2014

UNDERSTANDING TEACHERS' CONCURRENT KNOWLEDGE OF ASSESSMENT LITERACY AND CURRICULUM-BASED MEASUREMENT

Paige Hamilton
University of Rhode Island, phamilton@my.uri.edu

UNDERSTANDING TEACHERS' CONCURRENT KNOWLEDGE OF ASSESSMENT LITERACY AND CURRICULUM-BASED MEASUREMENT

BY

PAIGE HAMILTON

A THESIS SUBMITTED IN PARTIAL FULLFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

IN

PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND

2014

MASTER OF PSYCHOLOGY THESIS

OF

PAIGE HAMILTON

APPROVED:

Thesis Committee:

Major Professor Gary Stoner

Susan Loftus-Rattan

Julie Coiro

Nasser H. Zawia

DEAN OF GRADUATE SCHOOL

UNIVERSITY OF RHODE ISLAND

2014

ABSTRACT

As a result of the standards-based movement in education it is important for teachers to be knowledgeable about and understand the use of assessment in the general education setting. Prior studies have investigated teachers' understanding of sound assessment known as assessment literacy. This study explores teachers' concurrent knowledge of Curriculum-Based Measurement and sound assessment, as measured by the Assessment Literacy Inventory. More specifically, the extent of the relationship is examined by the Pearson product-moment correlation coefficient. Additionally, the relationship between training and knowledge about CBM is described.

Results of the current study revealed a moderate, positive relationship between scores on measures of knowledge about CBM and assessment literacy. It is unclear if training is related to knowledge about CBM or assessment literacy. Implications of these results and the conceptualization of assessment literacy, as well as limitations of the study, are discussed.

ACKNOWLEDGMENTS

Thank you to my major professor, Dr. Gary Stoner, for his feedback and support. I am grateful for the guidance he has provided me throughout the study. Thank you to my committee members Dr. Susan Loftus-Rattan and Dr. Julie Coiro for their thoughtful comments and time for discussions. I would also like to express my appreciation to the school professionals who participated in the validation process. Finally, thank you to my family and friends who have supported me throughout this process.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

FIGURE PAGE

CHAPTER I

INTRODUCTION

Statement of the Problem

The current study explores to what extent teacher assessment literacy, as measured by the ALI (Mertler & Campell, 2005), is related to teachers' knowledge about CBM. The adoption of CBM as a tool for assessment in the general education classroom has increased since the promotion of accountability and evidence-based practices by the Individuals with Disabilities Education Improvement Act of 2004 (IDEA) and the No Child Left Behind Act of 2002 (No Child Left Behind [NCLB], 2002; Individuals with Disabilities Education Improvement Act, 2004). Curriculum-Based Measurement was originally developed for use by special education teachers and specialists, because of their background training in educational measurement, for assessing students' basic skills (Deno. 1985). Now, general education teachers are expected to utilize these measures to make instructional decisions.

In schools, testing and assessment are common practice. The dissemination of the 1990 Standards for Teacher Competence in Educational Assessment of Students incited interest in examining teachers' knowledge and understanding of assessment, and in 1991, Richard Stiggins coined the term 'assessment literacy' (AFT, NCME, NEA, 1990; Stiggins, 1991). According to Stiggins (1991), an individual who is literate in assessment activities has the knowledge and skills to administer, score, and interpret assessments with high quality. In general, the literature has indicated low

levels of assessment literacy and gaps in both inservice and preservice teachers' knowledge (Plake, Impara, & Fager, 1993; Mertler, 2003; Volante & Fazio, 2007).

Regardless of the empirical evidence supporting the reliability and validity of CBM (Good & Jefferson, 1998) and the positive effects on student achievement when the measures inform data-based decision making (Stecker, Fuchs, & Fuchs, 2005; Tindal 2013), the implementation of these measures by teachers in schools varies (Ysseldyke & Bolt, 2007; Bolt, Ysseldyke, & Patterson, 2010). Recent research has indicated that variability of the implementation and interpretation of CBM is not related to teacher characteristics including teaching experience, and years of experience with CBM (Wayman et al., 2011). One possible explanation for the variability of implementation is the function of different levels of knowledge and understanding about CBM.

In a data-driven system, where decisions about student progress and achievement are based on assessment results, it is important that teachers are assessment literate. To make valid decisions, teachers need to use assessments with fidelity, which means they need to be knowledgeable and confident interpreting results in addition to the implementation. The current study explores the relationship between an established framework for conducting sound assessment and more specific assessment information comprising CBM. The nature and specifics of the identified relationships will provide implications on how training for adoption and use of CBM could be more effective.

The following primary research question was examined:

1. How is teacher assessment literacy, as measured by the Assessment Literacy Inventory, related to teacher knowledge of Curriculum-Based Measurement?

The following secondary research question was examined:

1. To what extent is the amount of teacher training in assessment related to knowledge about Curriculum-Based Measurement?

REVIEW OF LITERATURE

Gap between Research and Practice

The goal of educational research is to influence practice by discovering and promoting the use of evidence-based practices. As educational research becomes more advanced, dissemination of effective practices into the field setting is integral. After a direct relationship between a program, or intervention, and positive student outcomes is established, then the degree to which a program generalizes to a typical school setting may be explored (Dzewaltowski, Glasgow, Klesges, Estabrooks, & Brock, 2004). However, it has been documented that educational research has been undervalued by teachers because of insufficient communication of how to replicate, in natural settings, those practices specified in journal articles (Greenwood, 2001). As a result, oftentimes research is not translated into practice as intended by the researchers (Greenwood & Abbott, 2001). A benefit of research conducted in the field setting is that it may elicit if and to what extent modifications to the implementation are necessary while preserving the outcomes (Shulte, Easton, & Parker, 2009). Potential consequences of a gap between research and practice may be the inaccurate use of or the absence of evidence-based practices in schools.

Evidence-Based Practices. Current school reform efforts focus on the use of evidence-based practices to improve the quality of core programs, targeted programs, instructional strategies, and professional development to enhance student academic achievement. National laws and policies, including the No Child Left Behind Act (NCLB, 2002) define evidence-based practice as "research that involves the application of rigorous, systematic, and objective procedures to obtain reliable and

valid knowledge relevant to education activities and programs" (NCLB, 2002, p. 540). Methods that are supported by scientifically based research may have a greater opportunity of impacting student learning and achievement because a relationship with positive student outcomes has been documented (Dzewaltowski, Glasgow, Klesges, Estabrooks, & Brock, 2004). Typically, evidence-based practices are manualized, or at the least have specific procedures detailing the implementation of the practice. To maintain the relationship between evidence-based practice and positive student outcomes, it is essential that the practice is implemented as intended by the researchers (Sanetti & Kratochwill, 2009). Such implementation is known as "treatment integrity".

Treatment integrity. Treatment integrity refers to the extent to and level of skill in which the procedures of an evidence-based practice are implemented (Shulte, Easton, & Parker, 2009). This is dependent on how well the implementers were trained and how well they understood the training. Currently, a universal research based method does not exist for documenting adherence to implementation procedures (Shulte, Easton, & Parker, 2009). Without this information it cannot be determined if student outcomes are a result of the evidence-based practice or how the evidence-based practice has had to be altered in order to obtain results (Sanetti & Kratochwill, 2009).

The intricacies of a school setting increase the risk of implementing a treatment inaccurately (McIntyre, Gresham, DiGennaro, & Reed, 2007). Teachers have numerous responsibilities in the classroom that may limit their attention to treatment integrity, including but not limited to setting demands, classroom management, behavior management, and limited resources. Drifting from a protocol

may interfere with outcomes if such drift results from factors unrelated to student needs (Sanetti & Kratochwill, 2009). However, some adaptations may be necessary to accommodate the context or individual. The American Psychological Association (2005) considers adaptation to be a natural part of implementing evidence-based practices in schools, if the modifications are determined by an individual with expertise in the area. Individuals who implement evidence-based practices should be knowledgeable about how to implement the practice and how to implement it with high quality.

Response to Intervention/Instruction

The primary issues noted, regarding a research to practice gap, the need for evidence-based practices in schools, and the importance of treatment integrity for obtaining expected results, are relevant to the contemporary area of educational practices known as Response-to-Intervention (RtI). In 2004, IDEA incorporated a new regulation in regard to identifying students with specific learning disabilities. The new regulation permits local education agencies to consider a student's response to scientifically-based intervention as a procedure for determining eligibility for a specific learning disability (P.L. No. 108-446 614 [b][6][A]; 614 [b][2&3]). In other words, the national policy encourages schools to use evidence-based practices for instruction and intervention. Further, documentation of how the practice is implemented (i.e. treatment integrity) and student progress during implementation is necessary to determine students' response to the treatment.

RtI is a data-driven prevention, intervention, and problem solving model (Burns & Gibbons, 2008). In essence, the goal of RtI is to increase the number of

students who are successful in the general education setting, while simultaneously decreasing the number of students referred to special education. One core feature of RtI is data-based decision making; at each stage of the problem-solving model data are used to determine students' response to instruction and/or intervention (Fuchs & Fuchs, 2006). Academic assessments are one method for collecting data. Ysseldyke, Burns, Scholin, and Parker (2010) review the characteristics of instructionally valid assessments within an RtI framework. First, assessments should be precisely matched to student individual needs. In other words, the evidence collected needs to provide information specific to the instructional goals or outcomes for each student. Second, all students' progress are monitored on a frequent basis, from three times a year for students attaining grade-level expectations, to two times a week for students receiving intensive intervention services. As such, multiple, equivalent forms are necessary. Finally, to perform the function of progress monitoring, or to determine students' responsiveness, it is essential that assessments are sensitive to change. Data collected by an assessment method that possesses these characteristics would support decisions made within an RtI framework.

Assessment

Assessments are measures used to determine what students know and are able to do before, during, and after instruction (Green & Johnson, 2010). According to Pellegrino, Chudowsky, and Glaser (2002-3), first, information is gathered and second, information is used to make inferences about students' knowledge and understanding. An example of one type of assessment teacher's use is an end-of-unit

exam, or a chapter test. The purpose of making judgments about student learning is to enhance future outcomes.

Research has indicated that teachers may spend a large portion of their instructional time engaged in assessment-related activities (Stiggins 1991). Given the amount of time spent on assessing students and the required use of assessments within an RtI framework, it is important for teachers to be aware of and understand why they are assessing, how the assessment is administered and scored, how to interpret the results, and how to use the results for decision making. Under the NCLB act, and for the purpose of data-based decision making within an RtI framework, a well-supported, evidence-based approach to assessment is warranted. The extent to which teachers are knowledgeable about assessments has been discussed in the literature as 'assessment literacy'.

Assessment Literacy. Assessment literacy is an understanding of the basic principles of sound assessment (Stiggins, 2002). In other words, an individual who is literate in assessment activities has the knowledge and skills to administer, score, and interpret assessments with high quality. The primary impetus for examining teacher's assessment literacy was due to the dissemination of the 1990 Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990), delineating the competencies necessary for teachers to conduct sound assessments.

Research following the dissemination of the standards explored teachers' level of assessment literacy using multiple-choice content related measures, self-efficacy measures, and a combination of the two measures. In general, the literature has indicated that preservice and inservice teachers have low, insufficient levels of

assessment literacy (Plake, Impara, & Fager, 1993; Mertler 2003; Gotch & French, 2013). More specifically, studies using multiple-choice content-related measures, directly associated with the standards developed in 1990, yielded common areas of weakness including developing assessment methods appropriate for instructional decisions (Quilter & Gallini, 2000; Mertler 2009), developing valid pupil grading procedures which use pupil assessments (Mertler 2003, 2009), and communicating assessment results to students, parents, and other audiences (Plake, Impara, Fager, 1993; Mertler 2009). Additionally, two common strengths emerged including choosing assessment methods appropriate for instructional decisions and administering, scoring, and interpreting the results of both externally-produced and teacher-produced assessment methods (Plake, Impara, & Fager, 1993; Mertler, 2003). However, further analysis of teachers' interpretation of results from externally-produced assessment methods revealed teachers' level of knowledge is inconsistently developed within this domain (Gotch & French, 2013).

Further, studies comparing inservice and preservice teachers' level of assessment knowledge have reported mixed results, suggesting it is unclear if teaching experience in the field influences assessment knowledge (Mertler, 2003; Alkharusi, Kazem, Al-Musawai, 2011). In fact, two studies that examined the effects of teaching experience on knowledge of assessment reported mixed conclusions. Alkharusi, Kazem, and Al-Musawai (2011) found that teachers with less than seven years of experience had higher levels of knowledge than inservice teachers with more than seven years of experience, and further, inservice teachers who completed a preservice course in educational measurement had higher levels of knowledge than inservice

teachers who did not complete a preservice course in educational measurement. Conversely, results from Gotch and French (2013) indicated no relationship between the number of years of teaching experience and knowledge or the completion of professional development in educational measurement within the past three years and knowledge.

