Curriculum-Based Measurements:

Middle School Mathematics

by

Josh Chaplin

A Research Paper
Submitted in Partial Fulfillment of the
Requirements for the
Master of Science Degree
in

School Psychology

Approved: 2 Semester Credits

Jacalyn Weissenburger, Ph.D.

The Graduate School

University of Wisconsin-Stout

December, 2010

The Graduate School
University of Wisconsin-Stout
Menomonie, WI

Author:	**Chaplin, Josh T.**
Title:	***Curriculum-Based Measurements: Middle School Mathematics***
Graduate Degree/ Major:	**M.S.Ed. in School Psychology**
Research Adviser:	**Jacalyn Weissenburger, Ph.D.**
Month/Year:	**December, 2010**
Number of Pages:	**40**
Style Manual Used:	**American Psychological Association, 6th edition**

Abstract

The United States' students trail their peers around the world in the areas of science and mathematics. Educators and lawmakers have called for reform not only in the instruction of mathematics, but also for methods to screen and monitor students who do not meet basic standards of mathematics. This literature review examined (a) the historical context of the assessment process in school psychology, (b) the state and need for mathematic skills in middle school, and (c) curriculum-based measurements (CBM) and their need in middle school math to assess, inform instruction, and monitor student progress. The review found that limited research exists in the area of curriculum-based measurements of mathematics at the middle school level. Recommendations for future research and implications for practice are discussed.

The Graduate School
University of Wisconsin Stout
Menomonie, WI

Acknowledgments

I would like to acknowledge some individuals who have significantly contributed not only to my graduate education but also to my life. I would first and foremost like to acknowledge and thank my parents, Rodney and Sylvia Chaplin, for their unwavering support not only in my educational career, but in my life as well. Without them, I would not be in the position I am today. I would also like to thank my second family, the Hacketts, as they have provided me additional support in my academic career and my life. I would like to thank the Bangor School District and community as they have had both a direct and indirect positive impact on my graduate career. I need to also acknowledge Dr. Jackie Weissenburger because she is the person who selected me as a student for the school psychology program at the University of Wisconsin-Stout and is the advisor for this thesis. Finally, I also have to thank Dr. Carlos Dejud and Dr. Christine Peterson because they not only have been tremendous teachers and advisors, but they have also been willing to accommodate my educational plan.

TABLE OF CONTENTS

Chapter One: Introduction

A quote from the American Television Series *The Wonder Years* stated, “Growing up is never easy. You hold onto things that were. You wonder what’s to come.” This quote describes the state of the field of school psychology. School psychology has been said to be a relatively new specialty in the field of psychology; however, the field of school psychology has moved past its infancy and early developmental stages and is now in early adulthood. Like early adults, who are competent in many areas but are still searching to fulfill their potential in life, the field of school psychology is well developed in many areas, but school psychologists are still searching to best fulfill the field’s potential. School psychology’s founding principle was and still is to help solve children’s learning problems (Bardon & Bennett, 1974 as cited by Merrell, Ervin, & Gimpel, 2006).

Many different principles, theories, and tools have been developed as the basis of school psychology, and the field still holds onto many of the original ideas. However, as the need for the field of school psychology continues to grow, the field is at a crossroads regarding assessment. Should school psychology stick to traditional assessment models and tools, move away from assessment-related activities, as they have been integral in the role of many school psychologists, or move toward something new in terms of assessing children? A fundamental question arises when it comes to assessment in any field. Why are we assessing something? Is it to just gather data or is it to try and solve a problem?

In 1991, President George H.W. Bush stated, “If we want America to remain a leader, a force for good in the world, we must lead the way in educational innovation” (Alexander, 1993, p. 2). President Bush initiated AMERICA 2000, which was one of the boldest and most complex long-range strategies to move all school and children toward national education goals set the previous year by President Bush and all the governors. This act was enacted into federal legislation in 1994 and was known as the Goals 2000: Educate America Act (P.L. 103-227). GOALS 2000 had eight national educational goals. While the United States did not meet any of these proposed educational goals by the year 2000, the nation did make measurable progress (Cooper, 1999).

One of the national goals was that the United States would be first in the world in mathematics and science achievement. Disappointingly, the U.S. did not meet its goal of being first in the world in mathematics and science; moreover, the scores from the 2006 Program for International Student Assessment (PISA) showed that the U.S. 15-year-olds trailed their peers from many industrialized countries (Glod, 2007). The PISA test is given every three years and measures the ability of 15-year-olds to apply math and science knowledge in real-life contexts. The test was given to about 400,000 students worldwide, including 5,600 students in the U.S. In 2006, the United States ranked 23rd in the world in the math portion of the test, which was about the same standing as when the test was given in 2003 (Glod, 2007).

President Barrack Obama confirmed this ranking again in 2009, stating American students ranked 25th in math compared with students around the world. According to Glod (2007), “The PISA results underscore concerns that too few U.S. students are prepared to become engineers, scientists, and physicians, and that the country might lose ground to competitors.” While panels and committees have been appointed to recommend ways to improve public school math instruction, little has been done to monitor the progress of students in math.

The previous questions are raised in this particular paper because of the need to assist educators in improving student performance in mathematics. Numerous empirical studies on curriculum-based measurement (CBM) has confirmed its usefulness as a system of screening students at-risk for failure in academic domains (i.e., reading, math, or writing) and for progress monitoring student gains in achievement, particularly in the area of reading (Foegen, Jiban, & Deno, 2007); however, there has been less research conducted for CBM in the area of mathematics, specifically at the middle school level. This research will examine: (a) the historical context of the assessment process in school psychology, (b) the development and measures of CBMs of middles school mathematics, and (c) the current research on CBMs of middle school math to effectively assess, inform instruction, and monitor student progress.

Statement of Problem

Elley (1988) reported reading and literacy are the best indicators of intellectual ability and an accurate predictor of success at school, and literary-related skills are often the most heavily emphasized in school. While many people, such as Hosp, Hosp, and Howell (2007), have argued math is similarly important for success in life, the achievement of United States' students in math, particularly as children move into middle and high school, has continued to be lower than other countries around the world.

Many factors contribute to the success or failure of a middle school student in the area of math. For example, Bryan (2005) says research has shown children are more likely to succeed in academic achievement when their families are involved; however, research has also shown as children move from elementary school to middle school, parental involvement decreases. It is difficult to determine the exact reason for this decline; however, parental level of education, particularly in the case of mathematics, many be a contributing factor, as some parents are less knowledgeable about this content area (Constantino, 2007). While educators expect children to become more independent as they enter middle school and high school, school personnel need to be aware students may be getting less and less academic help at home, particularly in the area of math.

