to the CBM model (Jenkins, Deno, & Mirkin as cited in Shinn, 1989). These include:

1. The test items are developed from the actual curriculum.

2. The process is of short duration and can be frequently administered by teachers and other educators.

3. The process can have multiple forms.

4. It is inexpensive in terms of time and production.

5. It is sensitive to student improvement over time.

6. The technical adequacy has been well defined.

These characteristics along with many others make CBM attractive to educators and other school professionals.

According to Fewster and MacMillan (2002), "assessment becomes an integral component of instruction with the use of CBM for progress monitoring of student performance, instructional strategies, and program effectiveness" (p.155). Another advantage of CBM is that it brings together traditional behavioral and observational assessment methods and places them with a more innovative approach to measurement. The CBM model may be used to answer questions about an individual student's academic growth, but also about how a classroom can produce better academic growth for a particular student. In order to uphold and achieve these traditional methods, it is highly important for CBM measures to be scored and administered in a standardized way (Deno, Fuchs, & Marston, 2001). As schools begin to move from traditional systems of determining placement in special education programs to a more problem-solving or solution focused orientation, the use of CBM is beneficial because it can be administered efficiently and then directly linked to instruction and intervention (Howell et al., 2002). According to Good and Jefferson (as cited in Shinn, 1998), "the relevance of CBM in the Problem Solving model can be demonstrated with evidence that CBM can be used to make all of the decisions necessary to link assessment information directly to problem resolution" (p. 69).

Many studies have examined the technical adequacy of CBM. The initial validity study in reading fluency was done by Deno, Mirkin, and Chiang (as cited in Shinn, 1989). In this study, various types of reading probe measures were compared to see if they could be used to monitor student progress on a regular basis. Students were required to read aloud passages from stories in the basal readers, read aloud lists of words

randomly selected from basal readers, read aloud underlined words in basal readers,

supply words that were deleted from stories in the basal readers, and define words

selected from the basal readers. All of these measures were correlated with different

criterion tests of reading. Correlation coefficients from these measures ranged from .73 to

.91 with most coefficients above .80. Reliability coefficients were also examined for

reading using three different measures. Test-retest reliability ranged from .82 to .97 with

most above .90. Parallel forms reliability ranged from .84 to.96 with most above .90.

Finally, interrater reliability was .99. When examining all of these measures, there is

adequate evidence of the reliability of CBM reading according to Deno (as cited in Shin,

1989). Oral reading fluency of CBM can be used as

> "a 'vital sign' of reading achievement in much the same sense that heart rate or
>
> body temperature is used as a vital sign of physical health. We must immediately
>
> caution, however, that just as heart rate and body temperature do not reveal all
>
> that there is to know about physical health, the average number of words read
>
> aloud from text in 1 minute does not reveal all that can be known about the
>
> student's reading."(Deno, 1985, p. 224)

The role of CBM reading fluency benchmarks in predicting success on state and

local district standardized reading achievement in Grades 3 through 5 was examined in a

Northeastern Illinois suburban school district. This study calculated the predictive

validity of using CBM oral reading fluency in relation to state and local reading

assessments (Sibley et al., 2001). The results indicated strong predictive validity for

CBM oral reading fluency measures relative to student performance on state and local

standardized achievement tests. The results also indicated a significant positive

relationship between oral reading fluency benchmarks and local and state assessment in

Illinois. The state assessment is known as the Illinois Standards Achievement Test (ISAT). This study examined CBM oral reading fluency measures that were given each semester beginning with the spring semester for Grade 2 through the fall semester for Grade 5 and scores on the ISAT reading assessment which was given in the spring for Grades 3 and 5. Using the Pearson correlation coefficient, a significant correlation was found between CBM oral reading fluency and the ISAT scores. The correlation coefficient was .63 for second grade spring oral reading fluency and .75 for third grade fall oral reading fluency (Sibley et al., 2001). The CBM "…approach to assessment is based on the assumption that assessment must not only inform if students are learning, but also if they are learning at a rate that will allow them to attain the desired criteria on high stakes tests" (Good et al., as cited in Sibley et al., 2001, p. 3).

Another similar study was conducted to examine the relationship between student CBM oral reading fluency scores and the Washington Assessment of Student Learning (WASL) reading assessment scores (Stage & Jacobsen, 2001). The oral reading fluency low scores in September predicted WASL failure at a level of .41. The September oral reading fluency high scores predicted WASL success at a level of .91 (Stage & Jacobsen, 2001). These results support a significant relationship between CBM and standardized assessment procedures.