The Assessment Literacy Inventory is one instrument that was developed to assess the strengths and weaknesses of teacher's application of the seven standards for teacher competence in educational measurement. The instrument includes five scenarios mimicking real-life experiences with assessment in the classroom (Mertler & Campbell, 2005). The internal consistency of the measure, r_{KR} =.74, indicates acceptable reliability within the measure (Mertler & Campbell, 2005). The information gathered from this measure may suggest areas for refining knowledge and skills. A limitation of this measure is the focus on teacher-made assessments because the inventory does not address adherence to the guidelines for valid instructional assessment within an RtI framework (Ysseldyke, Burns, Scholin, Parker, 2010).

Based on the assessment literacy literature, it can be concluded that teachers' have variable levels of assessment knowledge and inconsistent areas of strengths and weaknesses have emerged. It is possible that the noted variability may be a result of inconsistencies between assessment requirements in teacher education programs, state standards for assessment education, and the culture of schools. More than a decade of research on assessment practices, in addition to standards-based reform efforts, led the Assessment Training Institute to develop an updated list of competencies necessary for teachers to understand, to conduct sound assessment. These competencies include: 1)

assessment processes and results serve clear and appropriate purposes, 2) assessments reflect clear and valued student learning targets, 3) learning targets are translated into assessments that yield accurate results, 4) assessment results are managed well and communicated effectively, and 5) students are involved in their own assessment (Stiggins, Arter, Chappuis, & Chappuis, 2004).

Curriculum-based measurement (CBM) is one approach to assessment that possesses the characteristics necessary to fit within an RtI framework (Ysseldyke, Burns, Scholin, Parker, 2010; Ball & Christ, 2012). Evidence suggests that CBM is precise, frequent, and sensitive to change (Shinn & Bamonto, 1998; Deno, 1985). A limitation of the assessment literacy research is the exploration of teachers' knowledge and understanding of CBM. This information would be particularly revealing because CBM was originally developed for use by special education teachers and specialists (Deno, 1985). Special education teachers and specialists typically have more in depth training in educational measurement than general education teachers. As a result, problems with treatment integrity may arise when teachers use these measures to make educational decisions.

Curriculum-Based Measurement (CBM)

CBM is a standard approach to assessment that allows for the efficient measurement of skills aligned to the existing curriculum taught in a classroom (Deno 1983; Shinn & Bamonto, 1998). The technical adequacy of CBM is well-supported in the literature (Good & Jefferson, 1998). (For a detailed review of the available literature on CBM since its' conception in the 1970s, the reader is referred to Tindal (2013). Current research has increasingly focused on examining the appropriate and

inappropriate uses of CBM for determining students' response to instruction and intervention (Ball & Christ, 2012). More specifically, there has been criticism about the use of CBM for the purpose of progress monitoring (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2012). However, CBM is currently being implemented in schools to support decisions about students' response to instruction/intervention. Similar to other approaches to assessment, the use of CBM alone is not sufficient; rather data-based decision making must accompany the use of CBM (Tindal, 2013). Following is a discussion of CBM broken down by the competencies identified by the Assessment Training Institute for conducting sound assessments (Stiggins, Arter, Chappuis, Chappuis, 2004).

Assessment processes and results serve clear and appropriate purposes. The three intended purposes for the use of CBM include screening students for academic difficulties, measuring student growth, and recognizing a need for a change in the instructional program when the current program is ineffective (Shinn & Bamonto, 1998). The first purpose for CBM is screening, or the process of identifying the students that may need additional educational support and the students that are reaching grade level expectations (i.e. benchmarks). All students are screened, rather than a selected set of students suspected of having difficulties, to reduce the chance of overlooking any student who is at-risk. Multiple screenings throughout the year allow teachers and principals to evaluate and determine if students are making progress toward an end-of-the-year goal. Screening measures inform teachers that classrooms have a group of heterogeneous students. Scores on these measures may assist teachers

in differentiating instruction to meet students' various needs. CBMs provide a quick and efficient method for assessing large groups of students (Deno, 2003).

The second purpose for CBM at the individual level is to monitor student progress. Unfortunately, it is unknown ahead of time whether or not a student will respond to a specific instructional program or intervention. A student's rate of progress during an instructional program is an indication of their responsiveness to and the effectiveness of the instructional program for that student (Fuchs & Deno, 1991). By collecting data points during instruction, or intervention, it can be determined whether the student's performance and/or learning is improving or not. If the graphed relationship between number of weeks and the students' scores does not indicate growth, as determined by the teacher or team, changes to the instruction would be indicated. However, if the student is making adequate progress, the current instructional program should remain unchanged.

In recent years, an examination of the technical characteristics of CBM for progress monitoring has been prominent in the literature. More specifically, passage equivalence (Christ & Ardoin, 2009; Betts, Pickart, & Heistad, 2009), decision rules (Ardoin, Christ, Morena, Cormier, & Klingbeil, 2012), domain sampling (Shapiro, 2013), probe-set development (Christ & Ardoin, 2009), and standard errors associated with commonly used CBM (i.e. DIBELS and AIMSWEB) (Christ & Ardoin, 2009) have been explicitly studied. An empirical base for the technical features is paramount to the standardization of CBM, and further, may lead to improvements in consistent decision making within an RtI framework. In other words, it may increase the level of confidence for making special education eligibility decisions.

The final purpose for the use of CBM is decision making within a problem-solving model. The data collected from CBM is valuable for informing decision makers about various questions. First, CBM is used to identify whether a problem exists or not. A problem is defined as a discrepancy between what is expected and what is occurring (Shinn & Bamonto, 1998). Second, the severity of the problem will validate if special services are required and what level of support is necessary to meet student needs. The severity of the problem is verified by the magnitude of discrepancy between a student's measured skills and a normative measure of those same skills. Third, CBM is used to explore solutions and set goals. To determine a goal, Shinn (2002) recommends that the local norms for the school are considered for typical performance. This means that the goal is specific to the curriculum, instruction, and the environment because it is based on students within a specific context. Fourth, after a goal is set CBM is used to determine if the student is making adequate progress toward achieving that goal. Finally, the initial screening tool will reevaluate the discrepancy between individual student performance and local normative performance to determine if the problem continues to exist or if the problem has been resolved.

Assessments reflect clear and valued student learning targets. CBMs are tools used to indicate basic skill level development (Deno, 1985). Similar to a thermometer in the medical profession, CBM measures the 'vital signs' of a student's ability in different areas of academics (Shinn & Bamonto, 1998). The purpose is to demonstrate if a problem exists using an efficient method. However, these measures are limited because they are skill based.

CBM is also known as a general outcome measurement approach to assessments. That is to say that Fuchs and Deno (1991) describe CBM as using a long-term measurement approach. This means that the skill/s students are expected to have at the end of the year are being assessed. This is contrary to common mastery measures that assess specific skills currently being taught. Instead, the tasks students perform on CBMs require students to apply numerous subskills learned throughout the year (Hosp, Hosp, & Howell, 2007). For example, oral reading fluency is a reading achievement measure by CBM. To read fluently, some of the skills students apply are decoding, vocabulary knowledge, syntax, and background knowledge (Adams, 1990). Poor reading fluency would indicate that the student is struggling with one or more subskills or the integration of subskills, and further diagnostic information is necessary. Results from a meta-analysis of 41 correlational studies examining CBM oral reading fluency as an indicator of reading achievement demonstrated that CBM oral reading fluency was a significant predictor of state-specific tests of reading standards, a significant predictor of third grade reading outcomes, and a significant predictor reading comprehension (Reschly, Busch, Betts, Deno, & Long, 2009).

Learning targets are translated into assessments that yield accurate results. CBM has standardized administration and scoring procedures. The purpose of standardization is to ensure that the procedures are consistent across students and testing periods, which minimizes error. Prior research has demonstrated CBM to be a valid and reliable approach to measurement (Good & Jefferson, 1998). Numerous studies have demonstrated the concurrent validity between CBMs and standardized measures of student achievement. For example, Fuchs, Fuchs, and Compton (2004)

found that CBM word identification fluency was highly correlated (i.e. r= .52-.82) with subtests on the Woodcock Reading Mastery Test-Revised and highly correlated (i.e. r= .73-.93) with measures on the Comprehensive Reading Assessment Battery. In another example, Ardoin et al. (2004) found that CBM reading was moderately to highly correlated (i.e. r = .35-.74) with subtests on the Woodcock-Johnson-III and the Iowa Test of Basic Skills. Additionally, CBM maze selection was moderately correlated (i.e. r= .31-.51) with subtests from the Woodcock-Johnson-III and the Iowa Test of Basic Skills. For an in depth review of the literature supporting the technical adequacy of CBM related to measures, materials, and representation of growth, the reader is referred to Wayman, Wallace, Wiley, Ticha, and Espin (2007).

Further, a strong link between the use of CBM in the classroom and student achievement has been documented (Fewster and MacMillan, 2002; Fuchs, Fuchs, and Hamlett, 1989). For example, Fuchs, Fuchs, and Hamlett (1989) examined reading achievement for students whose teachers used CBM. The study compared three groups of teachers: one group of teachers who used CBM only for measurement, one group of teachers who used CBM for measurement and instructional decision making, and a control group who did not use CBM at all. The results demonstrate that for teachers who used CBM for both measurement and decision-making, students had higher rates of growth than the measurement only group with a medium to large effect size of .72. No reliable differences were found between the measurement only group and the control group. A limitation of the study was the lack of random assignment. This study indicates that to achieve the most accurate results with the use of CBM, it is essential to use the results for instructional decision-making.

Assessment results are managed well and communicated effectively. The academic skill areas CBM measures are reading, written expression, spelling, and mathematics computation. CBM is valued for its sensitivity to intra-individual growth over time, repeatability of use, access to multiple equivalent forms, inexpensive cost, time efficiency, easy-to-learn administration procedures, and reliability of the measures (Deno 2003). These design characteristics were an integral feature of the development to provide teachers with a simple way to monitor student achievement and inform instruction (Deno, 1985).

To communicate results effectively, an accurate understanding of how to interpret results is necessary. Scores can be compared to two different types of normative scores: local or national. Normative scores provide a criterion for evaluating student success. National norms are available for both AIMSWEB (AIMSweb National Norms Technical Documentation, 2012) and DIBELS, and both sets of norms are based on research studies conducted with nationally representative samples. This provides an indication of how students' scores, or overall school scores, compare to other students, or schools, in the nation. Conversely, local norms provide an indication of how students' scores compare to other students in the district, or school. To understand the comparison between a student's score and a normative score, scores are attributed to a percentile rank. A percentile rank represents the percentage of scores that fall at or below the student's score. Students' scores may also be compared to a benchmark, which is a desired level of performance that indicates proficiency.

Additionally, teachers should be knowledgeable in sharing information from progress monitoring graphs. Wayman et al. (2011) examined special education

teachers' interpretations and understanding of progress monitoring data using a think-aloud approach, which means that teachers looked at a progress monitoring graph and expressed their thinking process. Results revealed strengths in teachers' understanding of goal attainment, function of the goal line, and setup of the graph and weaknesses in teachers' understanding of the slope, baseline data, and the meaning of the words read correctly measure (Wayman et al., 2011). Although these results are informative, it is important to note that general education teachers would not be expected to have the same depth of knowledge as special education teachers.

Students are involved in their own assessment. CBMs were originally developed as a formative evaluation tool (Deno, 1985). Formative assessments are administered during instruction, or an intervention. Learning occurs through the use of formative assessment for both students and teachers. Moreover, teachers can use formative assessment to enhance their instruction to meet student needs (Green & Johnson, 2010 p. 97). Students can become involved in their own learning when teachers share learning goals with the students, and when teachers provide feedback about the students' current performance in comparison to their goal. Hattie (2009) conducted a meta-analysis of 800 studies examining variables that impact student achievement and results demonstrated that formative evaluation had the strongest effect on student achievement (d=.9). Further, formative evaluation had a greater impact on student achievement than students' self-instructed strategies, students' socio-economic status, teacher expectations, and teachers' knowledge of subject matter.

Purpose of the study

CBM is a valid and reliable approach to assessing student needs. Within an RtI framework CBM can be utilized, in conjunction with other information, to make important educational decisions. The complexity of the measures may lead to challenges for those without a background in educational measurement. Assessment literacy is a field of study that explores teachers understanding of assessment based on the 1990 Standards for Teacher Competence in the Educational Measurement of Students. A limitation of the assessment literacy research is the absence of CBM. The current study explores to what extent a relationship exists between teachers' scores on the Assessment Literacy Inventory and their corresponding scores on a measure of CBM knowledge. Understanding the relationships between knowledge and understanding on these two measures may lead to a better understanding of the general assessment competencies needed to increase awareness and understanding of the use of CBM. An implication of this study may include how to improve the training teachers receive in using assessment in the classroom.