By using progress monitoring tools like CBMs, school personnel would have an easy and valid way of quickly tracking students' progress in math. Teachers can then use this data to inform their instruction to help students better succeed in mathematics at school. Most middle school mathematics teachers have various information sources about a student's achievement for specific topics within the classroom such as tests, quizzes, homework, and other assignments; yet, an indication of the student's overall understanding of mathematics can be lacking. Helwig, Anderson, and Tindal (2002) said, "Measures that give teachers snapshots of students' conceptual understanding of grade-level content can fill this void" (p. 103); however, little research in the particular area of CBMs in middle school mathematics has been conducted. This paper will review the limited research on CBMs in middle school mathematics and discuss implications of for research.

Research Objectives and Questions

The purpose of this research is to identify if CBMs are useful; and, more importantly, reliable and valid in assessing, informing instruction, and monitoring student progress in the area of mathematics. The research will look at the differences between CBA, CBM, and traditional assessments, discuss the current research in CBM in middle school mathematics, review the reliability and criterion validity of CBMs in middle school mathematics, determine how CBMs are used in middle school math, and how schools can further use CBMs to assess developing proficiency in mathematics at the middle school level.

The main questions for this study include the following:

1. What is the historical context of assessment in the field of school psychology and what is CBM's role in assessment?
2. What is the current state of research in the area of CBMs of middle school mathematics?
3. What is the reliability and criterion-related validity of the current CBM measures used to assess and monitor progress mathematic skills at the middle school level?

Definition of Terms

Traditional Assessments. Standardized, norm-referenced cognitive and achievement tests used to assess academic skills. Examples include the Wechsler Intelligence Scale for Children-IV, and the Woodcock-Johnson-III, Kaufman Test of Educational Achievement-Second Edition.

Curriculum-Based Assessment. Test stimuli taken from the student's curriculum; the student is tested repeatedly over time and the information from the tests is used to inform instruction.

Mastery Measurement. Taking global skills and breaking them into a set of sub-skills and assess to determine proficiencies for each sub-skill.

General Outcome Measurement. Using standardized procedures to assess overall proficiency in a given academic domain, such as reading, math, and writing, over a longer period of time (e.g., one academic year).

Curriculum-Based Measurement. A specific type of CBA which involves a set of specific procedures that are technically adequate, include standardized measurement tasks, intended to be time efficient, and are indicators of overall proficiency in an academic domain (i.e., reading, math, and writing).

Chapter Two: Literature Review

Cesar Chavez once said, "Once social change begins, it cannot be reversed. You cannot uneducate the person who has learned to read. You cannot humiliate the person who feels pride. You cannot oppress the people who are not afraid anymore. We have seen the future, and the future is ours." This quote describes the state of CBMs, specifically in the area of middle school mathematics. The following chapter provides a review of the historical context of assessment process in school psychology, information about the history and foundation of the development of CBM, the empirical research pertaining to CBMs of mathematics at the middle school level.

Historical Context of Assessment in School Psychology

In order to fully comprehend the assessment process, we must visit the history of school psychology. The publication of the Binet-Simon scales is synonymously linked with the beginning of the field of school psychology. In 1905, psychologist, Alfred Binet, and psychiatrist, Theophile Simon, were commissioned by the Minister of Public Education in Paris to develop a way to classify and sort children who were predicted to be unsuccessful in the general education classroom and could not benefit from the regular curriculum (Merrell, Ervin, and Gimpel, 2006). Binet and Simon developed the first modern intelligence test that was later adapted to an English version known as the Stanford-Binet Intelligence Scales. These events caused early school psychology history to be forever linked to intelligence testing, individual assessment, and the classification of students.

The field of school psychology, as well as the field of education as a whole, has continued to develop exponentially since the early 1900s. The development of education and school psychology specifically, can be accredited to the passing of new laws and major court decisions. Some of the most influential laws were Public Law 94-142 in 1975, the Education for

All Handicapped Children Act, and the individuals with Disabilities Education Act (IDEA 1990, 1997, and 2004) (Merrell et al., 2006). These laws were very important in the development of school psychology and drove the specific need for school psychologists as assessment specialists in the schools.

The Education for All Handicapped Children Act was a unified federal law mandating free and appropriate public education for all students with disabilities (Merrell et al., 2006). This mandate expanded the need for school psychologists because of the need to assess students for special education eligibility; however, this educational act also stymied the development of school psychology because it kept many school psychologists in psychometrician/sorter roles (Merrell, et al., 2006).

In recent years, there have been criticisms of traditional assessment (Foegen & Deno, 2001). Many school psychologists have been trying to move away from the psychometrician/sorter role; however, assessment continues to play a very important role in the provision of comprehensive services. Assessment is necessary to ensure that children's difficulties are adequately identified and that interventions are appropriate for the problem a child is experiencing and in the context in which the problems occur (Merrell et al., 2006).

Traditionally, school psychologists have used standardized tests, such as the Stanford-Binet Intelligence Scales-Fifth Edition and the Woodcock-Johnson Tests of Achievement-III to look at any discrepancies between cognitive ability and academic achievement scores in order to determine eligibility for services. While many times this is a necessary step for a child to become eligible for services, this process has become a reactive process rather than a proactive process and has caused a backlash toward traditional assessment from both school psychologists and others in the educational community (Merrell et al., 2006).

Merrell et al. (2006) stated that assessment should not be viewed as limited to determining eligibility for special education; rather, assessment should be viewed as an important part of the problem-solving process. This idea was expanded further by Howell and Nolet (2000), who differentiated between assessments done for purposes outside the classroom (traditional assessments for meeting criteria for a disability and selected services) and assessments conducted inside the classroom (to obtain data to identify students who lack specific skills, inform intervention decisions, and monitor the effects of the implemented interventions). While traditional assessments are needed and provide us with relatively good information about the child/student, they are less useful in monitoring and evaluating the effects of instruction and interventions on an ongoing basis.

In the past, education was provided to the very talented or the very privileged/wealthy; however, in the early 1900s, school became compulsory for all children in the United States (National Conference of State Legislatures, 2010). According to Deno (2002), the purpose of school is to foster the cognitive, affective, social, and physical developmental outcomes of students. Because school attendance is not a "natural" course of development, Deno (2002) has argued that society has said school itself is a universal intervention for all children. Schools have developed a general idea of what skills/education a student should be able to demonstrate by the time they have reached graduation. During the course of the educational process, schools have also established ways of indicating progress (i.e., benchmarks) of sequential competencies required to meet graduation criteria. However, Deno indicates the movement along this linear educational goal and mastery of skills and outcomes does not occur at the same rate for all students or occur evenly across different educational domains for each student (i.e., mathematics, reading, spelling, etc.). While we know education does not occur in a linear progression, traditional assessments to indicate progress are assessed linearly and against standardized norms.