Welch and Dean (2004) explored the utility of CBM as a predictive tool for student outcomes on a state mandated criterion-referenced test for English Language Arts. The CBM measure was known as the Screening To Enhance Educational Performance (STEEP) protocol and the state mandated test was known as the Louisiana Educational Assessment Program for the 21st century (LEAP 21). Researchers examined the utility of the STEEP protocol as a screening instrument to predict outcomes for fourth

grade students on the LEAP 21, a high-stakes test. A correlation of .70 ($p < .01$) was found indicating a significant relationship between these two measures (Welch & Dean, 2004). The researchers in this study also wanted to investigate whether or not there was a critical point at which predictions about the outcome of high-stakes testing can be made based upon a student's oral reading fluency score on CBM. They did, in fact, find the base value for instructional level fluency for fourth grade was 70 words read correctly per minute (WRCPM). The cut-score for the LEAP 21 was "Approaching Basic" which was the minimal passing score on this measure for the year the data were collected (Welch & Dean, 2004).

Lastly, Myers (2003) at Emporia State University conducted a similar research study investigating the relationships between CBM oral reading fluency scores and test scores obtained on two achievement tests: the Qualitative Reading Inventory and the Stanford Achievement Test, 9th Edition. Results from this study indicated the CBM scores and the Qualitative Reading Inventory test scores correlated at .79 (Myers, 2003). Therefore, CBM and the Qualitative Reading Inventory had a moderately strong relationship. Results also indicated that a moderately strong relationship ($r = .67$) existed between CBM scores and scores on the Stanford Achievement Test (9th ed.). There were no significant differences in the correlations when gender, race, and income were separately analyzed. The findings of this study show that CBM oral reading fluency scores can be a good predictor of how well students perform on other reading assessment procedures.

One area of further CBM research is mathematics measures' reliability and validity. Although the mathematics fluency measures of CBM appear to be useful and valid for evaluating and assessing math performance, the technical adequacy is still

somewhat unknown (Shapiro, 1989). In a discussion of mathematics validity evidence, Shinn (1998) states "the consistently lower correlations with published math tests have led some to question the validity of the criterion measures" (p. 66). Marston (as cited in Shinn, 1998) found the best correlations with published math tests appear to be a combination of both math and reading performance. When test re-test and parallel forms estimates were examined, single administrations of mathematics probes are reliable. Inter-scorer agreement is also high with mathematics probes (Shinn, 1989). In a summary of reliability studies of curriculum-based mathematics measures by Fuchs, Fuchs, and Hamlett, (as cited in Shinn, 1989), internal consistency correlation was .93 and interscorer agreement correlation was .98. Tindal, Germann, Marston, and Deno (as cited in Shinn, 1989) reported the test-retest (one week) reliability correlation was .93.

Standardized assessment procedures, including state assessments have been the foundation for educational testing done by psychologists to identify students for special education services. While traditional methods of assessments are good for screening and program decisions, they are not as good at dealing with the more direct and individualized needs of students. It is important for educators to examine alternative methods of assessment such as CBM as a way to determine the instructional needs of students based upon their ongoing performance in the classroom setting (Shapiro & Kratochwill, 1988).

Research Questions

Based on review of the research, the following research questions were developed to determine the utility of CBM scores to predict performance on the state assessments. More specifically:

Research Question 1: Are fourth grade Spring 2002 CBM reading fluency scores predictive of fifth grade Spring 2003 Kansas State Assessment reading proficiency levels?

Research Question 2: Do the predictive validity coefficients between the fourth grade Spring 2002 CBM reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for girls in contrast to boys?

Research Question 3: Do the predictive validity coefficients between fourth grade Spring 2002 CBM reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for white students in contrast to non-white students?

Research Question 4: Do the predictive validity coefficients between the fourth grade Spring 2002 reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for students from low income families in contrast to students from high income families?

Research Question 5: Are third grade Spring 2002 CBM mathematics fluency scores predictive of fourth grade Spring 2003 Kansas State Assessment mathematics proficiency levels?

Research Question 6: Do the predictive validity coefficients between the third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003 Kansas State Assessment proficiency levels for mathematics differ for girls in contrast to boys?

Research Question 7: Do the predictive validity coefficients between the third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003

Kansas State Assessment proficiency levels for mathematics differ for white students in contrast to non-white students?

Research Question 8: Do the predictive validity coefficients between the third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003 Kansas State Assessment proficiency levels for mathematics differ for students from low income families in contrast to students from high income families?

CHAPTER 2

METHOD

Participants

The participants in this study consisted of fourth grade and fifth grade elementary students from eight different schools in one midwestern school district. The participants included 236 fourth grade students (99 boys, 137 girls) and 163 fifth grade students (91 boys, 72 girls). All of these schools participated in the CBM norming process and the data were pre-existing. This sample was stratified based on ethnicity and socio-economic status level depending on the meeting of specific criteria for free or reduced lunches. The fourth grade sample consisted of 29% of students who qualified for free/reduced lunch and 71% of students who did not qualify. The fifth grade sample consisted of 26% of students who qualified for free/reduced lunch and 74% of students who did not qualify. The ethnicity of the fourth and fifth grade samples combined consisted of 80% White and 20% non-White.

Instruments

Curriculum-based measurement. Since the ability to report on the technical adequacy of CBM varies depending on the location of implementation, local norms were established in the Midwestern school district used in this study. In this particular school district, CBM probes were administered twice a year, once in the fall and once in the spring.