This study tested test the following primary hypothesis:

1. It is hypothesized that CBM related subareas of knowledge would be differentially, not similarly, related to Assessment Literacy competencies for conducting sound assessments.

The second research question was exploratory because prior research reported mixed results. Consequently, a specific hypothesis was not examined.

CHAPTER II

METHODOLOGY

Participants

For the primary research question, the sample is based on respondents who completed both the CBM instrument and the ALI. The total sample size was 27 participants. Approximately 3% of the teachers who were contacted completed both instruments. Participants were 24 inservice teachers and three school psychologists from elementary schools in the Northeast region of the United States. All participants were female and the majority (i.e. 89%) of participants identified with Caucasian ethnic backgrounds. Almost half (i.e. 44%) of the sample had ten or more years of teaching experience, 33% had between five and nine years of teaching experience, and 22% of the sample had less than five years of teaching experience. More than half (i.e. 63%) of the inservice teachers identified themselves as currently teaching general education and the remaining ones identified themselves as reading, or literacy, specialists (i.e. 19%) and special education teachers (i.e. 7.4%). Further, individuals teaching in the general education setting taught kindergarten (i.e. 11%), first grade (i.e. 19%), second grade (i.e. 11%), third grade (i.e. 7%), or fourth grade (i.e. 19%). The majority (i.e. 82%) of the sample earned a Master's degree as their highest level of education. All participants taught at a school that implements CBM. Finally, half of the participants (i.e. 48%) attended two or more trainings on CBM, whether the training was during their preservice or inservice experience (see Table 1). The sample demographic characteristics are not likely representative of the general teacher population. Greater diversity would be expected within the population. Additionally,

Table 1. Inservice Teachers who Completed Both Instruments (N=27)

	Characteristic	n	%
Gender	Male	0	0.0%
	Female	27	100.0%
Ethnicity	Asian or Pacific Islander	1	3.7%
	African American	0	0.0%
	American Indian or Alaskan Native	0	0.0%
	Hispanic or Latino American	2	7.4%
	Caucasian	24	88.9%
	Multiracial	0	0.0%
Years of Experience Teaching	2 years of less	2	7.4%
	3 - 4 years	4	14.8%
	5 - 9 years	9	33.3%
	10 + years	12	44.4%
Current Teaching Position	General education teacher	17	63.0%
	Special education teacher	2	7.4%
	Reading/Literacy specialist	5	18.5%
	Math Specialist	0	0.0%
	School Psychologist	3	11.1%
	Other	0	0.0%
Grade Level Currently Teaching	Kindergarten	3	11.1%
	1st grade	5	18.5%
	2nd grade	3	11.1%
	3rd grade	2	7.4%
	4th grade	4	14.8%
	5th grade	0	0.0%
Highest Level of Education	Bachelor's degree	1	3.7%
	Bachelor's degree with some CEU credits	3	11.1%
	Master's degree	22	81.5%
	PhD	1	3.7%
Number of Courses or Trainings on CBM Attended	1	14	51.9%
	2	5	18.5%
	3	4	14.8%
	4	2	7.4%
	5+	2	7.4%

school psychologists, specialists and special education teachers were included because of the exploratory nature of the study. Investigating the correlation by position may have provided insight to the relationship between training and knowledge about CBM.

For the secondary research question two samples were used; one sample was based on all respondents who completed the CBM instrument (N= 32) and one sample was based on all respondents who completed the ALI (N= 47). Thirty-two participants completed the CBM instrument, which is approximately 4% of the teachers who were contacted. All participants taught at a school that implements CBM. Almost all of the participants (i.e. 97%) were female and the majority (i.e. 91%) of participants identified with Caucasian ethnic backgrounds. Approximately half (i.e. 47%) of the sample had 10 or more years of teaching experience, 34% had between five and nine years of teaching experience, and the remaining ones (i.e. 19%) had less than five years of teaching experience. More than half (i.e. 66%) of respondents identified themselves as currently teaching general education, and the remaining ones identified themselves as school psychologists (i.e. 13%), reading specialists (i.e. 9%), special education teachers (i.e. 6%), and a math specialist (i.e. 3%). Further, individuals teaching in the general education setting were teaching Kindergarten (i.e. 13%), first grade (i.e. 22%), second grade (i.e. 9%), third grade (i.e. 6%), fourth grade (i.e. 6%), or fifth grade (9%). A majority (i.e. 81%) of respondents earned a Master's degree as their highest level of education. Finally, 60% of the sample completed two or more courses and/or trainings on CBM (see Table 2).

Forty-seven participants completed the ALI, which is approximately 6% of the teachers who were contacted. All participants were female and taught at a school

implementing CBM. The majority (i.e. 92%) of participants identified themselves with Caucasian ethnic backgrounds. Approximately half (i.e. 53%) of the sample had 10 or more years of teaching experience, 32% had between five and nine years of teaching experience, and 15% had less than five years of teaching experience. Over half (i.e. 55%) of the participants identified themselves as currently teaching in the general education setting, and the remaining ones identified themselves as reading specialists (i.e. 17%), special education teachers (i.e. 9%), other (8.5%), school psychologists (6%), and a math specialist (2%). Further, individuals teaching in the general education setting taught Kindergarten (i.e. 11%), first grade (i.e. 13%), second grade (i.e. 17%), third grade (i.e. 6%), or fourth grade (11%). A majority (i.e. 79%) of participants earned a Master's degree as their highest level of education. Finally, approximately half (i.e. 53%) of the sample completed two or more courses and/or trainings on CBM and the remaining ones (i.e. 47%) completed one course and/or training on CBM (see Table 3).

Measures

Curriculum-Based Measurement Instrument

Rationale for Item Construction. The CBM instrument used in the current study is an adapted version of an established CBM test. The original version (Contained in Appendix A) consisted of 13 items, all requiring written production responses, worth a total of 100 points and was developed in 1993 for use by school psychologists. The measures' content validity was established by thorough and intensive review and evaluation by faculty and graduate students with expertise in CBM and the field of school psychology. The committee, supported by a federal leadership training grant for CBM, included the principal investigator, Dr. Mark

Table 2: Inservice Teachers: CBM Instrument (N= 32)

	Characteristic	n	%
Gender	Male	1	3%
	Female	31	97%
Ethnicity	Asian or Pacific Islander	1	3%
	African American	0	0%
	American Indian or Alaskan Native	0	0%
	Hispanic or Latino American	2	6%
	Caucasian	29	91%
	Multiracial	0	0%
Years of Experience Teaching	2 years of less	2	6%
	3 - 4 years	4	13%
	5 - 9 years	11	34%
	10 + years	15	47%
Current Teaching Position	General education teacher	21	66%
	Special education teacher	2	6%
	Reading/Literacy specialist	3	9%
	Math Specialist	1	3%
	School Psychologist	4	13%
	Other	0	0%
Grade Level Currently Teaching	Kindergarten	4	13%
	1st grade	7	22%
	2nd grade	3	9%
	3rd grade	2	6%
	4th grade	2	6%
	5th grade	3	9%
Highest Level of Education	Bachelor's degree	1	3%
	Bachelor's degree with some CEU credits	3	9%
	Master's degree	26	81%
	PhD	1	3%
Number of Courses or Trainings on CBM Attended	1	13	41%
	2	8	25%
	3	6	19%
	4	2	6%
	5+	3	9%

Table 3: Inservice Teachers: ALI Instrument (N=47)

	Characteristic	n	%
Gender	Male	0	0%
	Female	47	100%
Ethnicity	Asian or Pacific Islander	1	2%
	African American	0	0%
	American Indian or Alaskan Native	0	0%
	Hispanic or Latino American	2	4%
	Caucasian	43	92%
	Multiracial	1	2%
Years of Experience Teaching	2 years of less	2	4%
	3 - 4 years	5	11%
	5 - 9 years	15	32%
	10 + years	25	53%
Current Teaching Position	General education teacher	26	55%
	Special education teacher	4	9%
	Reading/Literacy specialist	8	17%
	Math Specialist	1	2%
	School Psychologist	3	6%
	Other	4	9%
Grade Level Currently Teaching	Kindergarten	5	11%
	1st grade	6	13%
	2nd grade	8	17%
	3rd grade	3	6%
	4th grade	5	11%
	5th grade	0	0%
Highest Level of Education	Bachelor's degree	1	2%
	Bachelor's degree with some CEU credits	8	17%
	Master's degree	37	79%
	PhD	1	2%
Number of Courses or Trainings on CBM Attended	1	22	47%
	2	11	23%
	3	7	15%
	4	3	6%
	5+	2	4%

Shinn, and four advanced graduate students. The committee worked closely to ensure the appropriateness and the clarity of the content. For the purposes of the current study, the CBM test was adapted for the intended audience, teachers. Items were developed in cooperation with experts, including researchers and educators from the University of Rhode Island, who provided feedback and reviewed the items. A multiple choice item format was chosen to measure teachers' recognition of terms and concepts related to the use of CBM in the classroom. Items specific to the role of a school psychologist, including special education eligibility and Individualized Education Plan goals, were discarded. After careful review, it was decided that items relied heavily on definitions and recognition of concepts, and this may not be sufficient for measuring teachers' knowledge and understanding. Therefore two applied scenarios were added to the instrument. Each scenario included a data table accompanied by five questions. These items were added to support and measure the domain of interpretation. The first version of the instrument, a total of 48 items, consisted of 42 single-select multiple-choice items, including two scenarios, and six multi-select multiple choice (i.e. check all that apply) items. Validation procedures and the final version of the CBM instrument are described below.

Framework Conceptualizing Subdomains of Curriculum-Based Measurement Knowledge. The CBM instrument was designed to measure teachers' level of knowledge and understanding of the use of CBM in the classroom. Three major dimensions characterize the use of CBM: administration and scoring, interpretation, and data-based decision making. Knowledge of administration and scoring procedures are necessary, but not sufficient for using assessments in the classroom. Given CBMs

principal use as assessment for formative evaluation, described by Shinn and Bamonto (1998), results must be applied to instructional decision-making. To accomplish this, knowledge of interpretation and data-based decision making are necessary requisites. Items were developed to assess knowledge and understanding specific to all three domains. Please refer to Appendix B for a visual representation of the framework for item construction.

Validation Procedures. The first version of the instrument was reviewed by an expert in educational measurement in the teacher education department, at the University of Rhode Island. Specifically, items were reviewed to determine appropriateness and relevance for inservice teachers. Following feedback and discussion two items were removed, two items were added, and nine items were reworded for clarification. Additionally, two items were modified into applied scenarios (i.e. given the provided information, what would the next step be). As a result, the second version of the instrument, a total of 48 items, consisted of 43 single-select multiple-choice items, including four scenarios, and five multi-select multiple choice items.

Next, ten graduate students in school psychology, who completed a course in CBM and assessment one year prior, completed the instrument and were encouraged to provide feedback on the clarity and appropriateness of items. As a result one item was removed for poor wording and two items, both addressing the same term, were carefully reviewed. To review the items, numerous state department of education websites were explored to find a more appropriate term; however, it was noted that there were inconsistencies among states. The two items were retained and noted to

discuss with potential respondents, inservice teachers. Further, the five items following the second data table were modified to repeat the same five items from the preceding data table. Consequently, the same five items were presented for two different data tables.

Then, four inservice teachers volunteered to meet with the researcher to carefully review and provide feedback on the third version of the instrument, a total of 47 questions, consisting of 43 single-select multiple choice items, including four scenarios, and five multi-select multiple choice items. Three of the volunteers taught first grade and had five years of experience implementing CBMs in their school. One volunteer taught second grade, had two years of experience implementing CBMs, and had recently earned her dual Masters' degree in a special education and literacy. Items were reviewed one at a time and the teachers provided feedback regarding the clarity, appropriateness, and difficulty of the items. In general, teachers thought the instrument challenged them to think critically. More specifically, teachers suggested they were unfamiliar with terminology and theory, but they understood the basic concepts of CBM and how it is used in the classroom. In other words, teachers felt they possess the knowledge of how to implement and use CBM in the classroom, but are less familiar with why it is used. Additionally, teachers expressed concern with the length of item response options and how it influenced the overall time to complete the instrument. In response to teacher feedback, three of the multi-select multiple-choice items were removed, fourteen items were reworded for clarification, and three basic concept questions were added. With the noted revisions, the fourth and final version of the instrument consisted of 47 single-select multiple-choice items, including four

scenarios, and two multi-select multiple-choice items (See Appendix C). In the present study, the internal consistency coefficient for the CBM instrument was r_{KR20}= .469.