History and Foundation of Curriculum-Based Measurement

In the late 1970s, Stan Deno and his colleges at the University of Minnesota Institute for Research on Learning Disabilities developed curriculum-based measurement (CBM). CBM was designed to be an objective, ongoing measurement system of student outcomes that facilitated enhanced instructional planning (Deno, 1985). Because educational laws demand increased accountability and improved outcomes (Foegen, 2001), CBM allows data to be taken frequently to assess the students' progress. Further, through CBM, the effectiveness of instruction for that student can also be evaluated.

Some confusion has arisen about the differences between curriculum-based measurement (CBM) and curriculum-based assessment (CBA). CBMs can be thought of as one type of CBA (Hosp & Hosp, 2003). CBA is defined by three features: the test stimuli are taken from the curriculum the student is being taught, the student is tested repeatedly over time, and the information from the tests are used to inform instruction. Fuchs and Deno (1991) have identified two major measurement models: mastery measurement and global outcome measurement. Mastery measurement is defined as taking global skills and breaking them into a set of sub-skills. These sub-skills are then taught and measured in sequence, thereby identifying a child's short-term academic progress toward mastery of specific sub-skills (Fuchs & Deno, 1991). Most CBAs fall under mastery measurement because they usually include teacher-made tests with varying items to demonstrate the mastery of each sub-skill (i.e., teacher X makes a teacher-constructed test for 10th grade English students in Somewhere, WI). General outcome measurement is defined as using standardized procedures to assess overall proficiency in an academic domain in which the testing procedures remain constant over a longer period of time, typically one year (Fuchs & Deno, 1991).

CBM falls under the category of global outcome measurement because it uses repeated measures of the overall proficiency of global skills in an academic domain by using different but equivalent forms (Hosp & Hosp, 2003). While a need for both CBA and CBM exists, the reliability and validity of CBA are unknown because they are not standardized and individual teachers design curriculum-specific tests for a specific group of students. In contrast, CBMs are standardized, with guidelines and specific procedures for selecting test materials, which allows for the collection of reliable and valid data.

The use of CBA/CBM methods as alternatives to standardized, norm-referenced achievement testing has gained momentum over the past few years. As previously stated, CBM is a specific type of curriculum-based assessment, which involves a set of specific procedures that are reliable and valid, that includes standardized measurement tasks, and that have specific administration and scoring guidelines (Deno, 2003).

According to Deno (2003), one of the more common uses of CBM procedures is to evaluate students who are academically at risk. Because local norms can be developed with CBMs, children who perform toward the bottom of these local norms can be considered at-risk for academic difficulties compared to their local peers. CBMs not only help assess who is at-risk for academic difficulties in a specific area (i.e., mathematics, reading, etc.), but also can identify specific skill deficits. This is one of the biggest differences between CBM and traditional assessments: CBMs can be more closely linked with academic interventions. With the push of Response to Intervention (RTI), the usefulness and effectiveness of CBMs will only continue to rise.

One of the greatest benefits of using CBM is it creates a database not only for a particular student but also for a particular group (i.e., Mrs. X's 2nd hour math class) or local/district (i.e., District X 6th grade math). Individuals, then, can be compared across these groups because they take the same assessment (Fuchs & Fuchs, 1991). The resultant databases allow school personnel to graph and document concrete data about the child's progress. This information is very useful when discussing a student's progress with the student and/or the parent(s).

Hosp and Hosp (2003) reported that children who are aware of their own CBM data appear to be more knowledgeable about their own learning and see themselves more responsible for their own learning (Davis, Fuchs, Fuchs, & Whinnery, 1995). Additionally, educators can adjust what and how they teach to meet the specific needs of each student (Hosp & Hosp, 2003, p. 16). While this may sound as if the teacher will have to do more instruction, the teacher can actually provide better instruction because they know where to target their instruction. Targeting instruction can reduce having to spend time on instruction in areas the students have already mastered or to repeat instruction when unwarranted.

Fuchs, Butterworth, and Fuchs (1989) and Fuchs and Fuchs (1986) demonstrated that students of teachers who use CBM to inform instruction achieve increased test scores and subsequently higher grades than students of teachers who did not use CBMs. A final reason to use CBMs is they are easy to administer and require little time, which allows teachers to spend more time with instruction.

CBMs of Mathematics

Foegen (2008) reported early research in the investigation on CBMs of mathematics at the elementary level. Foegen focused on single operation basic fact tasks that included single-digit combinations for the four basic operations of addition, subtraction, multiplication, and division (Epstein, Polloway, & Patton, 1989; Tindal, Germann, & Deno, 1983) as well as mixed-operation facts measures that included single-digit combinations across all four operations (Espin, Deno, Maruyama, & Cohen, 1989; Tindal, Marston, & Deno, 1983). Fuchs, Hamlett, and Fuchs (1998, 1999), as cited by Foegen, Jiban, & Deno (2007), developed grade-level-specific measures for computation and then for mathematics concepts and applications.

While there is substantial research on the reliability, criterion validity, and usefulness in using CBMs to depict student growth at the elementary school level [(Fuchs et al., 1998, 1999, as cited by Foegen et al., (2007)], research in the area of CBMs in middle school mathematics is relatively limited. While there is some research on CBMs of mathematics at the elementary school level, only two groups of researches account for the research of curriculum-based measures at the middle school level. Foegen and her colleagues have researched the use of measures to assess and monitor student progress (basic facts and estimation), while Helwig and his colleges have conducted research using concept-based measure to predict performance on high-stakes achievement tests and to evaluation of adequate yearly progress (Foegen, 2008, p. 196). Because of the limited research in the area of CBM in middle school mathematics, both of Foegen's studies and Helwig's studies will be reviewed in this paper.

