

PREDICTING STUDENT SUCCESS: USING CURRICULUM BASED MEASUREMENTS
OF READING TO PREDICT STUDENT SUCCESS ON MINNESOTA'S
STATEWIDE ASSESSMENTS

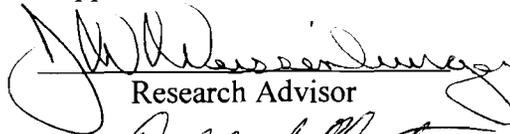
by

Kristine Louise Emery

A Research Paper
Submitted in Partial Fulfillment of the
Requirements for the
Education Specialist Degree
in

School Psychology

Approved: 6 Semester Credits


Research Advisor


Committee Member


Committee Member

The Graduate School
University of Wisconsin-Stout

November, 2005

The Graduate School
 University of Wisconsin-Stout
 Menomonie, WI 54751

ABSTRACT

	Emery	Kristine	L.
(Writer)	(Last Name)	(First)	(Initial)
Predicting Student Success: Using Curriculum Based Measurements of Reading to Predict Student Success on Minnesota's Statewide Assessments			
(Title)	School Psychology	Jacalyn Weissenburger	3/2006
(Graduate Major)	(Research Advisor)	(Month/Year)	(No. of Pages)
Publication Manual of the American Psychological Association, Fifth Edition (Name of Style Manual Used in this Study)			

Curriculum-based measurement (CBM) is an assessment tool that gives information on how a child is performing academically (i.e., reading, math, spelling, and writing). CBM can also be useful in helping teachers evaluate their programs and determine students' eligibility for special education services. This study examined the predictive validity of CBM in predicting future success on Minnesota's state assessment system, the Minnesota Comprehensive Assessments. Results indicate moderately strong correlations between students' words read correct per minute (WRCM) and their reading levels as assessed by a statewide assessment. Results suggest educators may use WRCM as a means of identifying those students who are at risk for reading failure. As such, using WRCM cut scores will assist educators to be proactive in their teaching methods as opposed to waiting until it is too late.

Acknowledgements

I would like to express my gratitude to Dr. Jacalyn Weissenburger for her patience, guidance, and expertise in completing this project. She is extremely knowledgeable in this research area and understood my vision from the start. I would also like to thank Dr. Donald Platz and Dr. Robert Sedlak for agreeing to serve as committee members when the situation wasn't always clear. In addition, I would like to thank my parents David and Beverly, for their unconditional support through the many stages of my educational career. Lastly, I would like to thank my 'future husband' Michael for his unending support and for agreeing to spend hours in various coffee shops as I typed away.

TABLE OF CONTENTS

	Page
ABSTRACT.....	ii
Chapter I: Introduction	1
<i>Statement of the Problem</i>	4
<i>Purpose of the study</i>	5
<i>Definition of Terms</i>	6
Chapter II: Literature Review	8
<i>Development of Curriculum-based Measurement</i>	8
<i>What is Curriculum-based Measurement in Reading</i>	12
<i>Using Oral Reading Rate to Predict Future Performance on State-Mandated Tests</i>	14
Chapter III: Method	18
<i>Participants</i>	18
<i>Materials</i>	18
<i>Procedure</i>	20
<i>Data Collection Procedures</i>	20
<i>Data Analysis</i>	21
Chapter IV: Results.....	23
<i>Preliminary Analyses</i>	23
<i>Research Question One</i>	23
<i>Research Question Two</i>	24
Chapter V: Discussion	26
<i>Research Question One</i>	26

Research Question Two27

Limitations28

Implications for Practice29

Suggestions for Future Research30

Summary30

References32

Appendix A: Sample of Reading Probe35

Appendix B: Tables37

Table 138

Table 239

Table 340

CHAPTER 1

Introduction

Traditionally, school psychologists predominately have worked with students in the area of special education; and, more specifically, to determine their eligibility for services. Today, that is changing. School psychologists still participate in the eligibility decision, but they also are asked to aid teachers in preparing students for state-mandated achievement tests. This change in function is due primarily to the new act referred to as No Child Left Behind (No Child Left Behind, n.d.) implemented by President Bush in 2002. Due to this act, more importance is placed on norm-referenced statewide achievement tests to measure student performance. However, these tests are given on an infrequent basis. As such, they are not ideal for teachers to monitor progress towards academic benchmarks (Crawford, Tindal, & Stieber, 2001). Because of the importance of these tests in measuring student performance, it is imperative that methods be developed to help monitor a student's ongoing progress. If educators wait to see results from statewide tests to decide if their instruction is working, it is too late. Therefore, it is essential that a system be developed that could aid educators in monitoring their instruction (McGlinchey & Hixson, 2004).

In the field of education, one system of measurement that is directly related to a student's progress in the general education curriculum is referred to as curriculum-based measurement (CBM). Curriculum-based measurement is a standardized measurement system used for recording student performance in the basic academic skills of reading, math, spelling, and written expression (Deno, 1985; Fuchs & Fuchs, 1991). CBM employs a set of standardized procedures to directly monitor students' academic performance over time.

CBM is an alternative approach used to assess student progress. Research on this system began over twenty years ago at the Institute for Research on Learning Disabilities at the

University of Minnesota. Stan Deno, one of the leaders in the development of this system, created a measurement procedure that routinely monitored student achievement in the curriculum. CBM was designed with four specific characteristics in mind. According to Deno (1985), CBM had to be:

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 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.*

The main intent of CBM was to implement an assessment system that allowed special education teachers to monitor the effectiveness of their instruction through assessing what their students had achieved. This system also allowed teachers to decipher if a student was not progressing in a specific academic area (Deno, 1992). Systematic, ongoing monitoring of a student's progress is deemed an important feature of good practice for an educator (Fuchs & Fuchs, 1991). Overall, the purpose of CBM was to decrease the separation between measurement and instruction, and to help make student achievement more integral to daily teaching and decision-making.