Assessment Literacy Inventory

The Assessment Literacy Inventory (ALI) was designed by two experts in the field of educational assessment to be aligned with the Standards for Teacher Competence in Educational Assessment of Students (AFT, NCME, & NEA, 1990), and questions were developed to mimic real-world applications of the competencies. The ALI is a 35-item measure of teacher's assessment knowledge. The measure includes five classroom-based scenarios that present teachers facing numerous assessment-related decisions. Participants are asked to answer seven questions for each scenario; one question pertaining to each of the seven standards (AFT, NCME, & NEA, 1990). Each correct answer is worth one point and an overall score out of a possible 35 points is awarded. When used with preservice teachers the internal consistency (KR20) of the measure is r_{KR20}= .74, demonstrating sufficient reliability within the measure (Mertler & Campbell, 2005). The item difficulty values range from .212 to .992 providing an effective range of difficulty (Mertler & Campbell, 2005; Chase 1999). Further, the mean item discrimination value .313 demonstrates that the majority of the items on the ALI are good quality questions (Mertler & Campbell, 2005; Chase, 1999). Further, the ALI was a modified, shorter and easier to read, version of the *Teacher Assessment Literacy Questionnaire* (TALQ; Plake, Impara, & Fager, 1993). Prior studies employing the *TALQ* with inservice teachers have reported internal consistency coefficients ranging from r_{KR20}= .5 (Quilter & Gallini, 2000) to r_{KR20}= .57 (Mertler, 2003). Moreover, the first, original study using

the TALQ (1993) reported an internal consistency coefficient of r_{KR20}= .54, which was based on a nationally represented sample of inservice teachers (Plake, Impara, & Fager, 1993). In the current study, the internal consistency coefficient for the ALI was r_{KR20}= .525.

Qualtrics Survey Software

The final version of the CBM instrument was uploaded onto Qualtrics, an online survey software system. Maintaining participant anonymity was a priority for producing an electronic version of the instruments. Multiple actions were performed to maintain participant anonymity. First, Qualtrics offers a feature to anonymize participant responses, which means an IP address was not recorded and therefore it was unknown to the researcher who chose to participate. To utilize this feature, the anonymize results option was selected in the survey flow and in the survey options during instrument development. Second, to associate responses on the two measures, a random number generator was used. Participants were assigned a random four-digit number at the beginning of the first instrument and were asked to enter the four-digit number during the second instrument. Accordingly, participant responses were associated with the random number and the researcher could not associate responses to any individual. Third, participants did not report information about their school or district, and individual responses were not shared with building principals or district superintendents.

Two versions of each measure were generated on Qualtrics to allow for counterbalancing the distribution of instruments. Therefore there was an ALI part one, CBM part one, ALI part two, and CBM part two. For purposes of distribution, ALI

part one was paired with CBM part two and CBM part one was paired with ALI part two. The part one and part two versions were not identical. Part one included assignment of the random number and the demographic questions. Part two included entering the previously assigned random number and a hyperlink to the incentive survey. The incentive survey offered participants the opportunity to win one of eight twenty-five dollar gift cards to amazon. Participants were asked to enter their email address to be eligible to win. A brief thank you note was added to the end of each instrument. Additionally, an email repertoire was created including the following emails: an initial invite to participate, a reminder to complete part one, an introduction to part two, and a reminder to complete part two. The student teacher versions were comparable in all aspects, except the demographic questions.

Following the creation of the instruments on Qualtrics, values were coded to indicate a correct or an incorrect response. For all items, a one indicated a correct response and a zero indicated an incorrect response. Additionally, multi-select multiple choice items were coded so that the participant received credit, one point, for the item only if they chose four or more correct responses. As a result, a total score of correct responses and an overall percentage of correct responses, in addition to individual items, were recorded for each participant. Finally, all questions were set as a 'forced-response,' which means that participants were required to respond to all questions. Consequently, incomplete responses were a results of discontinuing the instrument.

Finally, to pilot the use of the instruments on Qualtrics, five graduate students in the special education program, who recently completed a course in educational

measurement at the University of Rhode Island, completed the instruments electronically. Students provided positive feedback including ease of use and no concerns about the clarity of instructions. Data were downloaded from the website and reviewed to confirm the output provided what was expected.

Procedures

Recruitment

Participants for this study were recruited from multiple states in the Northeast region of the United States including New York, Massachusetts, Rhode Island, Connecticut, and New Hampshire. Recruitment followed a three-step process: obtain permission from superintendents, obtain permission from principals, and invite teachers to participate. Accordingly, teachers were contacted if their superintendent and school principal granted permission. Further, superintendent names and emails were obtained from state department of education websites. For each list of superintendents acquired, a subset of names was searched to determine if the list was current and accurate. Additionally, superintendent titles were obtained from individual district websites. An email requesting permission to conduct research (Appendix D) was sent to a total of 362 superintendents: 203 in New York, 87 in Connecticut, 45 in Massachusetts, 26 in Rhode Island, and two in New Hampshire. Initially, 75 superintendents were contacted. Based on response rates additional superintendents were contacted; approximately 50 superintendents were contacted bimonthly for three and a half months to recruit additional participants. Approximately seventeen percent of superintendents contacted responded to the email. More specifically, of the contacted superintendents, four percent granted permission, eleven percent denied

permission and two percent sent the request to an alternative person within the district. Two districts requested to meet with the researcher in person to discuss the project.

The researcher met with the Director of Special Education Services in two school districts that requested a meeting. The first meeting occurred during the first month of recruitment. The Director proposed minor changes to the project to make it more accessible in their district. First, he suggested the researcher meet with building principals to explain the study and answer questions. Second, he requested clarification for the term Curriculum-Based Measurement. More specifically, he recommended using Star (i.e. Renaissance Star) as an example of CBM because teachers were currently implementing it. Next, he recommended shortening the interval between the distributions of the two instruments from two weeks to three days. Finally, he selected online distribution rather than hard copy distribution. Three weeks following the initial meeting the researcher met with building principals to explain the project and answer questions. The principals agreed to speak with their teachers at their next faculty meeting and the first instrument was sent out immediately following.

The second meeting occurred during the fourth month of recruitment. The Director supported the purpose of project; however, he expressed concern regarding the expectation for teachers to complete both instruments and the time commitment it entailed. He volunteered to approach principals, special education teachers, and general education teachers to develop interest in the project. If interest was expressed, he would distribute the instruments via email to volunteers. Immediately following the

meeting the researcher sent the Director the invite to participate with hyperlinks to both instruments for distribution to the teachers.

Following approval from superintendents, fifty principals were sent an email requesting permission to contact teachers in their building. As a result of low response rates, principals were sent a follow-up email two weeks following the initial email. Approximately forty-eight percent of principals contacted responded to the email. More specifically, forty-four percent granted permission and four percent denied permission. Further, some principals (i.e. 27%) preferred to distribute the assessment instruments to their teachers, a few principals (i.e. 9%) collected names of volunteers to send to the researcher, and the remaining principals (i.e. 64%) preferred the researcher to distribute the instruments. Following principal requests to distribute, the assessment instruments were modified for this purpose. Originally, the assessment instruments were designed for distribution by the researcher to simplify the process for teachers. To permit principals to send the instruments while maintaining private access (i.e. invitation only), a password was added to all versions of the assessment instruments.

Upon receiving approval from principals, teacher emails were obtained from district and school websites. Teacher emails were first added to an excel file and then uploaded as a panel onto the Qualtrics website. Once uploaded onto the website, one email could be sent to all individuals within a panel. Teachers contacted by the researcher received a total of four emails each sent at one-week intervals. The four emails included an initial invitation to participate, a reminder to complete part one, an

introduction to part two, and a reminder to complete part two. Teachers contacted by their school principal received one email that included a hyperlink to each instrument.

Following three months of recruiting, response rates were significantly lower than expected. In particular, numerous participants completed the first instrument and failed to complete the second instrument. Further, although participants were allowed to exit and reenter the instrument for two weeks, a number of participants responded to fewer than half of the items. In pursuit of recruiting additional participants, the assessment instruments were modified for distribution to student teachers. Two Universities were contacted requesting permission to conduct research and one approved the project. The director of teacher education, at the University of Rhode Island, granted permission to contact students in their junior and senior years of the teacher education program. A total of 264 students were contacted at the end of the fall semester. Students received a follow-up, reminder email approximately one week following the initial invitation to participate. Approximately 2% of student teachers completed the first instrument and 1.5% of student teachers completed both instruments. As a result of the low response rate from student teachers, these respondents were not included in the analysis.

At the beginning of the study, there were three inclusionary criteria for participation. First, participants were required to be currently working in a school implementing CBM. Second, participants were required to have completed a preservice, either undergraduate or graduate, course in assessment or educational measurement. Third, participants had to have attended an inservice presentation or training on the use of CBM. These three questions appeared immediately following

the notice about informed consent. If participants responded ‘no’ to any of the three questions, they were redirected to the thank you note at the end of the instrument. After approximately six weeks of data collection it became evident that a large portion of participants were answering no to either the second or the third question. As a result, the researcher modified the questions so participants could continue with the instrument after responding no. However, the first question was retained as the only inclusionary criteria.

Informed Consent

Informed consent was obtained electronically on the Qualtrics website. When volunteers followed the hyperlink from the invite email to the instrument, a notice regarding informed consent (Appendix E) appeared. Participation in the study was voluntary and anonymous. Once informed consent was obtained, by participants selecting ‘yes,’ the instructions to the instrument appeared. If an individual selected ‘no,’ the site was redirected to the thank you note at the end of the instrument.

Design

The strength of the relationship between measures was examined using Pearson’s product-moment correlation coefficient. Pearson correlation is a method for exploring the direction and strength of a relationship between variables (Pallant, 2010). It is important to note that although two variables may be related, it does not mean that one variable causes the other variable. Further, correlations obtained with small sample sizes (i.e. N= < 100) are less reliable than correlations obtained with sample sizes larger than 100 (Nimon, Zientek, Henson, 2012).

As a result of the sample size, a descriptive study design was chosen for investigating the secondary research question. The purpose of descriptive research is to describe what is occurring, by organizing data into patterns that emerge. Descriptive statistics involves analyzing variables one at a time. In other words, the researcher described what the data showed for performance on the CBM measure in relation to training variables, separate from performance on the ALI measure in relation to training variables. As this study is exploratory in nature, it is important to describe the data to form explanations that can be tested in future research. For the purposes of this study summary data will be reported including measures of central tendency.

CHAPTER III

FINDINGS

Analyses

Descriptive statistics were used to examine the distributions of scores on both instruments. Further, descriptive analyses (i.e. means, standard deviations, skewness and kurtosis, normality, and boxplot graphs) are reported to describe the data. Then, a correlational analysis was used to determine if a relationship exists between teachers' knowledge of assessment literacy, as measured by the Assessment Literacy Inventory, and knowledge about Curriculum-Based Measurement, the first purpose of the study. Additionally, a correlation matrix was created to explore the relationships between teachers knowledge based on competencies outlined by Stiggns, Arter, Chappuis, Chappuis (2004). Finally, descriptive analyses are used to explore to what extent training may be related to knowledge about CBM, the secondary research question.

Primary Research Question

Descriptive Statistics and Preliminary Analyses. A total of 27 respondents completed both assessment instruments and were included in the correlational analysis. Thirty-five participants completed part one, but did not complete part two. Scores on the CBM instrument ranged from 55% to 81% accuracy (M = 70% ; SD = 7.458). Scores on the ALI instrument ranged from 43% to 74% accuracy (M= 62.11; SD = 8.187).

Multiple methods were used to assess the normality of the distribution of scores on both instruments (see Table 4). First, skewness and kurtosis were acceptable for both measures according to guidelines presented by Harlow (2005). Next, visual representations, including histograms, normal Q-Q plot, detrended normal Q-Q, and

boxplot were reviewed to verify normality of the data and to identify outliers. Based on visual representations, it was determined that the data were normal and no outliers existed. Finally, the Shapiro-Wilk statistic was obtained and a non-significant result confirmed the normality of the distributions for both measures. In addition to normality, a linear relationship between vairables is required for a correlational analysis. The linearity of the data was assessed using a scatter plot and fit line. A linear fit line was compared to a quadratic fit line to determine which shape best fit the data. A small difference (i.e. .007) between fit lines was observed, and both fit lines revealed a moderate relationship between instruments, verifying that a linear relationship exists between variables. Linearity is further examined in the correlational analysis.