Research by Foegen and Colleagues

Foegen and Deno (2001) originally examined the possibility of identifying indicators of growth in middle school mathematics. Their strategy was similar to Deno's approach when he developed CBMs in the late 1970s and early 1980s, in which they attempted to identify simple, economical, efficient, and technically adequate performance indicators that teachers could repeatedly use to measure growth. Foegen and Deno examined two types of measures to identify growth indicators: estimation fluency and basic fact fluency. "We selected estimation as the primary focus of the study because it is representative of the type of mathematical skill that is widely applied by adults in daily living situations and thus is likely to represent a generate outcome of many middle school mathematic curricula" (Foegen & Deno, 2001, p. 5). Further, Foegen and Deno quoted Reys (1992, p. 281) in advocating for their choice of investigating estimation as a CBM measure: "estimation is a basic skill, and its growing importance in a technological society is recognized. It is used much more than exact computation

Fluency with facts was also selected by Foegen and Deno (2001) because of the importance of gaining automaticity in basic skills as a foundation for gaining higher levels of math competence. Foegen and Deno cited studies (e.g., LaBerge & Samuels, 1974) of automatic-decoding skills in reading to allow for increased attention on comprehension, which they compared to the automaticity of mathematic basic skills for greater information-processing capacity and more complex tasks in mathematics.

Foegen (2008) also conducted another study to assess progress monitoring in middle school mathematics. As with the previous study, basic facts and estimation measures were investigated; however, four other measures were also examined. These included the Monitoring Basic Skills Progress (MBSP) measures of computation (MBSP-Comp) and for concepts and applications (MBSP-ConApp). The MBSP was developed by sampling representative skills and concepts from the Tennessee state mathematics curriculum for Grades 1 through 6 (Fuchs et al., 1998, 1999, as cited by Foegen). According to Foegen, the MBSP's measures have had substantial research and support documenting their reliability, criterion validity, and predicting student growth at the elementary level as reported by Fuchs et al., (1994, 1998, 1999). While the MBSP has been supported at the elementary level, its usefulness with students at the middle school level was unknown. The other new measures investigated by Foegen were two variations in assessing numeracy concepts (i.e., complex quantity discrimination and missing number).

Measures.

As previously stated, estimation is one of the most valuable skills in mathematics. As cited by Foegen and Deno (2001), Reys (1992) suggested over 80% of all math applications call for estimation rather than exact computation. However, mathematic instruction in schools often focuses on computation (Foegen & Deno, 2001). In Foegen and Deno's study, 100 (12 of which were identified as having learning or behavioral disabilities as identified by state guidelines) urban students in the sixth, seventh, and eighth grade were given three 3-minute estimation measures. The first measure was the Basic Estimation Task (BET). Students selected the best estimate for the four different operations (i.e., addition, subtraction, multiplication, and division) in both computation tasks and word problems. These computation estimate measures included basic problems (one digit numbers), intermediate problems (combinations of two-digit and one or two-digit numbers), and advanced problems (combinations of three-digit and one or two-digit numbers). The word problem measure contained problems similar to the basic and intermediate computation problems with respects to the digit sizes; however, they contained general topics such as money and purchasing; sports, hobbies, and music; and school and work (Foegen & Deno, 2001). Students selected answers from three alternatives. These items included distracters to simulate common student errors such as incorrect operations, digits out of order, etc.

The other two estimation tasks provided by Foegen and Deno (2001) comprised of Modified Estimation Tasks (METs). An increased amount of word problems, an increased amount of division problems, the elimination of basic computation problems, and the inclusion of problem involving rational numbers were included in the METs. Foegen and Deno used identical problems for each MET and only differed in the response format. The MET-A provided exact numbers (i.e., 12, 354, 1262), while the MET-B responses differed from one another by powers of 10 (i.e., 30, 300, 3,000). As with the BET, students had three minutes to respond to 40 items by selecting the letter corresponding to the correct answer. These estimation tasks were also used in a later study conducted by Foegen (2008).

The final task in Foegen and Deno's (2001) study looked at the usefulness of simple basic facts. The Basic Math Operations Task (BMOT) was designed to look at students' accuracy and fluency in mental computation of whole-number facts (0-9) in the four basic operations. The BMOT probes consisted of 80 randomly ordered single digit operation problems, and students had one minute to respond to the probe. This task was also in included in the study conducted by Foegen in 2008.

The MBSP-Comp measure required student to compute addition, subtraction, multiplication, and division with whole number, fractions, and decimals (ex: 4 6/7 + 8 3/5 =? and .783/2.1 =?). Students had six minutes to respond to 25 items, and the measures were scored by the digits correct in their answer. The MBSP-ConApp measure asked questions about concepts and applications related to numeration, applied computation, measurement, geometry, percentages, charts and graphs, word problems, ratios, and probability, proportions, and variables. Students were given seven minutes to respond to 24 items and needed to provide the correct answer for scoring purposes.

In 2008, Foegen also introduced looking at two new measures to investigate indication of overall math proficiency. "The Complex Quantity Discrimination task required students to analyze pairs of quantities and write the appropriate symbol (> , < , and =) in a box between the quantities (Foegen, 2008, p. 198). Foegen used these items to reflect number concepts and properties such as place value, associativity (an expression in which the order of the operations performed does not affect the final result as long as the order of terms is not changed); commutativity (an expression in which the order of terms does not affect the final result), and conceptual understanding of the four basic operations. The other measure used by Foegen was the Missing Number task in which students were presented a series of three numbers and indicated the missing number in the sequence. Both of the new measures were given for one minute and contained 44 items.

Reliability.

The Foegen and Deno (2001) study examined the reliability of each of the measures and explored the degree to which students' scores on the measures were related to other indicators of math proficiency (i.e., grades, teacher ratings, and standardized test score). The last measure also examined to determine the predicted growth of students. The reliability of the measures in this study was relatively high, as the BET and METs' internal consistency coefficients ranged from .77 to .93; test-retest reliability coefficients ranged from .67 to .88; and the parallel forms coefficient ranged from .67 to .86. The reliability of the BMOT was also relatively high, as the coefficients for the BMOT were .91 to .92 (internal consistency); .80 to .85 (test-retest); .79 to .82 (parallel forms).

The results of Foegen's 2008 study showed alternate-form reliability coefficients exceeded .70 for four of the six measures (MBSP-ConApp, Basic Facts, Estimation, Complex Quantity Discrimination). Excluding Estimation in Grade 7, alternate-form reliability exceeded .80 in the winter and spring administration periods. Foegen noted, "as students become more familiar with the task, their scores become more consistent across forms" (p. 201).

Foegen (2008) also investigated the test-retest reliability for the six measures. With the exception of MBSP-Comp at the 7th grade level, which was significantly lower at .55, the test-retest reliability coefficients were between .75 and .95. Additionally, in Foegen's 2008 study, the test-retest reliability for Estimation for Grade 7 was below the acceptable .80 level, which was inconsistent with previous research.