CBM Reading. The CBM oral reading fluency probes were developed by a school psychologist in the district who chose three reading passages taken from the curriculum. This particular district used a guided reading program based on children's literature books. All of the books were assigned a difficulty level with a corresponding letter of the

alphabet. All grade levels were assigned certain levels of difficulty that they were expected to cover within the school year. The selections for the reading probes were chosen from material within these books that were available to all the schools in the district. The idea was to have alignment between the curriculum and assessment. The student was given three passages and was asked to read aloud from each passage for one minute. The CBM reading procedures were defined as the number of words read correctly per one minute intervals The median number of words read correctly from each of the passages was calculated and recorded on the outside of the CBM packet.

CBM Mathematics. The CBM mathematics fluency probes used in this school district were generic probes that were purchased from the University of Oregon and used in the CBM norming process. The probes were composed of addition, subtraction, multiplication and division problems. The district made sure the level of the mathematics probe matched with the grade level at which skills were introduced in the curriculum. The CBM mathematics score is defined as the number of correct digits written per two-minute intervals. Students were given two minutes to complete as many math problems as possible.

Kansas State Assessments

The scores for the Kansas State Assessment are reported in both proficiency levels and weighted scores. The proficiency levels for the Kansas State Assessments include the following: advanced, proficient, satisfactory, basic, and unsatisfactory. The weighted scores are officially titled "mean percent correct weighted scores" (T. McEwen, personal communication, April 9, 2004). The weighting of the scores comes from a complex formula that is applied to each measure each year (T. McEwen, personal communication, April 9, 2004).

Kansas Reading Assessment. The fifth grade Kansas Reading Assessment consists of four authentic extended reading selections representing different text types in each test form. These text types include; narrative, expository, technical, and persuasive selections. The format of the assessment is multiple mark yes/no questions, and true/false questions (Kansas Department of Education, 2002b). Questions are posed and several choices are presented as alternatives to the question. The student responds yes or no to each question to indicate whether the option is correct or incorrect.

Kansas Mathematics Assessment. The fourth grade Kansas Mathematics Assessment follows a multiple choice, selected response testing format and is based upon indicators found within the state standards in mathematics. The multiple choice questions have only one correct answer to be selected from the response options provided to the questions. Each item on the test is linked to the standard in the indicator.

Data Collection

The CBM mixed mathematics probes were administered to third graders in Spring 2002, and CBM reading probes were administered to fourth graders in Spring 2002. The CBM data were collected by trained school psychologists and special education teachers. The CBM assessment was administered using standardized procedures during the spring of the 2001-2002 school years. In the spring of the 2002-2003 school year, the state assessments in mathematics and reading fluency were administered by trained teachers and were machine scored. The administration techniques were standardized according to the state testing regulations.

Procedure

Permission was obtained from Emporia State University Institutional Review Board (IRB) to conduct the study. A special request to the large, midwestern public

school district was made to obtain permission to access and use the pre-existing data, accompanied by the IRB chair's letter, indicating the research proposal was approved. Permission was obtained to analyze and interpret data from the school district.

This study analyzed archival data that was provided by the school district. The scores obtained on the CBM measures were in the form of words read correctly per minute for reading and number of correct digits for mathematics. The Kansas Reading and Mathematics Assessment scores were reported in the form of proficiency levels and weighted scores.

CHAPTER 3

RESULTS

The purpose of this study was to examine the utility of Curriculum-Based

Measurements (CBM) as predictive tools for the proficiency levels and weighted scores

obtained on the Kansas Reading and Mathematics Assessments. The statistical

procedures used in this study analyzed CBM scores obtained in the 2001-2002 school

year to determine the predictive validity of the CBM reading and mathematics scores to

predict the Kansas State Assessments of Mathematics and Reading scores given in the

2002-2003 school year, respectively. Specifically, third grade Spring 2002 CBM

mathematics fluency scores were correlated with Kansas Mathematics Assessment scores

obtained from the same students who were in fourth grade in the Spring 2003 semester.

Fourth grade Spring 2002 CBM reading fluency scores were correlated with Kansas

Reading Assessment scores obtained from the same students who were in the fifth grade

in the Spring 2003 semester. The scores obtained from the CBM mathematics probes

were reported in the form of digits correct and the CBM reading probes were reported in

the form of number of words read correctly. The Kansas Mathematics and Reading

Assessment were available in both categorical proficiency levels and weighted scores.

Both proficiency levels and the weighted scores on the Kansas State Assessments were

analyzed for the initial analysis. Hence, the CBM scores were the predictor variables and

the scores on the state assessments in mathematics and reading were the criterion

variables.

The statistical technique used to answer the research questions was the Pearson

Product-Moment correlation coefficient. An alpha level of .05 was used to determine

statistical significance. Further statistical analyses were conducted using separate Z tests

to investigate potential differences in these correlations between the following groups: gender, race, and socio-economic status. Gender was defined and measured by information on individual students' school record forms. Ethnicity was defined and measured by reports from parents, which were recorded on school record forms. Socio-economic status level was defined and measured by students receiving free/reduced lunches. Students receiving free/reduced lunches were classified as low income level and students not receiving free/reduced lunches were classified as high income level.