Table 4: Descriptive Statistics for ALI and CBM Measures (Paired Responses)

Measure	N	M	SD	Min	Max	Skewness	Kurtosis
Assessment Literacy Inventory (ALI)	27	62.11	8.187	43	74	-0.588	-0.488
Curriculum-Based Measurement (CBM)	27	70	7.458	55	81	-0.57	-0.651

Correlational Analysis. The relationship between the ALI and a CBM instrument was examined using Pearson product-moment correlation coefficient, as shown in Table 5, using SPSS 22.0 (IBM Corp, 2013). As expected, results revealed a positive, significant correlation between the two instruments ($r = .505$, $p < .01$). In other words, teachers with high scores on the ALI tended to have high scores on the CBM instrument, and teachers with low scores on the ALI tended to have low scores on the CBM instrument. According to guidelines offered by Harlow (2005), the strength of this relationship falls in the moderate range (i.e. r= .4-.59). The primary

purpose of thee current study was to examine if, and to what extent, a relationship exists between the ALI and a CBM instrument. According to the results, a moderate relationship exists between the instruments. To investigate more detailed relationships, hypothesized relationships between competencies of assessment literacy were made, and a correlation matrix was expected to provide insight on the relationships. Following initial analyses, it was concluded that the subcategories (i.e. competencies) were not independent from one another, and as a result a correlation matrix was not further investigated.

Table 5. Correlation between Measures

Measure		CBM	ALI
CBM	Pearson Correlation	1	.505**
	Sig. (2-tailed)		0.007
ALI	Pearson Correlation	.505**	1
	Sig. (2-tailed)	0.007	

**Correlation is significant at the 0.01 level (2-tailed)

Secondary Research Question: Training

Descriptive Statistics and Preliminary Analyses. The second research question explored to what extent training is related to knowledge about CBM. In the current study, the sample size was too small to run a correlational analysis, or to make statistical comparisons between groups, so data were examined using descriptive statistics. As a result, it was unnecessary to run preliminary analyses regarding the normality of the data.

The following descriptive statistics are based on 52 participants who completed both, or one, of the two instruments. All participants were teaching at a school that was implementing CBM.As demonstrated in Table 6, approximately half

of the participants (i.e. 25 out of 52) did not complete a preservice course in assessment or educational measurement. More specifically, as shown in Figure 1, all special education teachers (N=4), half of the math specialists (N=1), half of the literacy specialists (N= 4), and three out of four school psychologists completed a course in assessment. Notably, 16 out of 29 general education teachers did not complete a course in assessment despite spending an average 21% (range 10-50%) of instructional time engaged in assessment related activities (see Figure 2). This means that more general education teachers did not complete a course in assessment than those that did complete a course, which is consistent with previous studies that found less than half of teacher education programs require a course on assessment (O'Sullivan & Chalnick, 1991). Additionally, all participants teaching for less than five years reported completing a course in assessment. Thus, individuals who did not complete a course in assessment had more than five years of experience teaching (see Figure 3).

Figure 1

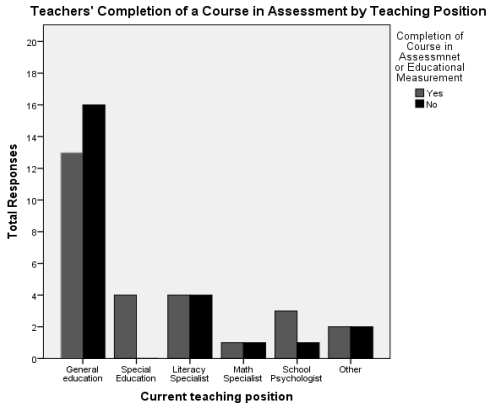


Figure 2

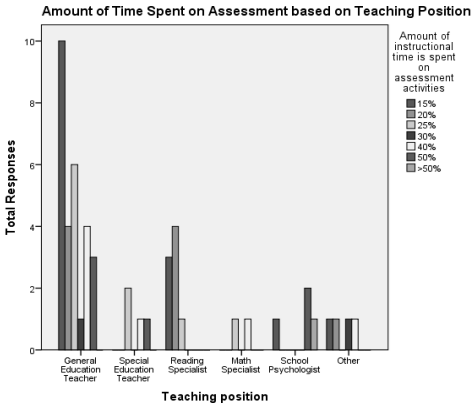
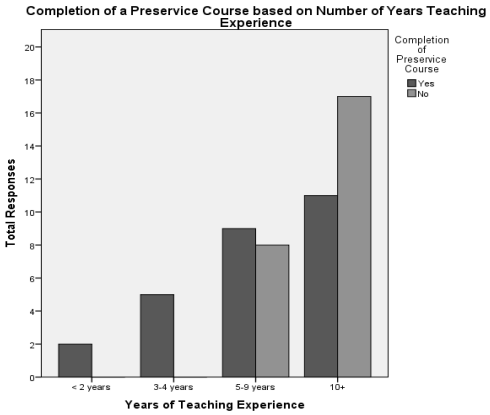


Figure 3



Approximately 38 out of 52 participants reported having attended an inservice, professional development, presentation or training, on the use of CBM (see Table 6). More specifically, as demonstrated in Figure 4, all special education teachers and specialists reported attending professional development on CBM; however, only 13 out of 21 general education teachers reported attending professional development on CBM. This means that some teachers implementing CBM in their classroom may not have received training specific to the use of CBMs. Interestingly, all 15 participants with 10 or more years of experience teaching attended a training compared to 3 out of 6 participants with between one and five years of teaching experience (see Figure 5). Further, 15 out of 19 participants who indicated they partake in the interpretation of CBMs have attended a professional development presentation or training.

Table 6. Descriptive Statistics for Training

	Yes	No	Total
Completed a Preservice Course	27	25	52
Attended an Inservice Training	38	14	52
Involved in Administration of CBM	43	9	52
Involved in Scoring CBM	28	24	52
Involved in Interpretation of CBM	36	16	52

Figure 4

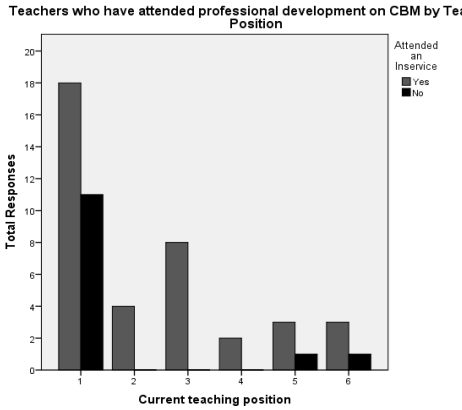
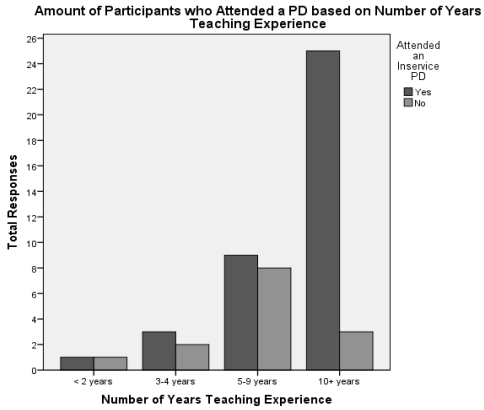


Figure 5



Curriculum-Based Measurement. A total of 32 participants completed the CBM instrument. Scores on the measure ranged from 55% (i.e. 26 out of 47 items) to 87% (i.e. 41 out of 47 items) accuracy (*M*= 71; *SD*= 8.106). Data were used to investigate the relationship between training and knowledge about CBM. Figure 6 presents average CBM scores based on participant completion of a course, in their undergraduate or graduate studies, which included training in the use of assessment, educational measurement, or CBM. Seventeen out of 32 teachers completed a course in the use of assessments (*M*= 74; *SD*= 7.11) compared to fifteen teachers who did not complete a course (*M*= 68.1; *SD*= 8.24). The difference in means between participants who completed a course and those who did not complete a course is approximately two or three items and does not appear to be meaningful.

Figure 7 presents mean CBM scores based on attending a professional development presentation, or training, on CBM. Teachers who attended an inservice presentation or training scored slightly higher (*N* = 23; *M*= 73.35; *SD*= 7.3) than teachers who did not attend an inservice presentation or training (*N*= 9; *M*= 65.72; *SD*= 7.8). The difference is equivalent to approximately four items and does not appear to be meaningful. Further, figure 8 depicts average CBM scores based on how

long ago teachers attended an inservice. As expected, participants who attended inservice five or more years ago had the lowest mean CBM score (*N*= 1; *M*= 68). Interestingly, the highest mean CBM score was found for participants who attended an inservice three to four years ago, which may be a result of learning from experience using CBMs in the classroom (*N*= 3; *M*= 75.89; *SD*= 9.83). Further, some teachers may collaborate with colleagues to learn more about the uses and implications of the measures. The difference in means between the groups is approximately five items. No difference was found for the mean score between participants who attended an inservice less than a year ago (*N*= 13; *M*= 73.32; *SD*= 7.8) and participants who attended an inservice one to two years ago (*N*= 6; *M*= 73.05; *SD*= 6.41).

Figure 9 demonstrates the mean CBM score based on the total number of courses and/or trainings completed on the use of CBM. It was expected that the data would demonstrate a linear relationship (i.e. the more courses and/or training completed the higher the score on CBM instrument); however, data revealed similar mean CBM scores at each level. More specifically, teachers who completed one course had the lowest mean CBM score (*N*= 13; *M*= 67.6; *SD*= 7.92), followed by teachers who completed three courses (*N*=6; *M*= 71.6; *SD*= 9.2), five courses (*N*=3; *M*= 73.1; *SD*= 13.84), four courses (*N*=2; *M*= 74.5; S*D*= 3.01), and finally the highest average for teachers who completed two courses (*N*=8; *M*= 75.3; *SD*= 4.4). The small sample size is one possible explanation for not finding the expected linear relationship.

Figure 6

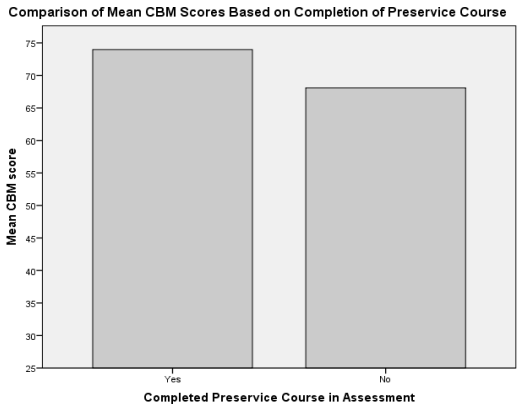


Figure 7

Figure 8

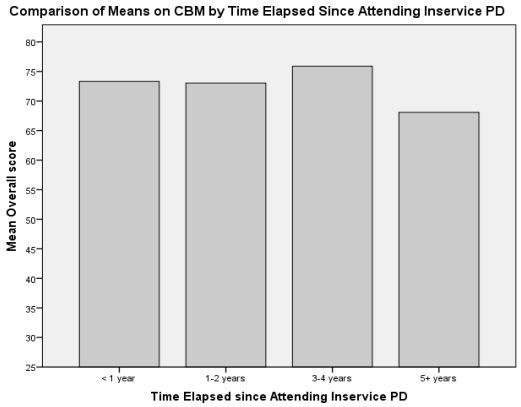


Figure 10 presents mean CBM scores based on the number of years of teaching experience. The figure demonstrates that teachers with less than two years of experience (*N*=2; *M*= 74.5; *SD*= 3.01) in the field and teachers with ten or more years experience in the field (*N*=15; *M*= 72.9; *SD*= 7.39) scored slightly higher than teachers with two to four years of teaching experience (*N*=4; *M*= 69.7; *SD*= 5.32) and teachers with five to ten years of teaching experience (*N*=11; *M*= 68.86; *SD*= 10.22). Similar to the total number of courses completed, a linear relationship was expected to emerge. Notably, teachers with the least amount of teaching experience and the most amount of teaching experience had the highest mean scores.

Figure 9

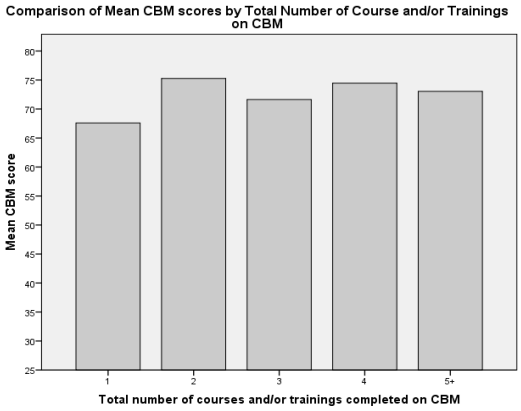
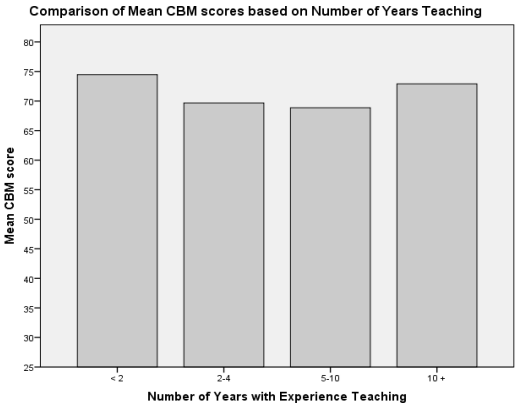


Figure 10



Assessment Literacy. A total of 47 participants completed the ALI. Scores on the instrument ranged from 31% (i.e. 11 out of 35 items) to 74% (i.e. 26 out of 35 items) accuracy (*M*= 58.7; *SD*= 10.4). Prior studies found similar mean ALI scores, which ranged from 56% to 68% of items answered correctly (Quilter & Gallini, 2000; Mertler, 2003; Mertler 2009) Data were used to investigate if training may be related to assessment literacy.