Criterion-Related Validity.

Foegen and Deno (2001) also selected four criterion variables for their measures to be compared against. The four measures were math grade point average, overall grade point average, standardized test scores [California Achievement Test (CAT)], and teacher ratings. Foegen and Deno selected these measures because they are the measures typically used in making decisions concerning a student's academic progress. Foegen (2008) also selected teacher ratings of students' overall math proficiency, as well as standardized test scores [Iowa Tests of Basic Skills (ITBS) and the Northwest Achievement Levels Test (NALT)] as criterion variables. Student mathematics GPA was also collected for the 2008 study; however, this data was only available for 8th grade students in one district.

The criterion validity, which looked at the how the measures compared to other indicators of math proficiency (i.e., grades, CAT scores, and teacher ratings) for the BET and METs ranged from .29 to .56, while the BMOT criterion validity ranged from .33 to .63. Most of the criterion validity scores were in the .40 to .50 range, which reflected a moderate relationship.

The strongest criterion-related validity coefficients were obtained for the MBSP-ConApp and ranged from .58 to .87, which far exceeded the criterion related validity evidence found for the other measures. Foegen also found the MBSP-Comp and Estimation had relatively strong coefficients at Grade 6; however, the MBSP-Comp was among the weakest at Grade 7 as well as estimation being weaker at Grades 7 and 8. Foegen additionally found the Complex Quantity Discrimination task and Missing Number tasks had weaker validity coefficients at grade 6 but were stronger at grades 7 and 8.

Growth.

Foegen also looked at changes in student performance levels over time. Like the criterion validity comparisons, the MBSP-ConApp produced the largest mean changes from fall to spring while the MBSP-Comp produced the smallest. According to Foegen (2008), "Among students in Grades 6 and 7, the Complex Quantity Discrimination and Missing Number measures produced the next largest (change). In Grade 8, Complex Quantity Discrimination produced the largest (change), followed by Estimation and Missing Number" (p. 203).

Weekly rates of growth on the estimation measures were examined for the BET and METs. Students increased .25 points/week on the estimation measure during a 10-week period. Finally, the students' growth on the BMOT was more than double that of the BET and METs, as the mean growth rate was .55 problems per week. "The results of this work suggested that both measures have acceptable levels of reliability and validity, but the (BMOT) measure is more likely to be sensitive to small changes in student performance" (Foegen, 2008, p. 196).

Remarks by Foegen and Colleagues

Foegen (2008) indicated the MBSP-Comp and MBSP-ConApp were included in the study because they are both widely practiced at secondary levels despite being designed to reflect elementary content. The MBSP-Comp produced the best results across reliability, criterion validity, and growth at Grades 6 and 7, but this measure was not given to Grade 8 students. While the MBSP-Comp had strong criterion validity and acceptable reliability at Grade 6, the MBSP-Comp was considerably lower at Grade 7. The Basic Facts and Estimation measures produced similar results to previous studies in reliability, criterion validity, and changes across time for Grade 6. The reliability and criterion validity coefficients in the 2008 study were lower than Foegen and Deno's (2001) results. The Complex Quantity Discrimination and Missing Number tasks were similar to the MBSP measures as they both produced different results from each other. While the Complex Quantity Discrimination produced acceptable to strong levels of reliability, criterion validity, and growth, most of the Missing Number indicators were among the lowest of the six different measures. Foegen indicated this result is surprising because one would expect this task to be strongly associated with math proficiency; however, "One possible explanation for these results is that the 1-minute duration of the task was insufficient to effectively discriminate among students of varying ability levels" (p. 205).

While better results are always desirable and Foegen and Deno (2001) and Foegen (2008) would agree, the results of the studies were consistent with prior research. Foegen and Deno indicated previous research (Espin & Deno, 1993; Espin & Foegen, 1996; Jenkins & Jewell, 1993) has indicated a decreasing relationship between all CBMs and external criteria as students progress past elementary school. Additionally, Marston (1989) indicated while there is good criterion validity for CBMs of mathematics at the elementary level, they are still lower than the criterion related validity coefficients of reading CBMs. Foegen and Deno (2001) and Foegen (2008) said while their studies only provided moderate relationships, they are consistent with previous research as well as the criterion-related validity coefficients of commercial achievement tests of mathematics; as such, they concluded the measures are promising indicators of math proficiency.

Research by Helwig and Colleagues

Robert Helwig, Lisbeth Anderson, and Gerald Tindal (2002) conducted the other major research in the area of CBMs in the area of middle school mathematics. The study by Helwig et al. focused on CBM tasks measuring conceptual knowledge rather than procedural knowledge. Helwig et al. also researched how well their measures predicted statewide test scores for middle school students, particularly for students with identified disabilities. They acknowledged the need for students to have knowledge of basic computation skills; however, they also asserted state standards require students to have conceptual knowledge beyond what can be measured by computation tasks.

Helwig et al. (2002) described two possible reasons why CBM tasks that measure conceptual knowledge may be stronger predictors of scores on achievement tests than procedural knowledge. The first reason is that conceptual understanding is a prerequisite for the successful application of mathematical procedures in problem solving situations. Carpenter (1986), as cited in Helwig et al. (2002), stated 70% of students ages 13-17 taking the National Assessment of Educational Progress could successfully multiply two common fractions, but only 20% could solve a one-step word problem involving multiplication of similar fractions. These results indicate while many students know how to do specific mathematical procedures; many students have difficulty in applying their procedural knowledge to solve application problems. A unit of conceptual knowledge cannot be thought of as an isolated piece of information because conceptual knowledge is only useful when the individual recognizes the relationship between different pieces of information. As, Greeno (1991), as cited in Helwig et al. (2002), said, “Knowing various domains (i.e., algebra, probability, computation, fractions, measurement), involves distinguishing the conceptual and theoretical entities and phenomena in the domain and knowing the principles of identity, invariance, composition, transformation, and causality that can be used to explain phenomena and events" (p. 174). Mathematical achievement at the middle school level requires students to understand interrelated concepts rather than applying a one or two step procedure commonly taught at the elementary level. Because many mathematical domains are related and overlap, conceptual knowledge in one domain has a direct effect on conceptual knowledge in other domains and subsequently on overall mathematical achievement.