Research Question 1 specifically asked: Are fourth grade Spring 2002 CBM reading fluency scores predictive of fifth grade Spring 2003 Kansas Reading Assessment proficiency levels? The correlation value for CBM reading fluency scores and the Kansas Reading Assessment proficiency levels was .60 ($p < .01$). Hence, the answer to Research Question 1 is yes; CBM and the Kansas Reading Assessment have a moderately strong relationship. The r values for CBM reading and Kansas Reading Assessment weighted scores are given in Table 1. The remaining tables also include data based upon the weighted scores. Since all results for both sets of scores, proficiency levels and weighted scores were statistically the same, only the results using the proficiency level are discussed in the following paragraphs.

Research Question 2 stated: Do the predictive validity coefficients between the fourth grade Spring 2002 CBM reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for girls in contrast to boys? For the CBM reading and the Kansas Reading Assessment boys' scores correlated at .65 ($p < .01$) and girls' scores correlated at .55 ($p < .01$) (see Table 1). There was no statistical difference between these correlations, $Z = 1.11, p > .05$ which indicates the boys' and girls' correlations are equivalent to each other (see Table 2).

Table 1

Correlations of Curriculum-Based Measurement Spring 2002 Probes with Kansas State Spring 2003 Reading Assessment Proficiency Levels and Weighted Scores

Group	n	Kansas Reading Assessment	
		Proficiency Levels	Weighted Scores
Total	163	.60**	.67**
Boys	91	.65**	.73**
Girls	72	.55**	.61**
White	131	.63**	.71**
Non-White	32	.59**	.64**
Low Income	43	.63**	.58**
High Income	120	.54**	.68**

* $p < .05$

**$p < .01$

Table 2

Z Values Comparing Correlations of Curriculum-Based Measurement with the Kansas State Reading Assessment.

Group	Proficiency Levels	Weighted Scores
Boys/Girls	1.11	1.39
White/Non-White	.31	.63
Low Income/High Income	.75	.91

Research Question 3 examined the following: Do the predictive validity coefficients between fourth grade Spring 2002 CBM reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for white students in contrast to non-White students? As noted on Table 1, Whites' scores on these two measures correlated at .63 ($p < .01$) and non-Whites' scores correlated at .59 ($p < .01$). There was no statistical difference between these correlations, $Z = .31, p > .05$, thus indicating White and non-White correlations are equivalent to each other (see Table 2).

Research Question 4 specifically asked: Do the predictive validity coefficients between the fourth grade Spring 2002 reading fluency scores and fifth grade Spring 2003 Kansas State Assessment proficiency levels for reading differ for students from low income families in contrast to students from high income families? For the CBM reading and the Kansas Reading Assessment, the correlation for low income students was .63 ($p < .01$) and the correlation for high income students correlated at .54 ($p < .01$) (see Table 1). There was no statistically significant difference between these correlations, $Z = .75, p > .05$, thus indicating the correlations for students from low income families and students from high income families were equivalent to each other (see Table 2).

Research Question 5 specifically stated: Are third grade Spring 2002 CBM mathematics fluency scores predictive of fourth grade Spring 2003 Kansas State Assessment mathematics proficiency levels? The correlation value for CBM mathematics fluency scores and the Kansas Mathematics Assessment proficiency levels was .37 ($p < .01$), thus indicating a modest correlation (see Table 3).

Research Question 6 stated: Do the predictive validity coefficients between third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003

Table 3

Correlations of Curriculum-Based Measurement Spring 2002 Probes with Kansas State

Spring 2003 Mathematics Assessment Proficiency Levels and Weighted Scores

Group	n	Kansas Mathematics Assessment	
		Proficiency Levels	Weighted Scores
Total	236	.37**	.38**
Boys	99	.33**	.32**
Girls	137	.40**	.42**
White	188	.41**	.40**
Non-White	48	.24	.28
Low Income	68	.28*	.26*
High Income	168	.39**	.41**

*$p < .05$

**$p < .01$

Kansas State Assessment proficiency levels for mathematics differ for girls in contrast to boys? For the CBM mathematics and the Kansas Mathematics Assessment boys correlated at .33 ($p < .01$) and girls' scores correlated at .40 ($p < .01$) as indicated in Table 3. There was no difference between these correlations, $Z = .62$, $p > .05$, which indicates the boys' and girls' correlations are equivalent to each other (see Table 4).

Research Question 7 examined the following: Do the predictive validity coefficients between the third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003 Kansas State Assessment proficiency levels for mathematics differ for White students in contrast to non-White students? As noted on Table 3 Whites' scores on these two measures correlated at .41 ($p < .01$) and non-Whites' scores correlated at .24 ($p > .05$). There was no statistical difference between these correlations, $Z = 1.16$, $p > .05$, thus indicating the predictive validity coefficients for the White and non-White students were equivalent to each other. This Z value is presented in Table 4.