Figure 11 presents a comparison of mean ALI scores based on participant completion of a preservice course in assessment. Participants who completed a course in assessment (*N*= 23; *M*= 59; *SD*= 8.7) scored similarly to participants who did not complete a course in assessment (*N*= 24; *M*= 58; *SD*= 12) and the difference, approximately one item, does not appear to be meaningful.

Figure 12 presents a comparison of mean ALI scores for teachers who attended an inservice on CBM and teachers who did not attend an inservice on CBM. Teachers who attended an inservice training on CBM scored similarly (*N*= 33; *M*= 58.7; *SD*= 10.2) to teachers who did not attend an inservice (*N*= 14; *M*= 58.6; *SD*= 11.4) and the difference does not appear to be meaningful. Further, Figure 13 demonstrates that

mean scores on ALI do not differ based on how long ago teachers attended an inservice. Although a majority of teachers who attended an inservice (i.e. 22 out of 33) did so one or two years ago, their scores were not greater than those who attended an inservice more than three years ago.

Figure 11

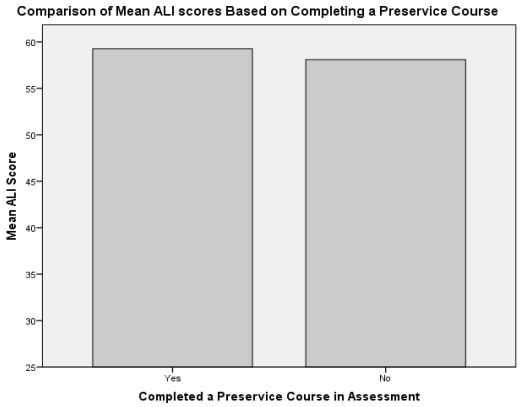


Figure 12

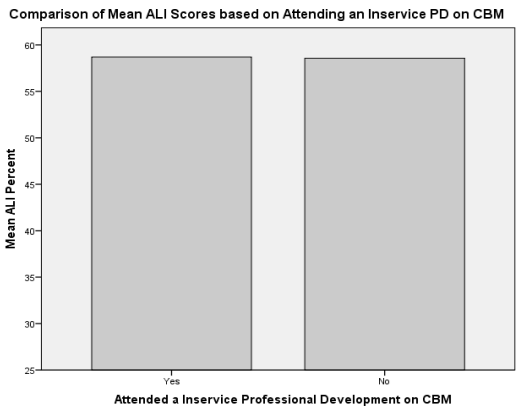


Figure 13

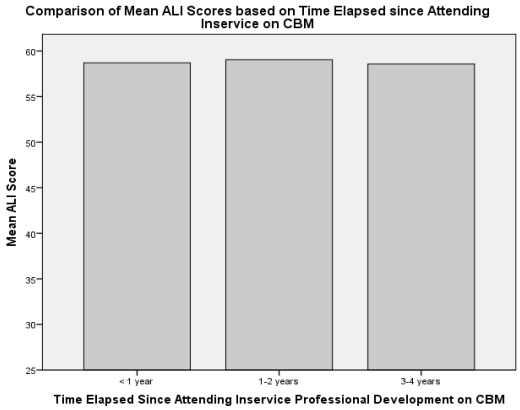


Figure 14 compares mean ALI scores based on number of years teaching in the field. Teachers with less than two years of teaching (*N*= 2; *M*= 60; *SD*= 8.1) and ten or more years of experience teaching (*N*= 25; *M*= 60.8; *SD*= 10.5) scored slightly higher than teachers with 3-5 years of experience (*N*= 5; *M*= 58.3; *SD*= 6.58) and teachers with 6-9 years of experience (*N*= 15; *M*= 55.5; *SD*= 11.3). This finding may suggest

that recent exposure to coursework and a minimum of 10 years of experience contribute to teachers' knowledge of assessment.

Figure 14

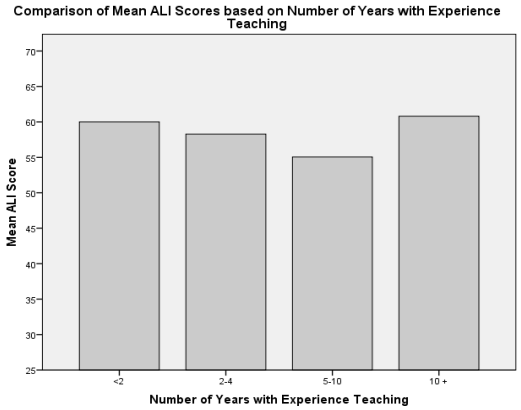
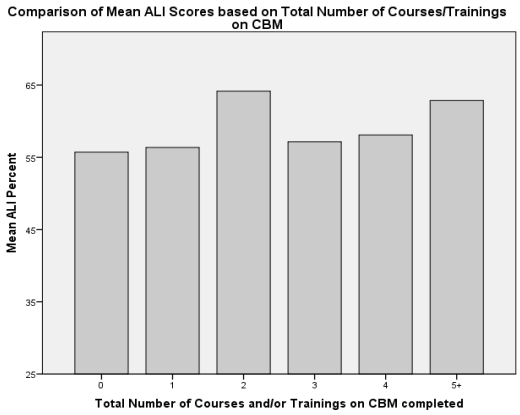


Figure 15 presents mean ALI scores based on the total number of courses and trainings on CBM that teachers completed. The majority of teachers (i.e. 22 out of 47) completed one course or training, followed by teachers who completed two courses and/or trainings (i.e.11 out of 47). Moreover, teachers who completed two courses and/or trainings (*N*= 11; *M*= 64; *SD*= 7.27) and five or more courses and/or trainings (*N*= 2; *M*= 62.9; *SD*= 4.04) scored highest on the ALI instrument. The greatest difference in means was noted between teachers who completed two courses and teachers who did not complete any courses or trainings (*N*= 2; *M*= 55.7; *SD*= 6.07) in CBM; however, the difference is equivalent to approximately three items and does not appear to be meaningful. Notably, there appeared to be a small spike in the mean for two courses; otherwise, a linear relationship emerged as expected. The total number of courses and/or trainings completed on CBM may be related to performance on the ALI and the relationship should be explored further.

Figure 15



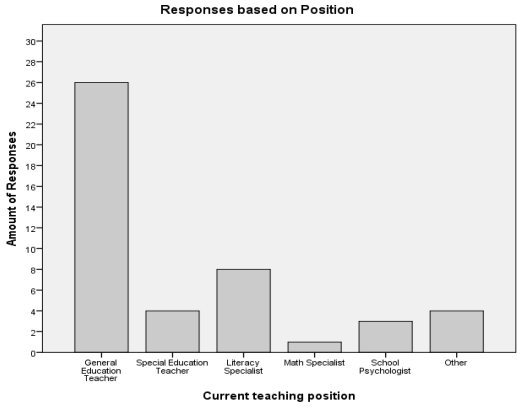
Other information related to Assessment.

Information is based on 52 participants who completed both, or one, of the two instruments. Over half (i.e. N= 29) of the participants were teaching in a general education classroom. The remaining participants, as demonstrated in Figure 16, were special education teachers (i.e. N=4), reading specialists (i.e. N= 8), math specialists (i.e. N=2), school psychologists (i.e. N= 4), and 'other' which was mainly defined as speech and language pathologists (i.e. N=4). A majority of teachers (N=48) reported using premade CBMs including web-based programs or materials from a curriculum. The most commonly identified premade CBMs included Renaissance Star, AIMSWEB, and DIBELS. Approximately half of participants (N=27) reported they develop their own teacher-made CBMs (see Table 7).

Table 7. Types of CBMs used by Participants

	Yes	No	Total
Premade CBMs	48	4	52
Teacher-made CBMs	27	25	52

Figure 16



Teachers reported a wide range of time spent engaged in assessment related activities (i.e. less than 10% to 50%). On average, general education teachers reported spending 19% of instructional time engaged in assessment related activities. Additionally, the average amount of time spent on assessment related activities was 26% for special education teachers, 14% for literacy specialists, 25% for math specialists, 32.5% for school psychologists, and 20% for participants who identified as 'other.'

In summary, the correlational analysis revealed a positive, moderate correlation between assessment literacy, as measured by the ALI, and knowledge about CBM ($r = .505$, $p < .01$). The relationship is significant, meaning the two instruments are measuring a similar construct, and the relationship indicates that as scores on the ALI increase, it is likely that scores on the CBM instrument will increase. The extent of the relationship could not be examined by exploring relationships between subareas of knowledge because the tools were not susceptible for this analysis. When examining the extent to which training influences knowledge about CBM, meaningful differences between groups did not appear to emerge and the

difference between groups could not be examined statistically due to a small sample size. Overall, knowledge about CBM is related to assessment literacy, as measured by the ALI.

CHAPTER IV

DISCUSSION

The primary purpose of this study was to explore to what extent teacher assessment literacy, as measured by the ALI, is related to teachers' knowledge about CBM. It is meaningful to explore this topic because teachers are becoming more involved with CBMs in the classroom in addition to other more commonly used assessments. Investigators interested in teachers' assessment literacy have investigated teachers' knowledge of assessment related activities based on the 1990 Standards for Teacher Competence in Educational Assessment of Students; however, teacher knowledge of CBM has not specifically been examined.

In the current study, teachers reported spending 10-50% of instructional time engaged in assessment related activities. There are three primary roles associated with the use of CBM: administration, scoring, and interpretation. Almost all participants (i.e. 43 out of 52) reported involvement in administering CBMs. More specifically, teachers from all positions reported involvement in administering CBMs, which means that the administration of CBM is not limited to one teaching position. Approximately half (i.e. N= 28) of the participants reported involvement in scoring CBMs and participants who do not score CBMs reported that specialists, special education teachers, and computer programs complete scoring in their school. Finally, 36 out of 52 participants reported involvement in the interpretation of CBM results. Similar to administration and scoring, interpretation is not limited to one teaching position. The small number of participants who reported involvement in scoring, in comparison to the amount of

professionals who administer and interpret CBMs, is likely a result of computer-programs that complete scoring for teachers. Further, a majority of teachers reported using premade CBMs, and approximately half of the teachers reported using teacher-made CBMs. Clearly, teachers are using CBM in the classroom.

A greater understanding of strengths and weaknesses in teachers' knowledge may inform appropriate training and supports. Additionally, teachers' involvement in CBM can vary between schools and districts, and as a result training should be focused on the role most appropriate for individual teachers. This study is the first to explore the relationship between assessment literacy and knowledge about CBM.

Examining the relationship between instruments

The first research question involved exploring if a correlational relationship existed between assessment literacy and knowledge about CBM for the participants. The ALI is purported to measure teachers' knowledge of assessment based on the 1990 Standards for Teacher Competence in Educational Assessment of Students. The CBM instrument was designed to measure multiple aspects of CBM to provide one general indication of CBM knowledge. Since both instruments are measuring teacher knowledge about assessment, it was hypothesized that the two instruments would be related with moderate strength. As expected, the current study found a positive, moderate correlation between the measures ($r=$; $p < .01$). In other words, the two instruments are likely measuring a similar construct. Additionally, specific relationships were predicted to emerge based on Stiggins, Arter, Chappuis, and Chappuis, (2004) conceptualization of assessment literacy into five components. In the current study, the tools were not amenable to this examination because with the CBM instrument the items forming the components were not found to have exclusive

membership in only one component. This means that items were not mutually exclusive to one component; rather, some items were related to multiple components. Overall, the instruments are related, but when items are categorized into previously identified subareas (i.e. Stiggns, Arter, Chappuis, & Chappuis 2004), those subareas were not differentially related as expected. An examination of the internal consistency of the instruments provided further insight into the relatedness of items and predicted subareas of knowledge (see Appendix F for a correlation matrix containing the inter-item relationships for the CBM instrument).