Because this approach easily fit both the school's curriculum and a teacher's daily instruction, it became one of the most popular assessment approaches during the 1980s for

special educators. CBMs can be used to create a database for each student, which allows the teacher to assess the success of the student's educational program (Deno, 1992). Through CBM, if the teacher found that a student was not attaining the basic academic skills proficiently, the teacher could then modify his or her teaching methods to help the student succeed (Elliott & Fuchs, 1997).

Studies have shown that when teachers use CBM information to assess the effectiveness of their student programs, not only does the instructional programming improve, but also the student's achievement (Fuchs & Fuchs, 1991). CBM is a way to clarify the lack of focus of both teachers and students in understanding the key indicators of growth and basic skills (Deno, 1992).

Not only is CBM an important component in the field of special education, it also has taken on a broader role in general education. With high expectations from the government and legislation, such as in the NCLB Act (No Child Left Behind, n.d.), school personnel are mandated to find assessment tools that measure the achievement levels of all students.

The NCLB Act was implemented in January of 2002 (No Child Left Behind, n.d.). This act sought out to create a solution that emphasized accountability, choice, and flexibility in Federal education programs. When the NCLB Act was passed in 2002, it reauthorized the Elementary and Secondary Education Act of 1985 (ESEA). The NCLB Act incorporates the following: a) increased accountability for states, school districts, and schools; b) greater choice for parents and students, particularly those attending low-performing schools; c) more flexibility for states and local education agencies (LEAs) in the use of Federal education dollars; and d) a stronger emphasis on reading, especially for the youngest children. The area of *increased accountability* is an area of high interest and concern for school districts and parents across the country.

Increased accountability, one of the pillars of the NCLB Act, puts forth strict guidelines for all school districts to meet (No Child Left Behind, n.d.). This pillar focuses on the achievement gap and the extent to how schools must close it. The increased accountability pillar states that all students, including those who are disadvantaged, will achieve academic proficiency in reading and mathematics by the year 2013-14. Proficiency is measured by requiring states to engage in annual testing for all students in grades 3-8. Once the testing is complete, the assessment results and statewide progress objectives are broken out by poverty, race, ethnicity, disability, and limited English proficiency. This is to ensure that no group is left behind. School districts that fail to meet adequate yearly progress (AYP) towards the state's proficiency goals will, over time, be subject to improvement and corrective action. Furthermore, if school districts do not make AYP after five years, dramatic changes will be made to how the school is run.

Like many states, Minnesota has its own state benchmarks and a tool for measuring those benchmarks. Minnesota's measurement system is referred to as the Minnesota Comprehensive Assessments (MCAs) (Minnesota Department of Education, 2003). In 1997, the state legislature enacted into law M.S. 121.113, the Statewide Testing and Reporting System, which established annual testing of all students in grades 3, 5, 7, 8, and in high school. Once the NCLB Act was passed in 2002, the MCAs were expanded to include the assessment of reading and mathematics in grades 4, 6, and 8. This test expansion aligns with the *increased accountability* pillar of the NCLB Act, stating that all states should engage in annual testing for students in grades 3-8 (No Child Left Behind, n.d.). The MCA-IIs will be aligned with the new academic standards and become operational in the year 2006 (Minnesota Department of Education, 2003).

The NCLB Act states, as one of the pillars, to put reading first. President Bush's goal is to ensure that every child can read by the end of third grade (No Child Left Behind, n.d.). Due to this requirement, it is too late to wait until the end of third grade, when the MCAs are given, to decipher which children can and cannot read. An earlier method of identifying students who lack reading skills is through CBM. From early on, research has supported the reliability and validity of CBM for elementary students (Deno, 1985). CBM has been found to correlate with measures of reading performance and to be highly effective in monitoring students' progress in reading programs (Fuchs & Fuchs, 1997).

Research has found that the amount of words a child reads correctly in one minute is a good indicator of general reading ability at the elementary level (Deno, Mirkin, & Chiang, 1982). However, less research has addressed how well reading curriculum-based measurements (R-CBM) can predict which students will meet the reading benchmarks on a statewide assessment. If R-CBMs can indeed aid in predicting performance on state assessments, school districts could use this measurement systematically to screen all students and implement interventions for those whose performance suggest future reading failure.

Purpose of Study

The current study was designed to examine the use of curriculum-based measurement in predicting success on a statewide assessment in the area of reading. The purpose of this paper, then, was to not only identify what constitutes CBM in reading, but also how well it predicts success or failure on Minnesota's statewide assessment (i.e., MCAs). Therefore, the following research questions guided this study:

1. What is the relation between oral reading fluency, as measured by curriculum-based measures of reading, and the Minnesota Comprehensive Assessments (MCAs)?

2. Can educators use oral reading cut scores, as derived from curriculum-based measures of reading, to predict success on the Minnesota Comprehensive Assessments (MCAs)?

Definition of Terms

Adequate Yearly Progress (AYP)- Minimum level of improvement that school districts and schools must achieve each year as determined by the No Child Left Behind Act (No Child Left Behind, n.d.).

Automatic-Activation Process- One of the two processes of word recognition where the stimulus information activates a memory location and spreads automatically to semantically related memory locations that are nearby in the network (Posner & Snyder, 1975a and 1975b as cited in Fuchs, et. al, 2001).

CBM- Curriculum-based measurement (CBM) is a set of measures that can serve as critical indicators of academic performance in the basic skill areas of reading, writing, spelling, and mathematical computation (Deno, 1985).

Conscious-Attention Mechanism- The second of the two processes of word recognition which relies on context to formulate a prediction about the upcoming word and directs the limited capacity processor to the memory location of the expected stimulus (Posner & Snyder, 1975a and 1975b as cited in Fuchs, et. al, 2001).