It is widely recognized in reading that decoding automaticity frees cognitive processes to concentrate on comprehension rather than word recognition, and the strong correlation between oral reading fluency rates and reading comprehension is likely influenced by this relationship. However, Helwig et al. (2002) states not all academic skills and CBMs are related to fluency. The tasks measured by Foegen and Deno (2001) and Foegen (2008) contain measures of fluency; however, Helwig et al. would argue conceptual knowledge and understanding cannot be memorized and subsequently CBM in secondary mathematics should not be fluency based.

Participants.

Helwig et al. (2002) study consisted of eight school districts from a western state to help develop and test different CBM measures. Students in Grades 3, 5, 8, and 9 were given measures in reading, writing, and mathematics. Only the results of 8th grade mathematics were reported in their 2002 study. The study included 199 eighth-grade students from moderately-sized schools to rural area schools. The participants were predominately white (77 to 90%) with students with minority status, and the socioeconomic status of the schools ranged from moderately low to high. Forty-seven percent of the students had an individualized education plan (IEP) in the area of mathematics, and 52% of the participants were boys and 48% were girls.

Measures.

Helwig et al. (2002) conducted a series of pilot studies to develop items for CBM in middle school mathematics that might be predictive of standardized math achievement scores. During these pilot studies, they selected 11 items that met the following criteria:

1. Each item had to differentiate between high and low achievers on some measure of general math achievement during field testing.
2. When analyzed with several other items, each item had to explain a significant portion of the variance in math achievement.
3. Each item had to be conducive to the creation of alternate forms.
4. Each item had to test a concept rather than a procedure.

Helwig et al. did not consider the resultant 11 items a representative sample of 8th grade subject matter curriculum because they wanted the items to assess conceptual understanding rather than memorization. Students were given a set of 48 items. The items included the 11 CBM items previously described, and students were not timed; however, most students completed the items in less than 35 minutes. On other field tests, which included only the 11 concept CBM items, showed students took between 10 to 15 minutes to complete the items.

Helwig et al. (2002) stated their participants were given a standardized achievement test; however, because the results of the tests were not available to them at the time they published their study, the students were given a custom computer adaptive test (CAT). The computer software was programmed to select items based on students' responses to the previous field tests. Students who answered items correctly were given progressively harder questions, while students who incorrectly answered items were given progressively less difficult questions. "Through this method of targeting each individual student's specific skill level, the CAT calculated a score with a similar standard error as the official state test using less than half the number of items" (p. 106).

Results

Helwig et al. (2002) used Pearson product-moment correlations to compare students' total number correct on the CBM math concept task and their corresponding CAT scores. The performance of general education students was significantly higher on both the CBM items and the CAT. General education students answered approximately half (5.57) of the CBM items correctly, while students with IEPs answered 1.77 items correctly (Helwig et al., 107, 2002). The Pearson product-moment correlation revealed a strong relationship between the 11-item CBM and the CAT scores ($r = .83$). The correlation coefficient was strong for general education students ($r = .80$), but the correlation was lower for students with IEPs ($r = .61$).

Helwig et al. (2002) additionally determined how much variance in the CAT scores could be explained by the CBM measures by using regression analyses. The results of simple regression using total CBM scores (0-11) showed the CBM items accounted for about two-thirds of the variance in the CAT for the entire population. Data from the step-wise regression analyses were consistent with the simple regression results.

Finally, Helwig et al. (2002) also used a discriminant function analysis (DFA) to investigate the effectiveness of using the CBM items and a student's status (general education vs. IEP) to predict success on the CAT. The application of the DFA resulted in an overall agreement between a student's actual and predicted status of 87.1%. 97 students did not meet the state benchmark standards, and the DFA accurately predicted 88 cases or 90.7% of the sample. Over 82% of the 74 students who did reach the state benchmark standards were accurately predicted by the measure. An additional DFA was conducted using only the concept CBM items as a predictive variable, and the DFA was shown as having a correct prediction rate of 81.3%.

Helwig et al. (2002) found the relationship between their CBM concept task and general math achievement was not as strong for students with learning disabilities as for students in general education. This result was inferred to be du to the difficulty of the items. The researchers hypothesized the correlation would be higher for the students with disabilities if they selected items targeted for lower levels of skill. Because of the poor performance of students' with disabilities on their measure, Helwig et al. said their instrument may not have been sensitive enough to measure progress over the course of a semester. Helwig et al. stated this was not a primary concern for them because of two reasons: 1) "the advantage of the specificity of information gained from a traditional math CBM over (theirs) is limited to the relatively small domain of computation, which becomes less important in favor of application and conceptual knowledge, and 2) as GOMs, their tasks were designed to assess student proficiency on the global outcomes toward which the entire curriculum is directed" (Helwig et al., 2002, p. 110). As with Foegen's studies, Helwig et al. found potential in their study, but they also were hesitant to make generalizations and made a strong argument for further research.

In conclusion, the assessment process has come a long way since the early days of Alfred Binet trying to sort children he thought would not be successful in the general education classroom. The role and need of the school psychologist has developed throughout the years, in part to the passing of federal legislation, such as IDEA. Traditionally, the school psychologist has been a psychometrician/sorter and utilized standardized assessments, such as the WISC-IV or WJ-III; however, these assessments have recently come under some criticism in their use as assessment tools for at-risk students. An alternative assessment system, developed by Stan

Deno, is CBM. CBMs were designed to be an objective, ongoing measurement system of student outcomes, which facilitates enhanced instructional planning. Most research on CBM has been conducted in the area of reading, but more and more research has been conducted in other academic areas such as mathematics. Research in the particular area of CBM in middle school mathematics is very limited and has primarily been conducted by two groups of researchers.

Foegen and her colleagues examined identifying indicators of growth in middle school mathematics for formal evaluation much like CBM. These indicators included: Estimation; Basic Facts; MBSP-Comp; MBSP-ConApp, Complex Quantity Discrimination; and Missing Number tasks. The MBSP-ConApp tasks were found to have the criterion validity. Helwig and his colleagues focused on CBM tasks to measure conceptual knowledge rather than procedural knowledge and researched how their measures could predict statewide test scores for middle school students, particularly those students with identified disabilities. Helwig et al. found the relationship between their CBM concept task and general math achievement was adequate for students in general education as well as predicting standard achievement test scores; however, this relationship was not as strong for students with learning disabilities. While there is great potential for CBMs of middle school mathematics, more research needs to be conducted.