The internal consistency of each instrument was below the generally acceptable level of .7; however, this was expected due to the similar focus of the instruments (Nimon, Zientek, Henson, 2012). In the current study, the Kuder Richardson coefficient for the ALI was r_{KR20}= .525. Similarly, the Kuder Richardson coefficient for the CBM measure was r_{KR20}= .469. One reason instruments may not reach the desired level of internal consistency is because the instrument may be measuring more than one construct (Tavakol & Dennick, 2011). Although the ALI and the CBM instruments are measures of teachers' knowledge about assessment, both instruments intend to measure the breadth of competencies as an indicator of overall assessment. In other words, the ALI and the CBM instrument may be more appropriately referred to as general outcome measures (Hosp, Hosp, & Howell, 2007). That is, the instruments measure a sample of multiple components to provide a general indication about the construct assessment knowledge, or assessment literacy. Accordingly, the internal consistency reliability coefficients obtained in the current study are adequate based on the nature of the instruments. Despite the indication that

both instruments are measuring multiple constructs, as expected, strengths and weaknesses in teachers' understanding of subcomponents of assessment could not be examined. Three possible reasons for the inability to examine a correlational matrix are described below.

First, the instruments used in the current study measured a small sample of each component and scales with a small amount of items (i.e. < 10 items) are typically not representative of a construct (Tavakol & Dennick, 2011). In contrast to general outcome measures, mastery measures assess specific subskills in depth. Mastery measures for each component may be more appropriate for investigating areas of strength and areas for improvement in teachers' understanding of assessment. Accordingly, conclusions could be drawn about specific skills rather than general knowledge about assessment. Additionally, mastery measures may be particularly informative for CBM because teachers' involvement varies between schools and districts. In the current study, 43 out of 52 teachers reported involvement in administration, 28 out of 52 reported involvement in scoring, and 36 out of 52 reported involvement in interpretation. Moreover, teachers may be involved with screening, but not progress monitoring or the use of CBM within a problem-solving model. This means that it may not be necessary for teachers to be 'experts' in all areas and as a result mastery measures would be a more informative method to understanding teacher knowledge. A few studies examining mastery of subskills have been explored. For example, Gotch and French (2013) examined teachers' knowledge specific to the component sound assessment, which included information related to externally produced, standardized instruments (i.e. CBMs included in curriculum

packages and state assessments). In addition to results revealing specific areas of strengths (i.e. concept of median score) and weaknesses (i.e. interpretation of Z-score), results also demonstrated inadequate internal consistency of the instrument. It is unclear why low internal consistency was obtained. More research is necessary to determine an effective way to measure teachers' subskills in assessment.

Second, some assessment competencies appear to be interrelated, although it is unclear how. Stiggins, Arter, Chappuis, and Chappuis, (2004) argue that two of the competencies, clear purpose and clear targets, lay the foundation for understanding the remaining competencies. In other words, without a clear understanding of those two components, it is more challenging to understand other aspects of assessment. This is one possible reason the competencies appear to be interrelated, or not independent from one another, in the current study.

Finally, rather than grouping items, or examining subskills based on the components of assessment literacy, an alternative classification system may be more appropriate. Arter (2006) argues that teachers are implementing effective practice, but they lack a conceptual framework to organize their current understanding and to incorporate professional learning. Similarly, DeLuca and Klinger (2010) used factor analysis to identify knowledge domains (i.e. components) on a questionnaire estimating teachers' assessment literacy by measuring teachers' confidence levels. Results revealed three knowledge domains with adequate internal consistency: practice, theory, and philosophy. In contrast, another study using factor analysis demonstrated assessment conceptualized into a framework with three components: format, purpose, and use (Brown, Lake, & Matters, 2011). Based on these studies, an

alternative conceptualization of assessment may be a more appropriate fit to the instruments, particularly the CBM instrument, used in the current study. This approach of alternative conceptualizations may be fruitful for use in future investigations of teachers' assessment skills and knowledge.

Examining relationship between training in assessment and knowledge about CBM.

The secondary research question involved investigating to what extent teacher training in assessment was related to knowledge about CBM. Due to the sample size obtained in the current study, statistical comparisons were not made between groups. Instead, data were examined descriptively and meaningful differences did not appear to emerge. It is important to note that the small sample size may contribute to the lack of meaningful results. Mean scores on the CBM instrument did not differ by more than 5 out of 47 items when mean scores were compared based on training variables including completion of a preservice course on assessment, attendance at an inservice training on CBM, the number of years teaching, and the total number of courses and/or trainings completed on CBM.

Mixed results about the relationship between training and assessment literacy have been reported in the literature. Consistent with the current study, Gotch and French (2013) investigated two training variables including number of years teaching and attending professional development training, and found that both variables were not related to knowledge about assessment. Conversely, studies have found that teachers with greater than seven years of experience in the field had higher levels of assessment knowledge than teachers with less than seven years of experience in the field (Alkharusi, Karem, Al-Musawai, 2011), teachers who completed a preservice course in assessment had higher levels of assessment knowledge than teachers who

did not complete a preservice course (Alkharusi, Karem, Al-Musawai, 2011; DeLuca & Klinger, 2010), and teachers with some training, or exposure, to assessment had higher levels of assessment knowledge than teachers without any training, or exposure, to assessment (Plake, Impara, & Fager, 1993; Volante & Fazio, 2007). Therefore, it is unclear if training is related to knowledge about CBM.

Limitations

While a relationship between teachers' knowledge of assessment literacy and teachers' knowledge of CBM clearly emerged, as expected, there were limitations to the research. An overall limitation to the research involves the small sample size, which limits the generalizability of the study, because the population is not adequately represented. According to Nimon, Zientek, and Henson (2012), a small sample size may reduce the strength of a correlational relationship. Additionally, the small sample size prohibited the investigation of the relationship between subcomponents of assessment and statistical analyses of the relationship between training and teacher knowledge of CBM. Further, the majority of the sample consisted of general education teachers, so comparisons could not be made between various teaching positions. Although the small sample size limited some aspects of the research, this was an exploratory study and provided valuable information that a relationship likely exists between the instruments. Replications of this study with larger sample sizes are needed to confirm the relationship and to generalize the results beyond the current sample.

Another limitation of the current research is the varying levels of training prior to teaching in the field. Prior to data collection, completion of a preservice course in

assessment or educational measurement was inclusionary criteria for participation; however, data revealed that almost half of the respondents did not complete a course. Completion of a course was considered inclusionary criteria because it may not be appropriate to expect teachers' to have knowledge and skills about assessment without having first received training. Further, receiving formal training (i.e. teacher education program) may differ than receiving informal training in the field (i.e. learning through experience and collaboration with colleagues). Formal training provides a foundation of knowledge to build on in the field. Therefore, it is unclear if the expectation for being 'assessment literate' could be the same for individuals who have completed a preservice course in assessment and individuals who have not completed a course in assessment. The expectation for 'assessment literate' would need to be more clearly defined.

Finally, it is not possible to determine if the observations (i.e. scores on instruments) were independent from one another. In other words, it is unclear if teachers completed the instruments independently, or if they worked with colleagues to complete them. Further, teachers within a school building may have been exposed to similar trainings and information about CBM offered in their school. Therefore, it cannot be concluded if some teachers have common influences that affected their outcomes on the instruments.

Implications for Practice

The results of the current study have two major implications for practice. First, the mean scores on the ALI and the CBM instrument in the current study were consistent with mean scores on the ALI in prior research. This means that teachers are

consistently answering approximately 56-68% of the items correctly. The question that remains is whether that level of understanding is adequate for practice, or if a higher number of items answered correctly would qualify teachers as assessment literate. Although the ALI has been used frequently in the literature, it is unclear what specific scores mean. As general outcome measures, the instruments provide a general indication of what teachers know about assessment, but without cut-offs, scores are less meaningful. In the current data-based accountability reform, the expectation for teachers to use assessment tools both accurately and effectively has increased; however, is it necessary for teachers to be experts in assessment? In a manner similar to a screening measure, it may make sense to assign cut-off scores to classify levels of teachers' understanding of assessment. This could inform practice if teachers could receive appropriate, individualized supports to feel more confident with the use of assessments. Furthermore, teachers may not need to be experts in all aspects of CBM. The definition of being literate in CBM may depend more specifically on the role teachers have, and further this may vary among schools and districts.

Second, results demonstrated minimal differences in the mean CBM scores based on the number of years of teaching experience. Due to the small sample size these results are interpreted with caution. A linear relationship was expected to occur, meaning the mean score would increase as the number of years increased. Interestingly, differences in the mean scores (i.e. approximately 1-3 items) did not appear meaningful. In fact, teachers who recently graduated from teacher education programs (i.e. less than two years of experience in the field) and teachers with greater than ten years of experience in the field exhibited slightly higher levels of CBM

knowledge than teachers with 2-10 years of experience. One possible explanation for the observed lack of a linear relationship in CBM knowledge is once teachers are in the field their role shifts from being a professional learner (i.e. a student) to the day-to-day activities of a teacher. Teaching is a multifaceted profession that requires knowledge of numerous areas in addition to assessment. This result has implications for professional development. Effective professional development is necessary to maintain, or increase, teachers' level of knowledge while in the field. It is unclear what is the most effective method of professional development; however, more recent studies are exploring this area. For example, Koh (2011) found that ongoing, sustained professional development more effectively increased teachers 'assessment literacy, as measured by student work samples, scoring rubrics, and focus group interviews, than a short-term, one-shot workshop. In another example, Mertler (2009) found that a two-week professional development workshop based on the 1990 Standards for Teacher Competence in Educational Assessment of Students significantly increased teaches scores on the ALI. Finally, providing a conceptual framework for sound assessment practice may assist teachers with organizing their knowledge about assessment (Arter, 2006). Professional development is important for maintaining teachers knowledge about assessment, and it is even more important to ensure that effective professional development is provided.

Future Directions

In this study, the relationship between teachers' assessment literacy and teachers' knowledge about CBM was examined. A positive, moderate correlation was found regardless of the small sample size. Future research could replicate the current

study with a larger sample size to determine if the relationship generalizes beyond the current sample. Furthermore, research can examine differences in assessment literacy and knowledge about CBM based on different teaching positions (i.e. general education, special education, specialists, and school psychologists).

Additionally, future research could examine to what extent these two areas of knowledge are related by investigating alternative conceptualizations of assessment. For example, rather than conceptualizing assessment based on the 1990 Standards for Teacher Competence in Educational Assesment of Students or Stiggins, Arter, Chappuis, and Chappuis (2004), other conceptualizations may be more appropriate for measuring knowledge, such as DeLuca and Klinger (2010) who identified practice, theory, and philosophy as domains of assessment. This line of research could have practical use because identifying specific strengths and areas for improvement would inform training programs and more individualized targets for professional development. Additionally, future research could improve the reliability of the ALI and CBM instrument as general outcome measures. Furthermore, validating measures for the purpose of mastery measurement is an area for future research. These two types of measurement have different purposes, and to more accurately assess knowledge about assessment, it is important to know and understand exactly what is being measure so results can inform practice.

Another suggestion for future research is to explore alternative methods for measuring teachers' knowledge about assessment. A test may not be comprehensive enough to capture a construct such as knowledge about assessment. Other forms of assessment such as performance assessment and observational assessment may

provide further information. For example, Wayman et al. (2011) used a think-aloud approach to measuring teachers' knowledge about progress monitoring in which teachers' were asked directed questions, their responses were coded by subcategories and then an overall score indicated low, middle, or high level of understanding the interpretation of progress monitoring data.

Finally, future studies could further examine the relationship between training and teachers' knowledge about CBM. More specifically, because teachers are differentially involved in CBM, the capacity in which they are practicing CBM in the field may moderate the effect of training.

In summary, the results revealed a moderate relationship between assessment literacy, as measured by the ALI, and knowledge about CBM. These findings suggest that the instruments are measuring similar constructs. The results of the present study contribute to the literature given that, to date, the assessment literacy research has not concurrently investigated teachers' knowledge of CBM

PRE-/POST-TEST ANSWER KEY AND SCORING CRITERIA

If we all follow these guidelines we will be more reliable:

1. **When in doubt, do not award a point.**
2. **Do not infer meaning that may not be there, use some judgment, but go with your first instinct rather than analyzing what the person meant to say.**
3. **Look for repetition within answers. People often say the same thing 2 or 3 different ways--award points for a concept one time within any answer.**

1. **For each of the steps of the Problem-Solving model, write about its purpose and the question(s) to be answered within each step (15 points, 1 for each answer). Note that some decisions have multiple purposes and questions to be answered.**

Decision	Purpose	Question(s) to be answered
Problem Identification	1. To decide whether a student's performance on academically important tasks is discrepant enough from expectations to indicate existence of a problem and to justify further assessment. **(1Point)**	1. Does a problem exist? or Is there a discrepancy between the student's actual and expected performance? **(1Point)**
Problem Certification	1. Determine the magnitude of the discrepancy. **(1Point)** 2. Determine what resources are necessary to enable the student to make progress. **(1Point)**	1. How important are the existing discrepancies between the student's actual and expected performance? AND/OR mention of certification/eligibility question **(1Point)** 2. What resources are required to help the student benefit from their education? **(1Point)**
Exploring Solutions	1. Goal setting **(.5 Point)** and intervention planning **(.5 Point)**	1. What is an appropriate long-term goal for this student? **(1Point)** 2.What is the content of the intervention? (e.g., curriculum) **(1Point)** 3.What is the process of the intervention? (e.g., instructional strategies) **(1Point)**

Evaluating Solutions	1. To monitor student progress toward goals **(.5 Point)** and make changes when appropriate **(.5 Point).** OR To determine effectiveness of an intervention **(1 point)**	1. Has this student made progress toward their IEP goal? **(1Point)** **OR** Did the intervention work? **(1 Point)**
Problem Solution	1. To determine whether a problem still exists **(.5 Point)** and whether special services are still required to meet the student's educational needs **(.5 Point).**	1. Has this student made progress toward their IEP goal(s)? **(1Point)** 2. How important are the existing discrepancies between the student's actual and expected performance? **(1Point)**

---

2. Identify 4 problems associated with the use of published, norm-referenced tests (PNTs) in special education decision making and how Curriculum-Based Measurement (CBM) may help alleviate them.