Increased Accountability- Part of the NCLB Act which requires states to implement statewide accountability systems based on challenging state standards in reading and mathematics, annual testing for all students in grades 3-8, and annual statewide progress objectives ensuring that all groups of students reach proficiency within 12 years (No Child Left Behind, n.d.).

Minnesota Comprehensive Assessments (MCAs)- The Minnesota Comprehensive

Assessment is a standardized, curriculum-based test administered to all 3-8 graders in Minnesota to help schools and districts measure student progress toward the state's academic standards (Minnesota Department of Children, Families, and Learning, 2003).

No Child Left Behind Act (NCLBA)- The most recent reauthorization of the Elementary and

Secondary Education Act (ESEA), which is the principal federal law affecting K-12 education (No Child Left Behind, n.d.).

Reading-Curriculum Based Measurement (R-CBM)- CBM in reading typically consists of 1-

minute reading probes that are developed from the student's curriculum and administered on a regular basis over the course of the school year (Wiley & Deno, 2005).

Treatment Validity- Treatment validity is the degree to which any assessment procedure

contributes to beneficial outcomes (Elliott & Fuchs, 1997).

CHAPTER 2

Review of Literature

This chapter will focus on the relevant literature describing the initial development of curriculum-based measurement. It will then examine curriculum-based measurement in reading and conclude with some of the research regarding the use of curriculum-based measurement in predicting outcomes on state-mandated assessments. Each section will be devoted to answering the primary research questions addressed in chapter one.

Development of Curriculum-based Measurement

In the field of education, school personnel have always struggled with finding ways to best measure student achievement. There has been general agreement that educators should routinely assess the outcomes of instruction (Deno, 1985). How we should do this, however, has always been less clear. Traditionally, the most widely accepted measures of achievement have been commercially developed, standardized, and norm-referenced tests. However, these tools have been criticized for a variety of reasons: a) they are biased regarding curriculum content, b) they are inadequate for making decisions about individual students, and c) they are not useful for making instructional decisions (Shinn, 1998). In an effort to find a measurement tool that solved the problems of norm-referenced tests, curriculum-based measurements emerged.

Deno (1985), from the University of Minnesota, developed curriculum-based measures because he believed that many of the procedures used at that time did not leave the teachers able to measure student performance. In the early 1980s, Salmon-Cox (1981) asked teachers if they would miss standardized tests if they were abolished. Results indicated that the supporters of norm-referenced tests were often people outside of the classroom (e.g., parents, school board members, principals). Many of the teachers believed these tests did not give them anything

worthwhile for their daily instructional decisions. Therefore, Deno (1985) began a research program intended to develop measurement and evaluation procedures that teachers could use every day to make decisions about whether and when to modify a student's instructional program. Researchers then made efforts to expand the role of CBM and examine its usefulness in making eligibility decisions.

A study done by Fuchs and Fuchs (1997) looked at the use of CBM in identifying students with disabilities. The researchers developed a three-stage model, which was operationalized in an elementary school in Nashville during the 1995-1996 academic school year. The first stage, Phase I, consisted of documenting adequate classroom instruction and dual discrepancies. Weekly CBM assessments were done for all students in the school. Every six weeks, meetings were held to review class reports and to formulate two decisions based on the following information: 1) reviews of each classroom to determine whether the overall classroom's progress was adequate (an indication of adequate classroom instruction) and 2) reviews of individual student progress to determine which students met a dual discrepancy criterion. The dual discrepancy criterion was defined as one standard deviation difference between the student's CBM median score and that of classroom peers plus a one standard deviation difference between the student's CBM slope of improvement and that of classroom peers. An assessment team that consisted of the principal, school psychologist, special education teacher, and a social worker conducted these reviews.

Phase II of the Fuchs and Fuchs (1997) study was a prereferral intervention which consisted of at least one member of the assessment team who worked with the general education teacher to design an intervention that addressed the student's underachievement as measured by the dual discrepancy criterion. CBM monitoring continued, and data from the CBMs were used

to assess the effectiveness of the prereferral intervention. Students, for whom the prereferral intervention had not successfully addressed the dual discrepancy, proceeded to Phase III.

Phase III consisted of the assessment team designing and implementing an extended plan (Fuchs and Fuchs, 1997). The purpose of this plan was a diagnostic special education trial period. The plan assessed the potential contribution of special education to the student's learning. CBM continued and was used to determine whether the intervention used during the special education trial period was successful in reducing the student's dual discrepancy. No later than eight weeks into Phase III, the team met again with the students' parents and reviewed the assessment information. If progress was demonstrated during the trial period, the intervention was continued and an IEP (individual education plan) was developed to enhance the program. If the intervention was not successful, the student progressed to the second stage of Phase III.

The second stage of Phase III of the Fuchs and Fuchs (1997) study consisted of the assessment team and parents discussing additional assessment methods for addressing the dual discrepancy (i.e., placing the student back into general education with accommodations or opening an IEP that required a more restrictive placement). No matter what the outcome, CBM was used to monitor the student's progress. The goal of this Three Phase Model was to identify and correct learning problems in the general education classroom environment whenever possible. A second goal was to increase the probability that children who are designated for special education actually require and benefit from those additional services.

In another study done by Elliott and Fuchs (1997), CBM's treatment validity was assessed. First, CBM data was used to document that a student demonstrated a dramatic failure to benefit from an otherwise effective general education learning environment in mathematics. To identify successful strategies, CBM was used. The CBM data revealed that even with

adaptations, the student continued to experience significant problems. Special education was introduced on a trial basis, where the student's CBM data was looked at every day and instructional features were found to help this student's problem. With this individually tailored program, implemented on a one-to-one basis, the data directed the educator toward a successful special education intervention for the student.