Chapter Three: Summary and Discussion

The research among CBMs and progress monitoring for middle school mathematics is not extensive enough for any researcher to make strong recommendations for any particular measure. Research conducted by Foegen and Deno (2001) and Foegen (2008) show adequate technical adequacy for certain measures and their use as progress indicators. Of the six measures discussed, four (MBSP-ConApp, Basic Facts, Estimation, Complex Quantity Discrimination) consistently produced reliability coefficients that exceeded .70 across grade levels. The only exception to this finding was for the Estimation task for Grade 7, which found inconsistent reliability levels between the two Foegen studies. Foegen reported additional refinements to the Estimation measure may be needed to improve the reliability to a level necessary for implementation in educational practice. Helwig et al. (2002) did not report any reliability for their conceptual measures, which causes some concern for their use as CBMs.

The strongest relationships across grades and the criterion variables were found with the MBSP-ConApp task, which was only investigated for students in Grades 6 and 7 (Foegen, 2008). These correlation coefficients (i.e., .60 to .87) were significantly higher than any other measures. The MBSP-Comp and Estimation tasks were the next strongest measures for Grade 6, but they produced the weakest criterion validity coefficient for Grade 7 and 8 (Estimation only).

Helwig et al. (2002) found a strong relationship ($r = .83$) between their CBM items and a standardized achievement test. Helwig et al. also analyzed their data to determine predictive validity for success on state standardized achievement scores. They found their measures to be predictive of state test scores (i.e., .49 for Wilk's lambda and .84 for standardized discriminant function coefficient). Although both studies looked at their measures for use as progress monitors, little can be inferred from their findings as data was only collected a few times throughout the academic year. Additional studies need to be conducted to confirm the technical adequacy of the Foegen and Helwig measures, and additional research needs to be conducted on their specific use for monitoring the progress middle school students.

Limitations

As with all studies, there are limitations, which have an impact not only on the particular study, but also about the conclusions and implications for the field. The first limitation of this research study is it is only a literature review. As such, it is not add any research to the field of CBM or CBMs of middle school mathematics. The biggest limitation of this literature review is that there has been limited research investigating the use of CBMs beyond the elementary level, especially in the area of mathematics.

There are also several limitations of the particular studies that were cited in this review. These limitations have an impact on the studies as well as the field of CBM in middle school mathematics. The first and probably biggest limitation of Foegen and Deno's (2001) study is the technical aspects of the study. The study was conducted with 100 participants with varied ethnic and educational diversity (i.e., students with disabilities); however, their sample size was not large enough to generalize results about CBMs of middle school mathematics to other populations. Additionally, the Foegen study was only conducted in one school district in the Midwest.

Other limitations of the Foegen and Deno study were the format of the probes and the administration procedures. While there was high reliability and moderate validity coefficients found in Foegen and Deno's study, the impact of changes in the format of probes, ways of responding, and timing are unknown. Foegen and Deno acknowledged this limitation and say the number of samples needed to determine generalized reliability is unknown. In order to make definite conclusions and generalize results, Foegen and Deno would have needed to have more participants representing various parts of the country.

Foegen's 2008 study had 563 participants and included ethnically and educationally diverse students. However, as with the 2001 study, the 2008 study was conducted in the Midwest. As such, the results may not generalize to other student populations around the country. Foegen also acknowledges that the district's curriculum and emphasis on different content from grade to grade may have had an effect on the results. Further, while the best results were obtained using the MBSP-ConApp measure, Foegen did not administer this measure to Grade 8 students. While one would expect the findings related to this measure to be similar for Grade 8 students, educators cannot be sure because it was not included in the study. The other major limitation of the study was the data for the study was only collected at three points during the school year. As Foegen (2008) stated, "Although we used these data to generalize estimates of weekly growth, it is possible that the growth rates obtained with more frequent data collection would differ from those reported in this study" (p. 205).

As with Foegen's studies, Helwig et al. (2002) study also had limitations. The biggest limitation was choosing untimed measures. Using untimed measures raises questions about the reliability and feasibility of the assessments for progress monitoring. The other limitations are similar to Foegen's studies, as the study was conducted with a relatively small sample size of primarily white students in one geographical area. As previously stated, Helwig et al. acknowledged that their measures may not be sensitive enough to progress of a student over a semester.

Implications for Future Research

The area of CBMs and progress monitoring measures for the middle school level, particularly in mathematics, is a relatively uncharted field. There is a high need for more research in this area. As indicated in the limitations section, there is a need for more studies conducted with larger and more diverse populations as well as researching these measures to determine their technical adequacy for frequent progress monitoring. More evidence is needed to support progress monitoring in verifying the changes in student performance. Without such research and data, instructors will continue to struggle with deciding when to change instruction and provide supplemental instruction.

While the Estimation task produced good reliability coefficients (Foegen and Deno, 2001), the results of Foegen (2008) were inconsistent. Because of this inconsistency, more research is needed to refine the measure and improve its reliability. Additionally, Foegen (2008) indicates future research should examine the technical qualities (i.e. reliability, criterion validity, growth, etc.) of the previously discussed measures for students who are at-risk and/or those individuals with identified disabilities to determine if the findings are similar for those populations. Further, additional research should be conducted with the MBSP-ConApp at the Grade 8 level to determine if its technical adequacy is the same as was found for Grades 6 and 7. More research also needs to be conducted to investigate the effects of adding additional time to the Missing Number task to determine if this changes its technical adequacy.

Helwig et al. (2002) study showed great promise for both predicting academic placement and predicting achievement on standardized tests; however, identifying these measures as CBMs can be disputed. Helwig et al. did not report any reliability coefficients for their measures, which is one of the foundational requirements of CBM (Deno, 2003). Because Helwig et al. chose not to implement time constraints to their measures, the reliability of the measures has to be questioned. Additional research needs to be conducted to determine the reliability of Helwig et al.'s CBM task items.

Implications for Practice

The main function of CBMs are to provide teachers with a tool and subsequent information from that tool to help improve student performance (Stecker, Fuchs, & Fuchs, 2005); moreover, "federal legislation, such as the No Child Left Behind Act of 2001 and IDEA of 2004, call for assessment tools like CBMs to quickly and accurately screen for students who appear at risk for failing to meet academic standards" (Diercks-Gransee, Weissenburger, Johnson, & Christensen, 2008, p. 361). Based on the current research, the measures previously discussed would be best served for screening students for academic risk in the area of middle school mathematics. While Foegen and Deno (2001), Foegen (2008), and Helwig et al. (2002) all inferred the usefulness of their CBM measures for monitoring the progress of students, their studies did not thoroughly research their adequacy for monitoring progress purposes.