Research Question 8 specifically asked: Do the predictive validity coefficients between the third grade Spring 2002 CBM mathematics fluency scores and fourth grade Spring 2003 Kansas State Assessment proficiency levels for mathematics differ for students from low income families in contrast to students from high income families? For the CBM mathematics and Kansas Mathematics Assessment, the correlation for students from low income families was .28 ($p < .05$) and for students from high income families the correlation was .39 ($p < .01$) (see Table 3. There were no differences between these correlations, $Z = .86$, $p > .05$, thus indicating that the predictive validity coefficients for these two group were equivalent to each other (see Table 4).

Tables 5 and 6 provide the means and standard deviations for all obtained scores. These scores were not statistically analyzed.

Table 4

Z Values Comparing Correlations of Curriculum-Based Measurement with the Kansas State Mathematics Assessment

Group	Proficiency Levels	Weighted Scores
Boys/Girls	.62	.89
White/Non-White	1.16	.83
Low Income/High Income	.86	1.17

Table 5

Means and Standard Deviations of CBM Reading Scores and Kansas Reading

Assessment Proficiency Levels by Gender, Race, and Income

	CBM Reading			Kansas Reading Assessment		
	n	*M*	*SD*	*n*	*M*	*SD*
Gender						
Boys	91	106.41	41.31	91	3.53	1.23
Girls	72	102.78	40.67	72	3.33	1.37
Race						
White	131	104.39	42.35	131	3.56	1.27
Non-White	32	106.50	35.46	32	2.97	1.31
Income						
Low	43	87.74	38.26	43	2.63	1.31
High	120	110.92	40.28	120	3.73	1.16
Total Sample	163	104.80	40.94	163	3.44	1.30

Table 6

Means and Standard Deviations of CBM Mathematics Scores and Kansas Mathematics

Assessment Proficiency Levels by Gender, Race, and Income

	CBM Mathematics			Kansas Mathematics Assessment		
	n	*M*	*SD*	*n*	*M*	*SD*
Gender						
Boys	99	19.25	8.40	99	3.37	1.24
Girls	137	19.26	7.41	137	3.28	1.22
Race						
White	188	19.52	7.59	188	3.53	1.17
Non-White	48	18.23	8.67	48	2.48	1.11
Income						
Low	68	18.16	8.06	68	2.87	1.23
High	168	19.70	7.71	168	3.50	1.18
Total Sample	236	19.26	7.82	236	3.32	1.23

In summary, no correlations between CBM and the Kansas State Assessments were significant for contrasting types of student groups. Both CBM reading and mathematics were equally predictive of the Kansas State Assessments across types within each of the three demographic categories.

CHAPTER 4

DISCUSSION

Several studies have been done to examine whether Curriculum-based

measurement can be used to accurately predict performance on high-stakes testing as well

as other assessment measures. The results of these studies have been positive, indicating

that CBM is a useful method of alternative assessment and can be used to help identify

students at risk of failing these state-mandated tests. The following paragraphs will relate

this previous research to the current study.

Reading

Research Question 1 asked whether CBM reading fluency scores are predictive of

the Kansas Reading Assessment proficiency levels. The correlation obtained indicated a

moderately strong relationship between performance on CBM reading probes and the

Kansas Reading Assessment. These results support previous research by Stage and

Jacobsen (2001) who concluded CBM oral reading fluency could be used to establish

statistically reliable cut scores. Fourth grade CBM cut scores were used to identify

students in danger of failing the state-mandated Washington Assessment of Student

Learning (WASL) reading assessment given in fourth grade.

The results of the present study are also similar to results obtained by Welch and

Dean (2004) who examined the concurrent validity of CBM oral reading fluency scores

to predict English Language Arts scores on the Louisiana Educational Assessment

Program for the 21st Century (LEAP 21) for fourth grade students. They found a

statistically significant relationship between students' scores on the STEEP CBM reading

measure and students' English Language Arts scores as measured by the LEAP 21

assessment.

In another similar study regarding the utility of CBM reading scores to predict state reading assessment scores, Sibley et al. (2001) found a strong relationship between CBM oral reading fluency and scores on the Illinois Standards Achievement Test (ISAT). Students who had the ability to read at or above 90 words correct per minute in the spring had a higher probability of falling into the "meet or exceeds" category on the Reading Test portion of the ISAT the following school year.

Myers (2003) conducted a similar study examining the utility of CBM reading as a predictive tool for students' performance on both the Qualitative Reading Inventory and the Stanford Achievement Test (9th ed.). Her results revealed moderately strong relationships existed between CBM oral reading fluency scores and both of the standardized, norm-referenced measures, thus indicating the predictive power of CBM oral reading fluency relative to other assessment measures.

Hence, like similar studies from three different states, the current findings support the usefulness of CBM reading assessment to predict performance on Kansas State Reading Assessment, which is aligned with the Kansas state curricular standards. Likewise, based upon Myers (2003) one could conclude CBM reading scores are predictive of performance on standardized norm-referenced tests.