Another study investigated whether a well-developed CBM system could be used to establish growth standards for students with learning disabilities in the area of reading (Deno, Fuchs, Marston, & Shinn, 2001). The participants in this study consisted of a convenience sample obtained from four different local education agencies (LEAs) across the United States. The LEAs were located in an urban, Midwestern, Southeastern, and a medium-sized town in a western mountain state. The student sample was separated by grade level and whether they did or did not receive special education services. The number of words read aloud correctly in one minute was scored and analyzed. The CBM passages remained constant in difficulty across the school year so that the changes in scores reflected changes in overall reading proficiency. The CBM median scores for three passages were obtained in fall, winter, and spring for each student. The weekly increase in number of words read correctly was computed and averaged with the weekly increases across students within grade levels and within general or special education programs. The researchers found significant growth rate differences between the general and special education students in first grade. It was found that these differences progressively decreased until fifth and sixth grades, where a decline in the slope line for general education students resulted in identical growth rates between the general and special education students. Overall, the researchers found that special education students were behind after their first year of reading instruction. Other results indicate they continued to fall further behind through the early

grades as the reading skills of their general education peers increased more rapidly until fifth or sixth grade. Based on this study, the researchers concluded that it is possible to set growth standards for both general and special education students using repeated measures of students' performance through CBM (Deno, Fuchs, Marston, & Shinn, 2001).

What is Curriculum-based Measurement in Reading?

Stanley Deno and colleagues (1985) set out to establish a measurement system for academics that a) teachers could use quickly; b) would produce accurate, meaningful information with which to index growth; c) could answer questions about the effectiveness of programs in producing academic growth; and d) would provide information to help teachers plan better instructional programs (Fuchs and Fuchs, 1997). Reading comprehension was the first academic skill addressed by the CBM researchers. In the classroom, the student's ability to comprehend text typically is based on whether or not the student can answer questions related to the content of the text material. However, writing comprehension questions for student use proved to be too time consuming. Supplying words deleted from text, saying the meanings of words underlined in the text, reading aloud from isolated word lists, or reading aloud from text passages were subsequent methods developed to assess reading comprehension. However, the question arose as to whether those tasks could be used validly to measure reading achievement. Researchers found that all of the curriculum-based measures (i.e., supplying words deleted from text, reading aloud from text passages) were highly correlated with performance on the standardized, norm referenced tests except for the word-meaning task. They also found that reading aloud from text discriminated between students in special education and students in general education, with students in general education reading more words correctly from text material than students in Chapter 1 programs. All in all, Deno found the use of CBM improved communication

(developed clear and effective communication of student performance), sensitively measured student progress (were sensitive to growth in student performance over short periods of time), were peer referenced (teachers were able to obtain a normative perspective on student performance by sampling regular classroom peers), and were cost effective.

To further analyze why reading aloud from text discriminated between students in special education and students in general education, Posner and Snyder (1975a and 1975b as cited in Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R., 2001) studied the process of word recognition. According to the researchers, there are two processes of word recognition. The first is referred to as the automatic-activation process, or where stimulus information activates a memory location and spreads automatically to semantically-related memory locations that are nearby in the network. The automatic-activation process is fast acting and requires little capacity for attention. The second process, conscious-attention mechanism, relies on context to formulate a prediction about the upcoming word and directs the limited capacity processor to the memory location of the expected stimulus. This process is slow acting and needs attention. According to Posner and Snyder, comprehension results from the combined effect of the two processes for poor readers. Therefore, poor readers need to use their attention capacity for predicting processes to aid in word recognition, leaving little attention for comprehension. Fuchs et al. (2001) concluded that fluent oral reading from text serves as a performance indicator of overall reading competence, including the readers capacity to process meaningful connections within and between sentences, to infer the macrostructure of a passage, to relate text meaning by checking consistencies with prior information, and to make inferences to supply missing information (i.e., comprehension).

To determine if oral reading fluency measures correlated with direct measures of reading comprehension, a study was done by Fuchs, Fuchs, and Maxwell (1988) in the late 1980s. Fuchs and colleagues correlated four alternative measures to the Reading Comprehension subtest of the Stanford Achievement Test (SAT). Seventy middle school and junior high school students with reading disabilities were administered the SAT and the alternative measures. The alternative measures included question answering (students read two 400 word passages for five minutes and then provided oral answers to 10 short-answer questions); passage recall (students read one 400 word passage for five minutes and had 10 minutes to retell the passage); a cloze method (every seventh word was deleted from each of the 400-word passages and replaced each deleted word with a blank); and oral reading fluency (students read two of the 400-word passages aloud, each for five minutes, while the examiner scored omissions, repetitions, substitutions, and mispronunciations as errors). The criterion validity coefficients for the question answering, the recall, and the cloze measures were .82, .70, and .72, respectively. The coefficient for oral reading fluency was .91. This study's researchers concluded that oral reading fluency was most strongly associated with the capacity to read passages and answer questions about those passages as assessed by a widely used, commercial achievement test of reading comprehension.

Using Oral Reading Rate to Predict Future Performance on State-Mandated Tests

Because it has been proven that oral reading fluency is strongly correlated with reading comprehension (Deno, Mirkin, & Chiang, 1982; Fuchs, Fuchs, & Maxwell, 1988; Fuchs et al., 2001) and the most commonly used classroom assessments of reading comprehension are commercial standardized tests, it was deemed imperative that researchers begin incorporating curriculum-based measurements as a way to predict success on high stakes, state mandated tests.

Currently, much research is underway to investigate the use of curriculum-based measurements to predict students' outcome on state mandated tests.

Stage and Jacobsen (2001) conducted one of the first studies to appear on this topic. In this study, 173 fourth graders were administered curriculum-based oral reading fluency probes in September, January, and May. The researchers then conducted growth curve analysis (GCA) to determine the relations between students' slope in oral reading fluency and the reading portion of the state mandated test administered in the state of Washington during the month of May (i.e., the Washington Assessment of Student Learning or WASL). The researchers found that both the slope in oral reading fluency across the school year and the oral reading fluency probes administered in September, January, and May reliably predicted WASL reading performance in May. More specifically, this study found the low scores in oral reading fluency predicted WASL failure ($r = .41$), and the high scores in oral reading fluency predicted WASL success ($r = .90$).