When selecting a measure, Foegen, Jiban, and Deno (2007) indicated instructors must consider the advantages and limitations as well as the degree to which the measure represents the instructional curriculum in the particular school/district. If a district emphasizes the acquisition of conceptual mathematical knowledge, computation measures may not be sufficient as general outcome measures. Furthermore, when selecting a measure to screen for academic risk in the area of math at the middle school level, educators should also consider the specific grade level of the student or students as different measures may be better indicators at different grade levels.

While not enough research has been conducted to make a strong recommendation for universal practice, the measures previously discussed show promise as screeners. Further research is needed to determine their adequacy for monitoring progress. Sixth grade mathematics teachers should consider using measures such as the MBSP-Comp, Basic Facts, or Estimation tasks, as they were shown to be the most reliable and valid for this grade level. These educators should also consider using the MBSP-ConApp task, as it also was shown to have technical adequacy for this grade level.

Seventh and eighth grade mathematics teachers should consider using more conceptual measures, such as the MBSP-ConApp, as they showed the greatest level of empirical support and greatest potential for Grade 7. While educators cannot make assumptions about the MBSP-ConApp for Grade 8 students, we can speculate that measures that assess conceptual knowledge, as shown by Helwig et al. (2002), will do well in predicting Grade 8 students' achievement in math.

One possible explanation of why procedural measures more were most reliable and valid for sixth grade mathematics is that the sixth grade mathematics curriculum builds on students' foundational math skills. As possible explanation as to why the Grade 7 and 8 criterion measures were more correlated with conceptual measures is because the curriculum at these grade levels requires students to demonstrate proficiencies beyond simple procedural knowledge.

Summary

Testing results indicates students in the United States trail their peers around the world in the areas of science and mathematics. Educators and lawmakers have called for reform not only in the instruction of mathematics but also for the periodic screening and monitoring of students who are not acquiring proficiency in math. This research looked at (a) the historical context of the assessment process in school psychology, (b) the state and need for mathematic skills in middle school, and (c) CBMs of middle school math to assess, inform instruction, and monitor student progress. Although there is limited research in the area of CBM of mathematics at the middle school level, two groups of researchers account for all the research in the area of CBMs of middle school mathematics. Preliminary results indicate the MBSP-ConApp and MBSP-Comp show the most promise for sixth grade students, whereas the MBSP-ConApp and Basic Facts tasks showed the most promise for seventh grade students. Based on the research, Helwig et al.'s task showed the most promise for screening eighth grade students. However, no one measure has been confirmed as having the technical adequacy needed to monitor the developing mathematical skills of students at the middle school level.

References

Alexander, L. (1991). AMERICA 2000: An educational strategy. Retrieved on September 10, 2010, from http://books.google.com/books?id=3fpFSuBAbSkC&printsec=frontcover&dq=alexander+1993+goals+2000&source=bl&ots=6l8bcwCumJ&sig=hyBe3lXxnBjXEmH4wO2SngLFNK0&hl=en&ei=CgcJTf39B8epnQfchKmCAQ&sa=X&oi=book_result&ct=result&resnum=1&ved=0CBgQ6AEwAA#v=onepage&q&f=false.

Bryan, J. B. (2005). Fostering educational resilience and achievement in urban schools through school-family-community partnerships. *Professional School of Counseling, 8*, 219-227.

Constantino, S. M. (2007). Families welcome: Tips for moving parents to the secondary school. *Principal Leadership*, *7,* 34-39.

Cooper, K. J. *'89 Educational summit's goals still unmet.* The Washington Post. December 10, 1999.

Davis, L. B., Fuchs, L. S., Fuchs, D., & Whinnery, K. (1995). Will CBM help me learn? Students' perception of the benefits of curriculum-based measurement. *Education and Treatment of Children, 18,* 19-32.

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

Deno, S. L. (2002). Problem-solving as "best practice." In A. Thomas & J. Grimes (Eds.) *Best practices in school psychology: IV* (Vol. 1, pp.37-55). Bethesda, MD: National Association of School Psychologists.

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

Diercks-Gransee, B., Weissenburger, J. W., Johnson, C. L., Christensen, P. (2008). Curriculum-based measures of writing for high school students. *Remedial and Special Education, 30* (6), 360-371.

Elley, W. B. (1988). New vocabulary: How do children learn new words? *Research Information for Teachers*, Item 10 No.1.

Foegen, A. (2008). Progress monitoring in middle school mathematics: Options and issues. *Remedial and Special Education*, *29*(4), 195-207.

Foegen, A. & Deno, S. L. (2001). Identifying growth indicators for low-achieving students in middle school mathematics. *The Journal of Special Education, 35*, 4-16.

Foegen, A., Jiban, C., & Deno, S. L. (2007). Progress monitoring measures in mathematics: A review of literature. *The Journal of Special Education*, *41*(2), 121-139.

Fuchs, L. S., Butterworth, J. R., & Fuchs, D. (1989). Effects of ongoing curriculum-based measurement on student awareness of goals and progress. *Education and Treatment of Children*, *12,* 63-72.

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

Fuchs, L. S., & Fuchs, D. (1991). Curriculum-based measurements: Current applications and future directions. *Preventing School Failure, 35*(3), 6-11.

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

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

Hosp, M. K., Hosp, J. L., & Howell, K. W. (2007). The ABCs of CBM: A practical guide to curriculum-based measurement. New York, NY: The Guilford Press.

Hosp, M. K. & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure,* 48, 10-17.

Howell, K. W., & Nolet, V. (2000). *Curriculum-based evaluation: Teaching and decision making* (3rd ed.). Belmont, CA: Wadsworth.

Greeno, J. G. (1991). Number sense as situated knowing in a conceptual domain. *Journal of Research in Mathematics Education, 22*, 170-218.

Glod, M. (2007). *U.S. Teens Trail Peers Around World on Math-Science Test.* The Washington Post. Retrieved on August 26, 2010, from http://www.washingtonpost.com/wp-dyn/content/article/2007/12/04/AR2007120400730.html

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293-323.

Merrell, K. W., Ervin, R. A., & Gimpel, G. A. (2006). *School psychology for the 21st century.* New York, NY: The Guilford Press.

N.A. (2009). CNN politics Obama pushes math, science education. *CNN.com-Breaking News, U.S., World, Weather, Entertainment & Video News.* Retrieved on August 26, 2010, from http://www.cnn.com/2009/POLITICS/11/23/obama.science/index.html?iref=allsearch

National Conference of State Legislatures. (2010). *Compulsory Education.* Retrieved on December 14, 2010, from http://www.ncsl.org/default.aspx?tabid=12943