The demographic information in the current study depicting differences between gender, race, and income was investigated for a number of different reasons. One reason for this examination was due to the fact that the previous research studies that were reviewed had not examined these differences and the researcher was curious if differences could be detected. If possible significant differences were detected between these groups; it would provide useful information to educators as to what types of alternative assessment should be given. Another example of why this demographic

information was examined is because if correlations between CBM and the Kansas State Assessments were not significantly different for contrasting groups, then educators could be assured CBM and the state assessments are equally predictive across gender, race, and income. Therefore, the tests are not biased.

Research Questions 2, 3, and 4 queried about the predictive validity of CBM reading scores with the Kansas Reading Assessment proficiency levels in relation to three demographic variables: gender, race, and income level. When looking at individual demographic categories such as gender, on the CBM reading and Kansas Reading Assessment, the correlations were relatively close. Hence, for these individual categories CBM reading predicted performance on the Kansas State Reading Assessment equally well. Although not analyzed statistically, boys did perform slightly better than girls. Considering that the CBM reading is an individually administered test, one reason for the discrepancy in scores could be due to the fact that some of these girls may have characteristics of being shy and reserved.

When different races were compared, the correlations between CBM reading and the Kansas Reading Assessment were moderately strong and very close for both white students and non-white students. The sample sizes were vastly different being 131 and 32 respectively. This similarity between the two groups could be due to the geographic location in which the data were obtained.

Lastly, when examining the differences in income level as reported by comparing scores of students who received free/reduced lunches with students who did not receive free/reduced lunches, there were similar correlations for both groups on CBM reading and the Kansas Reading Assessment. Both correlations were in the moderately strong category. There were no significant differences between these groups of students.

The statistical analysis completed on all of these demographic categories found no significance for any of the groups. This indicates boys and girls, whites and non-whites, and students from lower socio-economic backgrounds and higher socio-economic backgrounds were all equivalent.

These results support Myers' (2003) findings. She examined the relationship between CBM oral reading fluency and performance on the Qualitative Reading Inventory and the Stanford Achievement Test (9th ed.). The current study investigated these differences in gender, race, and income level in a similar study and also found no significant differences between groups.

Research Question 5 asked whether CBM mathematics fluency scores are predictive of Kansas Mathematics Assessment proficiency levels. The correlations obtained for CBM mathematics probes and the Kansas Mathematics Assessment indicated a modest relationship. The overall total correlation between CBM mathematics and the Kansas Mathematics Assessment was a .37 which is considered to be a modest correlation at the $p < .01$ level. Given the fact that a minimum amount of information exists regarding the technical adequacy of CBM mathematics measures, this correlation was expected. According to Thurber, Shinn, and Smolkowski (2002), CBM mathematics validity studies in relation to other commercial norm-referenced mathematics tests, few reported correlations exceeding .60. Two hypotheses have been suggested to explain these lower correlations:

First, the limited content validity of the criterion commercial mathematics tests may make them inadequate criterion measures. Second, these criterion math tests could be measuring more than just mathematics skills because many of the items rely on silent reading of the instructions and problems. Thus reading skills may

influence performance on the mathematics test. (Marston, as cited in

Thurber et al., 2002, pp. 499-500)

More research into the technical adequacy of CBM math measures would benefit the

overall effectiveness of CBM as a method of alternative assessment.

Mathematics

Research Questions 6, 7, and 8 questioned the predictive validity of CBM

mathematics fluency scores for the Kansas Mathematics Assessment proficiency levels

based upon demographic differences among students including gender, race, and income

level. When gender differences in the utility of CBM mathematics to predict Kansas

Mathematics Assessment were examined, the correlations were again relatively close.

Although, girls performed slightly better than boys, the difference was not tested

significantly. Although correlations between CBM mathematics and the Kansas

Mathematics Assessment were modest for White students at .41 and weak for non-White

students at .24, there was no statistical difference. Note the sample sizes for this group

were vastly different, 188 and 48, respectively. The correlations for the mathematics data

were significantly lower than the correlations for the reading data, which were to be

expected based upon the previous research. CBM mathematics only requires students to

complete math calculations, whereas the Kansas Mathematics Assessments requires

students to be able to read in order to complete the math problems. Research has been

done on the effectiveness of CBM reading as a predictor of performance on standardized

reading assessments. There are no previous ethnicity studies which have investigated the

relationship between CBM mathematics and the state mathematics assessments.

Lastly, when looking at the differences in income level on the CBM

mathematics and the Kansas mathematics assessment based upon receipt of free/reduced

lunches and not being eligible for free/reduced lunches, low income students correlated at .28 and high income students correlated at .39. As noted previously, these weak correlations were not significantly different. No previous research studies have investigated the CBM mathematics and standardized mathematics tests, including state assessments and norm-referenced tests.