Another study was conducted by Crawford, Tindal, and Stieber (2001). In this study, researchers collected longitudinal data from students over two years as they moved from second to third grade. The researchers not only compared the relations between the oral reading measures and reading scores on the statewide achievement test, but they also correlated their measures with the math scores on the achievement test. The researchers decided to examine the math scores for three reasons: a) math multiple-choice achievement tests require proficient reading skills, b) earlier research (McGrew & Pehl, 1988 as seen in Crawford et al., 2001) demonstrated that reading achievement is a strong predictor of math achievement; and c) initial research on accommodations indicated that reading math multiple-choice tests aloud may help a student with reading problems (Tindal, Heath, Hollenbeck, Almond, & Harniss, 1998 as seen in Crawford et al., 2001).

In the Crawford, Tindal, and Stieber (2001) study, the participants were administered three CBM reading probes one day in January for the two years of the study. In March in the second year of the study, participants were administered the math and reading portions from the statewide assessment. The results indicated that 81% of students reading at the 50th percentile and above passed the statewide reading test. Further, researchers found that a reading rate of 119 words per minute almost ensured that a student passed the statewide reading test. In the across-year analysis, 100% of the 2nd grade students reading at least 72 correct words per minute passed the statewide reading test in third grade. The findings in math, using the same set of analyses, however, were not significant. The researchers found, however, that 82% of the students reading at or below 54 correct words per minute in second grade failed the statewide math test the following year. This study was important because it confirmed past findings that CBMs are sensitive enough to detect growth for almost every student and teachers can rely on the accuracy of CBMs to monitor the reading progress of all students, regardless of skill level.

More recently, a study done by McGlinchey and Hixson (2004) expanded the predictive validity research by examining 1,362 students over eight years. Like the other studies, the researchers investigated the correlational data and predictive value of reading CBMs in Michigan. Michigan's statewide assessment, or the Michigan Educational Assessment Program (MEAP), was used as the outcome measure. Slightly different from other research, researchers administered only one one-minute oral reading sample for the first five years and three one-minute reading samples during the last three years. Similar results were found, regardless of how many probes were used. The validity coefficients in the last 3 years, when three one-minute probes were used, ranged .65 to .81. In the first five years, when one one-minute probe was used, validity coefficients ranged from .49 to .77. Findings suggested that using a single one-

minute probe did not substantially affect the stability of the measure. This study used 100 words read correct per minute (WRCM) as the cut score. Using that cut score, the probability of correctly identifying students who scored below *Satisfactory* on the MEAP was 77%. The probability of correctly identifying students who achieved *Satisfactory* was 72%. Once again, these results confirmed the predictive power of R-CBM.

Overall, preliminary research has concluded that R-CBMs are good indicators of predicting outcomes on state assessments. However, because each state uses different assessment tools, generalizing between these studies always raises a question. For example, a study done by Havey, Herrera, and McCormick (presented at the 2005 NASP convention) looked at the relations between R-CBM and the Illinois Standards Achievement Test (ISAT). This study used students in the fifth grade. The following results were found: 38% of students who read 91 words per minute met standards on the ISAT; 22% of students who read at that same rate did not meet standards; and 10% of students who read below that rate failed to meet standards. Overall, the CBM results correctly predicted 80% of those students who passed the ISAT. The researchers concluded that these were modest results. Thus, the researchers did not believe that R-CBM was a strong predictor of performance on the ISAT. However, concerns were raised about the state assessment's validity. Therefore, it is important for each state to determine the predictive power of R-CBM for their respective state assessments. Because of this importance, this study will focus on the state assessment administered in Minnesota, the Minnesota Comprehensive Assessment (MCA).

CHAPTER 3

Method

Participants

The study took place in two districts in Western Minnesota during the 2004-2005 school year. The two districts had a combined enrollment of approximately 1600 students, including approximately 93 students in the third grade. There were three elementary buildings with third grade students. In one district, the non-Caucasian population was 3%, and *free and reduced* lunch status (an indicator of socioeconomic need) was 32%. In the other district, the non-Caucasian population was 8%, and *free and reduced* lunch status was 38%.

The students involved in this study were third graders, as third grade is the first year that the MCAs are administered in the state of Minnesota. There were a total of 54 (58% of the total third grade population) students who returned with signed parent permission slips. Forty-six percent of the subject sample was male, and 54% were female.

Materials

CBM probe. Three passages were selected from the Jim Wright (www.interventioncentral.org/htmldocs/interventions/cbmwarehouse.shtml) web site after permission was granted. Each student was given three passages at grade level with a maximum of 250-words (see Appendix A for a sample passage). The students read aloud all three probes for one-minute each. The correct words read aloud were recorded for each of the three passages. The median score was then recorded for each student.

Minnesota Comprehensive Assessments (MCAs). The MCAs are Minnesota's statewide tests based on the benchmarks approved by the Minnesota Department of Education. The assessment system assesses reading, math, and writing in grades 3-8.

The third grade reading MCAs include both multiple-choice and short answer questions. There is no penalty for guessing. To ensure validity and reliability, the test vendor scores the assessment, including the reading response items. The test blueprints are aligned with the Minnesota High Standards. These standards were developed with the assistance of teachers from across Minnesota, teaching at or near the requisite grade level, as well as with the help of national experts. These blueprints determine the requirement of each test and each test form. The general guidelines for developing the test items of the grade three reading MCAs are of three types:

1. *Informational (nonfiction selections that would commonly be found in age-appropriate periodicals, reference book or textbooks)*
2. *Practical (informational selections intended for a specific application such as simple recipes, "how to" instructions, or advertisements)*
3. *Literary (fictional selections such as short stories, poems or excerpts from longer passages, novels or other age-appropriate literary publications).*

The grade three selections range in length from 250 to 800 words. Students perform within achievement levels ranging from Level 1 to Level 5. The level definitions include:

Level 1- Gaps in knowledge and skill

Level 2- Partial knowledge and skills

Level 3- Solid grade level skills

Level 4- Working above grade level

Level 5- Superior performance beyond grade level

For the purpose of this study, students who score at Level 3 or above will be regarded as passing.