In summary, on all of these demographic categories, no significant differences were found for any of the groups. This information indicated boys and girls, Whites and non-Whites, and students from low income families and students from high income families were all equivalent. In part, this supports Myers' (2003) results that found no significant differences for the same demographic categories. However, she compared CBM reading data to other reading achievement measures.

The rationale for conducting the current study was based upon previous research in the area of alternative assessment measures. Relatively little research has been done in this area and the results of the studies that have been done are extremely positive. The fact that past studies have been so positive, led the researcher to inquire whether the same results could be obtained on a smaller sample of students in a particular area of the state. Lastly, the idea that an alternative assessment procedure such as CBM can be used to identify students who are at risk for failing high stakes tests and provide them with academic interventions to improve outcomes, should provoke enthusiasm in educators. Not only are interventions beneficial to students, they are also beneficial to school districts in regards to the federal legislation of No Child Left Behind (NCLB) and the challenges in meeting Adequate Yearly Progress (AYP).

Correlations between CBM and Kansas State Assessments remained fairly consistent even though they were higher for reading than for math. All of the reading

correlations were moderately high. All of the mathematics correlations were modest with the exception of very weak correlations for non-whites and students from low incomes. These are two different types of assessment that measure similar concepts. CBM focuses on reading fluency and mathematics calculation fluency and the Kansas State Assessments focus on comprehension and vocabulary for the reading assessment and problem solving skills for the mathematics assessment.

Assumptions and Limitations

There are a variety of assumptions that exist for this current research study. Most importantly, it is assumed that both the CBM and the Kansas State Assessments were given in the same standardized fashion. It is assumed that the rules of standardization were followed for each assessment measure. It is also assumed that only trained personnel administered the assessments to students. Another assumption is that all of the scores obtained were accurately reported and the correct scoring procedures were put into place. It is also assumed that students who received free/reduced lunches and those who did not were appropriately eligible.

Unfortunately, limitations always exist in the area of scientific research. One limitation of this study is the loss of students over the 12-month period of time. The loss could be due to a number of different reasons. One reason is due to student movement within and out of the district. Although this did occur, it was not biased based upon demographic characteristics. The state assessment data were only available for students who had received mathematics and reading assessments 12 months previously in the same building. Since the data were only available for students who attended school in the same building one year later, the sample is slightly restricted. There was also some loss in the CBM norming due to kids who moved, got sick on the day of testing, or moved to

another building in the district. Lastly, another limitation is student loss due to the fact that CBM reading was not given in two of the schools because reading is individually administered and those particular schools only participated in group administered testing.

Further Research

CBM is an excellent tool that provides educators with a progress monitoring instrument to measure the effectiveness of intervention strategies used to help meet the individual needs of students. This method of assessment has been proven to help teachers identify kids who are "at risk" of doing poorly on other types of assessments. Future research studies on the effectiveness of CBM at predicting performance on other types of assessment measures is needed to further increase the reliability and validity of the CBM measures. Studies such as the current one are helpful in validating the ability of CBM to be used as a red flag indicator of students who need intervention.

Research relating to high stakes testing such as the Kansas State Assessments is extremely valuable given the increasing pressure that is placed upon school districts to achieve Adequate Yearly Progress. One area that could be developed in future research studies is to determine the CBM cut scores that could be used to identify students likely to score in the two lowest proficiency levels (basic and unsatisfactory) on the Kansas State Assessments. The benefit of determining these cut scores is to have teachers put into place interventions before students are required to take the state assessments.

Conclusion

CBM is a very valuable tool in the field of education and can serve a variety of different functions. The alternative assessment procedure known as CBM continues to gain respect among educators as a successful method of evaluating student instructional

outcomes. CBM can be used to evaluate students in order to meet their unique individual needs in the classroom and monitor progress.

Another important function of CBM is that it has the potential to be used as an intervention technique to examine a student's instructional level and to bring them up to speed with the curriculum. As school districts are faced with the ever challenging requirements of the No Child Left Behind legislation and the push to meet adequate yearly progress, Curriculum-based measurement is a very effective tool that can be used to help pinpoint students who are in danger of failing and help them to succeed both in the classroom in general and on high stakes tests. The general idea is to facilitate instructional programming and to help students early rather than waiting for them to fail. We are moving away from this "wait to fail" approach (Welch & Dean, 2004).

As seen in the current study, a significant relationship exists between CBM scores and scores on the Kansas State Assessments. Although correlations were consistent for both reading and mathematics, the correlations were higher for reading. Although more research needs to be conducted on the validity of CBM mathematics as a predictor for other assessment measures, CBM reading has the potential to predict proficiency levels on the Kansas Reading Assessment as well as other assessment measures. This research indicates that it is possible to identify students who are "at risk" of failing both in the classroom and on high-stakes tests. Early identification of these children who are considered to be "at risk" of failing is necessary to ensure that interventions are provided and monitored in order to help them succeed. The Institute for the Development of Educational Achievement at the University of Oregon focuses on "Big Ideas in Beginning Reading" with the goal of having all children reading by the end of third grade (n. d.). Dr. Edward Kam'enui, the institute director, emphasizes the concept that children

are "learning to read" in the first two to three years of school and "reading to learn" in the next following years of education. Educators need to be aware of the unique individual needs of students in order to help them succeed throughout their education. Educators can use the data provided by CBM to benefit both students and school districts.