Procedure

The school psychologist for both districts administered and scored the CBM reading probes as described by Shinn (1989). All students were administered the same three reading passages two to three weeks prior to the administration of the MCAs. The probes were administered in mid-April, and the MCAs were administered in early May. The CBM reading probes were administered in the school psychologist's office in two of the elementary schools, and in a quiet room next to the library in the third elementary school.

The child's regular teacher typically administered the MCA test. However, in one school, the school counselor was the test administrator. Students identified for special education may have had an accommodation that allowed for individualized administration or breaking the test into smaller units. A 2-week testing window was established in which all schools were expected to complete both the reading and math portions of the MCAs.

Data Collection Procedures

Approximately four weeks before the MCAs were administered, a permission packet was sent home with each third grade student. Each packet included a cover letter explaining the research study and a parent permission form for the student's parent or guardian to sign and return.

About one week later, the students with parent permission were asked to read each of the three reading probes. The number of words read correctly per minute for each reading probe was recorded as well as the median score of the three probes. When the MCA results were sent to the school, the achievement level was recorded for the reading portion.

Confidentiality of participants was maintained by coding all student records and their respective test scores. After the achievement level from the reading portion of the MCA was recorded, all identifying information was destroyed.

Data Analysis

Individual student data were analyzed and diagnostic efficiency statistics were used to determine the accuracy of the reading rate cut score (Stage & Jacobsen, 2001). Ninety WRCM was selected as the cut score set forth by the Dynamic Indicators of Basic Early Literacy Skills (DIBELS). The DIBELS homepage specified the range of 80 WRCM to 110 WRCM as the benchmark for spring of third grade (Official DIBELS Homepage, n.d.).

The first research question addressed the relation between oral reading fluency, as measured by curriculum-based measures of reading, and the Minnesota Comprehensive Assessment Reading Test. To examine this relation, Pearson product-moment correlation coefficients were computed between the oral reading fluency median scores (WRCM) and the scale and level scores of the reading portion of the MCA. A probability level of .01 was used to determine statistical significance.

The second research question addressed whether educators could use oral reading cut scores, as derived from curriculum-based measures of reading, to predict success on the Minnesota Comprehensive Assessments (MCAs). The following diagnostic accuracy statistics were used:

1. *True-Positive: the proportion of those students who were predicted to pass the MCA and did pass the MCA.*
2. *True-Negative: the proportion of those students who were predicted to fail the MCA and did fail the MCA.*

3. *False-Positive: the proportion of those who were predicted to pass the MCA and failed the MCA.*
4. *False-Negative: the proportion of those who were predicted to fail the MCA and passed the MCA.*

CHAPTER 4

Results

This chapter presents the results of the study to determine the relation between oral reading fluency, as measured by curriculum-based measures of reading, and the Minnesota Comprehensive Assessments (MCA). The ability of oral reading cut scores to predict success on the MCA Reading Test is also presented. Mean scores, standard deviations, Pearson product-moment correlation coefficients, and diagnostic efficiency statistics were utilized to answer the research questions.

Preliminary Analyses

Data Screening

After scatter plots of the data were analyzed, one outlier appeared. One student's WRCM median score was low ($Mdn = 79$), and his MCA scale and level scores were exceptionally high (i.e., 2150 and 5, respectively). Because of this uncharacteristic pattern of scores, that student's scores were removed from the analysis, resulting in a total of 53 participants.

Descriptive Statistics

Table 1 (Appendix B) provides the means and standard deviations of the CBM oral reading fluency median scores (i.e., WRCM). The mean and standard deviation for the MCA scale and level scores are also provided. Most participants scored at or above the cut level score of 3 on the reading portion of the MCA ($n = 46$; 86.8%)

Research Question One

The first research question addressed the relation between oral reading fluency, as measured by curriculum-based measures of reading, and the Minnesota Comprehensive Assessment Reading Test. Table 2 (Appendix B) shows the Pearson product-moment correlation

coefficients between the oral reading fluency median scores (WRCM) and the scale and level scores of the reading portion of the MCA. As can be seen, the correlation between WRCM and the MCA scale scores ($r = .687$) and the MCA level scores ($r = .697$) were statistically significant at the $p < .01$ level. Both correlation coefficients suggest moderate to strong correlations between the measures. The correlation size suggests a 47% (MCA raw scores) and 49% (MCA level scores) shared variance between the predictor and the criterion measures.

Research Question Number Two

The second research question addresses whether educators use oral reading cut scores, as derived from curriculum-based measures of reading, to predict success on the Minnesota Comprehensive Assessments (MCAs). As shown in Table 3 (Appendix B), with using 90 WRCM as the cut score, the following diagnostic accuracy statistics were obtained. The True-Positive score shows the percentage of those who earned scores at or above the 90 WRCM criterion and passed the MCA Reading Test (i.e., scored at or above level 3). As can be seen in Table 3, 44 students were predicted to pass the MCA Reading Test based on their WRCM median score. Results indicate 43 out of the 44 (98%) passed the MCA, as predicted.

The True-Negative score shows the percentage of those who earned scores below the 90 WRCM criterion and failed the MCA Reading Test (i.e., scored below level 3). As indicated in Table 3, only 9 students were predicted to fail the MCA Reading test based on their WRCM median score. Results indicate 6 out of the 9 (67%) failed the MCA, as predicted by their WRCM score.

The False-Positive score shows the percent of those who earned scores at or above the 90 WRCM criterion and failed the MCA Reading Test. As displayed in Table 3, only 1 student out

of 44 students (2%) was predicted to pass the MCA Reading Test but failed to pass the MCA Reading Test based on the student's level score.

The False-Negative score shows the percent of those who earned scores below the 90 WRCM criterion and passed the MCA Reading Test. As indicated in Table 3, only 3 out of the 9 students (33%) were predicted to fail the MCA Reading Test, but did pass the MCA Reading test based on their level scores.

CHAPTER 5

Discussion

This chapter will discuss the findings of the study. The results will then be compared with the findings of previous research. Finally, the limitations of the study will be discussed with suggestions for future research and educational practice.

Research Question One:

The first research question addressed the relation between oral reading fluency and the Minnesota Comprehensive Assessments (MCAs). The findings indicate that a student's WRCM is moderately related to a student's score on the MCA reading portion. These results appear to be consistent with previous research. Other studies (Crawford, Tindal, & Stieber, 2001) also found a moderate to strong correlation when the reading probes were administered the same year as the state assessment as well as when the probes were administered one year prior to the statewide assessments. A possible explanation for these correlations is the strong relationship between WRCM and a student's overall reading ability (Fuchs, Fuchs, & Maxwell, 1988).

As stated, the current study found moderately strong correlations between WRCM and Minnesota's state assessment, the MCAs. The MCAs are first given to students in the third grade, which has been stated as too late to increase reading skills in students (Speece, Mills, Ritchey, & Hillman, 2003). A more recent study done by Hintze and Silbergliitt (2005) supported the usage of WRCM as a good predictor for overall reading ability. The study also expanded the research and found that WRCM could predict students who were likely to pass the reading portions of the MCAs as early as the first grade. This research supports the use of WRCM to screen students in the first grade, as educators would be better able to prepare their students years before the statewide test is administered. Such a strategy might result in having

fewer students fall less behind in overall reading skills, resulting in less pressure from the No Child Left Behind Act. Teachers also would be able to document their attempts to ensure that every child establishes proficiency in reading.

Research Question Two:

The second research question addressed whether educators could use cut scores to predict success on the Minnesota Comprehensive Assessments (MCAs). The current study found a strong True-Positive score, suggesting that using 90 WRCM as the cut score correctly identified 98% of the students predicted to pass the reading portion of the MCA. Further analysis of this data found that all students (i.e, 100%) who had a WRCM score of 75 or less failed the MCA reading portion. This information is clear and simple, and indicates teachers can use cut scores to accurately predict student success or failure on the MCA.

Although identifying a specific cut score helped predict success on the MCA reading portion, the True-Negative score (identifying those students who obtained a WRCM below the 90 WRCM cut score and did not pass the MCA reading portion) was not as powerful. This study found that using a cut score of below 90 WRCM correctly identified only 67% of students as not passing the MCAs. It is confusing that a tool would be deemed so useful in predicting success, but be found less accurate in predicting failure. Stan Deno (1985) made the following assertion:

one minute reading samples can be used by teachers to monitor growth in reading and that the number of words read correctly and incorrectly in those repeated samples of reading aloud 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. (p. 224)

Stage and Jacobsen (2001) also found their True-Negative score to be weaker than their True-Positive score. One explanation for the weaker True-Negative score is that the adopted cut score (90 WRCM) was too high. The higher the cut score, the higher percentage of students who meet the cut score will receive a satisfactory score; however, many students receiving less than that will also receive a satisfactory score. In this study, there also were a few instances in which a student minimally performed above the cutoff score, yet failed the reading portion of the MCAs. In those cases, it is difficult to say why those students did not pass the MCAs. Just as Deno (1985) stated, WRCM does not give us a complete picture of a student's reading proficiency.

Limitations

There are three possible limitations of the present study. The first limitation of this research study was the limited sample size. Since only those students who obtained parent permission were allowed to participate, it is difficult to say if the correlation would have been as strong had all third grade students across the two districts participated. Due to the limited number of participants, examining the effects other possible variables was virtually impossible. The effects of gender, English as a second language (Wiley & Deno, 2005), or special programming on the correlations and the predictive validity of WRCM would have enriched the results of this study.

A second limitation was the geographic location of the participant sample. The students of both districts were from rural Minnesota and predominately Caucasian. Thus, this regional study limits the ability to generalize the results to other parts of Minnesota or other states where

other ethnicities are present. The ability to generalize these results to other areas may be restricted to other small, rural regions of Minnesota.

A third limitation would be the use of Minnesota Comprehensive Assessments as a criterion measure. Each state appears to have its own specific assessment related to their respective standards. Given each state has a different assessment, researchers and educators should be cautious in generalizing these findings to other states. Currently, a variety of research exists to examine the use of CBM to predict success on other state assessments (Crawford, Tindal, & Stieber, 2001; McGlinchey & Hixson, 2004; Stage & Jacobsen, 2001). A school district looking to use curriculum-based measurements to help predict success on their state assessment should look for previous research and the specific correlations between their state assessment and students' WRCM.

Implications for Practice

Several implications can be drawn from the results of this study and the information provided by previous research. Educators can be confident in using a student's WRCM score to design programs to meet the needs of students at all elementary levels. Finding a student's WRCM score is an inexpensive, predictive, and useful tool to aid school personnel in identifying students who may be at-risk for falling behind in reading and performing poorly on statewide assessments. Many times, educators wait for students to show signs of failure on statewide, norm-referenced tests and then turn to special education to teach the deficient basic skills. Finding WRCM scores as early as the first grade, or even as early as preschool through using DIBELS (Dynamic Indicators of Basic Early Literacy Skills), eliminates the 'wait to fail' process and takes a proactive approach to assure all students are gaining the basic skills to be successful in school. If a student is found to have a low WRCM score, educators can try an appropriate

intervention for the student and continue CBM use to measure the progress of the intervention. This tool can be used at all elementary levels from the district, school, grade, and individual level (McGlinchey & Hixson, 2004). Overall, this tool allows for the consultant role of the school psychologist and makes the case for school psychologists to expand their services from special education students to all students.