REFERENCES

American Educational Research Association (2000) . *AERA position statement concerning high-stakes testing in PreK-12 education*. Retrieved February 18, 2004, from http://www.aera.net/about/policy/stakes.htm

Amrein, A. L., & Berliner, D. C. (2002). *High-stakes testing, uncertainty, and student learning*. Education Policy Analysis Archives, *10*,18. Retrieved October 1, 2003, from http://epaa.usu.edu/epaa/V10n18/

Braden, J. P. (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 IV* (pp. 301-317). Bethesda, MD: National Association of School Psychologists.

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

Deno, S. L., Fuchs, L. S., & Marston, D. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *The School Psychology Review, 30*, 507-524.

Fewster, S., & MacMillan, P. D. (2002). School based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial and Special Education, 23*, 149-156.

Glasnapp, D. R., Poggio, J. P., & Omar, M. H. (2000). *Technical report year 2000 Kansas assessments in mathematics, reading, and writing*. Retrieved March 10, 2004, from University of Kansas, Lawrence, The Center for Educational Testing and Evaluation Web Site:http://www.cete.ku.edu/distools/docs/technicmanuals/

techreport-2000-final.doc

Hartman, J. M., & Fuller, M. L. (1997). The development of curriculum-based measurement norms in literature based classrooms. *Journal of School Psychology, 35*, 377-401.

Hintze, J. M., & Shapiro, E. S. (1997). Curriculum-based measurement and literature-based reading: Is curriculum-based measurement meeting the needs of changing reading curricula? *Journal of School Psychology, 35*, 351-375.

Howell, K., Kurns, S., & Antil, L. (2002). Best practices in curriculum-based evaluation. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 753-772). Bethesda, MD: National Association of School Psychologists.

Institute for the Development of Educational Achievement (n. d). *Big ideas in beginning reading.* Retrieved April 21, 2004, from http://reading.uoregon.edu/index.php

Kansas State Department of Education (2002a). *Kansas state assessments in mathematics, reading, and writing.* Retrieved March 8, 2003, from http://www.ksbe.state.ks.us/assessment/assessment_features.htm

Kansas State Department of Education (2002b). *Kansas state assessment highlights.* Retrieved March 8, 2003, from http://www.ksbe.state.ks.us/assessment/assessment_highlights.htm

Kansas State Department of Education (2001). *Assessments and grade levels.* Retrieved March 8, 2003, from http://www.ksde.org/assessment/assessmentupdateseptember2001.htm

McVey, D., & Wright, L. (2003, October). *Modified and alternate state assessments.* Presentation presented at the meeting of the Kansas Association of School Psychologists, Wichita, KS.

Myers, L. (2003). An examination of the relationships between curriculum-based measures and other achievement measures. Unpublished Specialist in Education thesis, Emporia State University, Emporia, KS.

National Association of School Psychologists (2002). *NASP fact sheet- large scale assessments and high stakes decisions:facts, cautions and guidelines.* Retrieved February 3, 2004 from http://naspcenter.org/factsheets/highstakes_fs.html

Shapiro, E. S. (1989). *Academic skills problems: Direct assessment and intervention.* New York: The Guilford Press

Shapiro, E. S., & Kratochwill, T. R. (Eds.). (1988). *Behavioral assessment in schools: Conceptual foundations and practical applications.* New York: Guilford.

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

Shinn, M. R. (Ed.). (1998). *Advanced applications of curriculum-based measurement.* New York: Guilford.

Sibley, D., Biwer, D., & Hesch, A., (2001, April). *Establishing curriculum-based measurement oral reading fluency performance standards to predict success on local and state tests of reading achievement.* Paper presented at the annual meeting of the National Association of School Psychologists, Washington, DC.

Stage, S. A., & Jacobsen, M. D. (2001). Predicting student success on a state-mandated performance-based assessment using oral reading fluency. *School Psychology Review, 30,* 407-419.

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

Welch, M. D., & Dean, M. (2004, March). *Using curriculum-based measurement (CBM) to predict high-stakes testing outcomes.* Paper presented at the meeting of the National Association of School Psychologists, Dallas, TX.

Permission to Copy Page

I, _____, hereby submit this thesis to Emporia State University as partial fulfillment of the requirements for an advanced degree. I agree that the Library of the University may make it available for use in accordance with its regulations governing materials of this type. I further agree that quoting, photocopying, or other reproduction of this document is allowed for private study, scholarship (including teaching) and research purposes of a nonprofit nature. No copying which involves potential financial gain will be allowed without written permission of the author.

_____
Signature of Author

_____
Date

_____
An Examination of Curriculum-Based measurement as a Predictor of State Assessment Proficiency Levels
Title of Thesis

_____
Signature of Graduate Office Staff

_____
Date Received