Suggestions for Future Research

This study poses many questions that may be answered in future research. It would be beneficial to conduct a similar study, but examine the predictive validity of the WRCM screeners longitudinally. As teachers receive more pressure to ensure no student is falling behind, emphasis will be put on screening tools to identify those students who are at-risk for reading failure.

A second suggestion would be to take all of the research on using CBM and WRCM in predicting success on state mandated tests and conduct further research on what interventions work for those students identified as at-risk. As researchers, we sometimes focus our attention on what does not work; however, investigating what does work is somewhat lacking in the field of education.

Summary

Curriculum-based measurement (CBM) is an assessment tool that gives information on how a child is performing academically (i.e., reading, math, spelling, and writing). CBM can also be useful in helping teachers evaluate their programs and determine students' eligibility for special education services. This study examined the predictive validity of CBM in predicting future success on Minnesota's state assessment system, the Minnesota Comprehensive Assessments. Results indicate moderately strong correlations between students' WRCM and

their reading levels as assessed by a statewide assessment. Results suggest educators may use WRCM as a means of identifying those students who are at risk for reading failure. As such, using WRCM cut scores will assist educators to be proactive in their teaching methods as opposed to waiting until it is too late.

Reading is a foundational skill. If a student lacks basic reading skills, all other subjects will suffer. Therefore, it is essential for teachers to help those students who are falling behind before it impacts the rest of their lives. Using CBM screeners can effectively determine those most at risk for reading failure, thereby allowing teachers to address the reading problems of their students before reading deficits affect student success in other curricular areas.

References

- 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. (1992). The nature and development of curriculum-based measurement. *Preventing School Failure, 36*(2), 5-10.
- Deno, S. L., Fuchs, L. S., Marston, D., & Shinn, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*, 507-524.
- Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*, 36-45.
- Dynamic Indicators of Basic Early Literacy Skills (DIBELS) Homepage. (n.d.) Retrieved March 20, 2005, from <http://dibels.uoregon.edu/>
- Elliott, S. N. & Fuchs, L. S. (1997). The utility of curriculum-based measurement and performance assessment as alternatives to traditional intelligence and achievement tests. *School Psychology Review, 26*, 224-234.
- Fuchs, L. S. & Fuchs, D. (1991). Curriculum-based measurements. *Preventing School Failure, 35*, 6-11.
- Fuchs, L. S. & Fuchs, D. (1997). Use of curriculum-based measurement in identifying students with disabilities. *Focus on Exceptional Children, 30*, 1-16.

- 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).
- Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal measures of reading comprehension. *Remedial and Special Education, 9*, 20-28.
- Havey, J. M., Herrera, A., & McCormick, C. (2005, March). *The relationship between reading cbm and the Illinois standards achievement test*. Poster presented at the meeting of the National Association of School Psychologists, Atlanta, GA.
- Hintze J. M. & 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. .
- McGlinchey, M. T. & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 193-203.
- Minnesota Department of Children, Families, and Learning. (2003). *Minnesota Comprehensive Assessments: Grade 3 reading test specifications*. Rosevill, MN: Author.
- No Child Left Behind Act of 2001, *Pub. L. No. 107-110*.
- Salmon-Cox, L. (1981). Teachers and standardized achievement tests: What's really happening? *Phi Delta Kappan, 62*(9), 631-634.
- 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 Press.

- Shinn, M. R. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology-IV* (pp. 671-697). Bethesda, MD: National Association of School Psychologists.
- Shinn, M. R. & Habedank, L. (1992). Curriculum-based measurement in special education problem identification and certification decisions. *Preventing School Failure, 36*, 11-16.
- Speece, D. L., Mills, C., Ritchey, K. D., & Hillman, E. (2003). Initial evidence that letter fluency tasks are valid indicators of early reading skill. *Journal of Special Education, 36*(4).
- 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.
- Wiley, H. I. & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for english learners on a state standards assessment. *Remedial and Special Education, 26*, 207-214.

Appendix A

Sample of Reading Probe

Once upon a time there was a little girl whose home was down a deserted road deep in the forest. The little girl loved living in the forest among the animals. She was a friend to all the animals of the forest. Yet, she was also a lonely little girl because she had no family. One day while skipping through the woods, she met a beautiful fairy. The fairy questioned the girl about where she lived. The girl told the fairy about her forest home. The fairy asked the girl about her family. The girl told the fairy that she did not have a family. The fairy could feel that the girl was lonely, so she asked the girl if she wanted to live with all the fairies. The girl happily answered yes and went to live with the fairy. For the rest of her life the girl had her friends, the animals, and a family of her very own.

Appendix B

Tables

Table 1

Means and Standard Deviations of Oral Reading Fluency and MCA Reading Scores

Measure	<i>M</i>	<i>SD</i>
WRCM	115.81	26.583
MCA Scale Score	1565.47	176.249
MCA Level Score	3.72	1.063

Note: WRCM = Words Read Correctly Per Minute; MCA = Minnesota Comprehensive Assessment Reading Test; $n = 53$.

Table 2

Correlations Between the MCA and Words Read Correctly Per Minute (WRCM)

Measure	MCA scale score	MCA level score
WRCM	.687*	.697*

Note: $n = 53$. * $p < .01$.

Table 3

Diagnostic Accuracy Statistics

Statistic	<i>n</i>	<i>P</i>
True-Positive ^a	43/44	98%
True-Negative ^b	6/9	67%
False-Positive ^c	1/44	2%
False-Negative ^d	3/9	33%

Note. ^aRepresents the proportion those who were predicted to pass the MCA and did pass the MCA. ^bRepresents the proportion of those who were predicted to fail the MCA and did fail the MCA. ^cRepresents the proportion of those who were predicted to pass the MCA and failed the MCA. ^dRepresents the proportion of those who were predicted to fail the MCA and passed the MCA.