Identifying a problem and its solution are worth 1 point each.
Describing that problem and its solution are worth 1 additional point each.

Four different sets of "Problem with PNTs" followed by reasonable "How CBM may reduce" responses for full 16 points. Like this:

1. Identify problem (1Point)
Describe problem (1Point)

1. Identify how CBM may reduce (1Point)
Describe how CBM may reduce (1Point)

No more than 4 points per set! No more than 4 sets.

NOTE: Following are examples of responses, some "How CBM may reduce" responses apply to other problems as well. Use your best judgement. This is an especially hard item to score so please fill in any possible answers I left off or give me ideas...

Problem with PNTs (1 point for each response)	**How CBM May Reduce... (1 point for each response)**
1. ID: The content of PNTs may have little overlap with the students' curricula DES: making it impossible to draw accurate conclusions about student performance.	1. ID: Items for assessment are drawn from the students' own curricula. DES:
2. ID: Many of the most frequently used PNTs have questionable or unknown reliability and validity. DES:	2. ID: There is evidence of reliability and validity for CBM. DES:

3. ID: PNTs employ selection-type responses DES: indirect assessment.	3. ID: Production-type responses are employed DES: direct assessment.
4. ID: Most PNTs are tied to national or regional norms. DES:	4. ID: CBM is tied to local norms DES:
5. ID: Often used for pre-post testing for evaluation of intervention effects. DES:	5. ID: Designed for frequent testing of intervention effects. DES:
6. ID: Insensitive to change DES:	6. ID: Sensitive to change DES:
7. ID: Most have only 1 or 2 forms. DES:	7. ID: CBM has multiple forms to facilitate continuous, frequent measurement for monitoring progress. DES:
8. ID: No link to interventions DES:	8. ID: CBM links assessment and interventions. DES:
9. ID: Expensive DES:	9. ID: Inexpensive DES:

3. At least 7 essential features of CBM have been identified in the literature. Identify and describe 4 of them. (8 points)

1 point for identification and 1 point for description. Must have 4 ID/DES sets for full 8 points. 8 IDs will not earn 8 points, 8 IDs would be 4 points.

1. Tied to a problem-solving model of decision making (1)
   CBM is designed to facilitate continuity across special education decisions (1).

2. Direct measurement of performance in the curriculum (1)
   CBM is based on the major premise that assessment and decision making are referenced to the local school curriculum--content validity is crucial (1).

3. Production-type responses (1)
   Students must actually perform the behavior of concern (1).

4. Valid and reliable indicators of academic performance (1)
   Many studies have determined that CBM are valid and reliable measures of student performance in the basic skill curricula (1).

5. Cost efficient in terms of time and money (1)
   Special education teachers can carry more responsibility for assessment in less time than it took to utilize fewer published achievement tests. School psychologists can engage in other activities such as consultation, direct interventions and teacher training. Less money is spent on test materials (1).

6. Designed for simple, repeatable administration (1)
   Curriculum-based measures are short duration, multiple form measures. Pupil performance can be viewed across several days at any stage of the decision-making process (1).

7. Designed to determine the effectiveness of instruction and monitor student progress (1)
   Pupil progress is viewed under standardized conditions across time. The functional relationship between student outcome data and instructional interventions can be established through time series analysis (1).

8. Local norms (1)
   Local norms operationalize expectations and facilitate decision making. Local norms are easy to establish using CBM because of their ease of administration and time efficiency (1).

9. Standardized administration (1)
   Description of how CBM is standardized...(1)

There may be other responses that are acceptable. Descriptions may vary, use your judgement.

---

4. Local norms can be developed at the classroom, school, and district/county level. Identify three appropriate uses for each of the different levels of local norms. (9 points)

Level of Norm	Best Uses
Classroom **(no more than 3 points available)**	1. Initial problem identification of referred students, goal setting, and problem solution , progress monitoring--**(1 point available for mention of activities in each of the problem-solving steps with the exception of problem certification)** 2. Assisting teachers and MDTs in prioritizing students' needs (1) 3. Program Evaluation (1) 4. Instructional Planning (1) 5. Get feet wet (1)
School **(no more than 3 points available)**	1. Initial problem identification of referred students (1) 2. Assisting teachers and MDTs in prioritizing students' needs (1); 3. Writing goals and objectives for monitoring progress (1); 4. Periodic and annual reviews (1) 5. Program evaluation (1)

School District or County **(no more than 3 points available)**	1.Problem certification of referred students (1); 2.Assisting teachers and MDTs in prioritizing students' needs (1); 3.Writing goals and objectives for monitoring progress (1); 4.Periodic and annual reviews (1) 5. Program evaluation (1)

5a. It has been observed in the professional literature that the prevalence of mild disabilities varies considerably from state to state, school to school, and district to district. Some critics claim that CBM and Problem-Solving Assessment may contribute to or encourage these differences. Explain the current variances in prevalence rates from a disability or person-centered model. (4 points)

1 point for an explanation of the disability model as "organic cause"
A disability or person-centered model suggested that a person's attempts to meet specified expectations are limited due to some condition of the individual. This disability has historically meant a limitation based on an identifiable organic cause.

1 point for elaboration of disability model
Therefore, an inability to read, compute or write might be attributed to organic causes (visual, hearing, brain damage, etc.) whether the etiology was directly observable and measurable or inferred from indirect measures.

The disability model assumes that:

1) the same disability may have different causes
2) some of these are organically based
3) some are not organically based, and
4) the distinction between the two can be made

However, distinguishing between organic and nonorganic causes is not so clear cut. The interactionist view is probably a more realistic model: human abilities result from a combination of biological (organic) causes as well as environmental, ecological contributions. Therefore, prevelance rates from a disabilty model are variable due to faulty assumptions and a lack of distinction between the organic and nonorganic causes of low academic achievememt.

1 point for each reason for variance mentioned--2 points available (following are examples, not an exhaustive list)
Variances in prevalence rates under the disability model could result from any of these:

•As additional disability conditions are "identified" and strong advocacy groups form, increased numbers of students are identified (i.e. ADHD, LD).

•The organic impairment approach suggests that problems cannot be solved, just managed.

•High levels of inference lead to inconsistent diagnoses.

•There are so many definitions of single disabilities that consistant diagnosis is difficult (e.g., 88% of "normal" students could be labeled LD using one of the current definitions (p. 28). OR something about how states differ in criteria for certification.

•Local SES (etc.) differ from national norms so prevalence in communities can vary widely.

5b. Explain the current variances in prevalence rates from a handicap or situation-centered, problem-solving model. (4 points)

1 point for an definition of the situation-centered model as discrepancy between expectations and current performance.
In the situation-centered, Problem-Solving model, a problem exists when the current behavior does not meet the expected level of behavior.

1 point for elaboration of situation-centered model
Therefore, a problem may occur within specific situations but not in others. They are situationally defined. Variances in prevalence rates occur when a person can meet the expectations in one situation but not another. Expectations are defined locally. A student may be handicapped in one classroom but moves to another classroom within the same school and is not handicapped. A student may be considered learning disabled in their current school district but not disabled in the district they just left.

1 point for each reason for variance mentioned--2 points available (following are examples, not an exhaustive list)
•Handicap is the normative difference in learning from cultural imperatives (p. 9-10). Cultural imperatives and significant differences vary from place to place.

•Situationally defined problems are set up to be solved and, therefore, have a better chance of being solved.

•Reliance on local norms to define problems lead to variances in prevalence rates from one community to the next, but should result in consistent decisions within any one environment.

•Teachers vary in the levels of intervention they are capable of implementing. Therefore problems are defined/solved differentially even within the same building.

•Some curricula arebetter than others. A poor curriculum creates more problems in learning and may result in more children being referred for what appears to be learning problems.

---

6. CBM is predicated on the use of short duration, fluency measures. What are the advantages of this kind of evaluation metric? (4 points)

Fluency of responding is an important component in many areas of academic achievement as well as in other important behaviors. **(1 point for some valid explanation of fluency)**

CBM offers a set of procedures for monitoring a student's progress that are: **(1 point for each advantage listed, up to 3 points)**

1) Tied to the student's curriculum
2) Short in duration thereby facilitating frequency of administration
3) Capable of repeated measurement with multiple forms
4) Inexpensive in terms of cost of materials and personnel time to create and administer
5) Sensitive to student progress over time
6) Reliable and valid
7) Production responses with direct measures of the behaviors of concern
8) Useful in making decisions within a problem-solving model

7. The following data was collected for Diane, a 4th-grade student who was referred by her teacher for performing "far below grade level" in reading. The teacher had no concerns about Diane's performance in other academic areas.

Results of CBM Survey-Level Assessment in Reading

Level of Reading Curriculum	Median WRC*	Percentile Rank**
Grade 4	44	3
Grade 3	40	14
Grade 2	65	64
Grade 1	92	99

*WRC=Words Read Correctly **District norms

a. Are additional resources warranted to resolve the reading problem for Diane? Provide a rationale for your answer. (3 points)

Yes, this reading problem may warrant additional resources **(1 point)**.

(2 points for some combination of the following)
Diane's median WRC is below the cutting score for problem certification (1). Referred students may be deemed eligible for special education services if their normative level (the point in the general education curriculum at which the student's scores on CBM are similar to those of typical students) falls 2 or more years below their current placement. This score translates functionally to performing below the 16th percentile one grade-level below grade placement on curriculum-based measures. Diane's median WRC in 3rd-grade materials fell at the 14th percentile compared to 3rd-grade students in her district (1).

An instructional placement logic would also be acceptable. Something about the suggested instructional placement of Grade 1-2 material being 40-60 WRC and Grade 3-6 material being 70-100 WRC (1). Diane's median scores suggest that her instructional level is 2 years below her placement (1).

b. If the student meets exclusionary criteria under P.L. 94-142, what variables might influence your decision regarding whether special education is needed? (3 points)

The MDT must determine whether the difference between Diane's actual and expected performance is important enough that special services may be required to resolve the problem. Several variables may be considered:

- Available general education resources that may improve student performance such as: a lower reading group (1); Chapter I (1), peer tutoring (1).
- Evidence regarding the effectiveness of interventions in the general education environment (1).
- Analysis of student errors may lead to identification of specific skill deficits that are limited (although important) and remediable in the general education setting (1).
- Analysis of the instructional environment may lead to identification of variables related the interaction between student behavior and teacher behavior that could be modified and lead to improved student performance (1).
- Priority ratings established by the MDT (1).

8. Why should we change our strategies for writing IEP annual goals? Provide 3 reasons. (3 points)

Underlined portions are key phrases to look for. Mention of any of these phrases/ideas is worth 1 point. Again there may be more acceptable answers.

•Goals are either too vague/global or too specific and numerous. Vague/global goals have no clear, measurable outcome; very specific and numerous goals are cumbersome/overwhelming. Neither extreme is easily operationalizable.

•We need to have reintegration as the ultimate goal. IEP goals should be based on expectations in the general education setting.

•A long-term measurement perspective allows for meaningful progress monitoring. Goals should be written using this perspective with decision rules in place for short-term goals. This allows for a formative approach rather than a summative approach to assessing student progress. (see #9 answers)

•Goals should be operationalizable, therefore they should include: conditions, expected behavior, and criterion.

•Goals should be ambitious because ambitious goals lead to better outcomes for students.

9. Given that teachers work with students daily and collect lots of information, why should we change how we monitor progress in the basic skill areas? Provide 3 reasons. (3 points)

NOTE: Two approaches can be taken when answering this question. First, problems with short-term, or mastery measurement, could be discussed and exemplified. Or second, advantages of long-term measurement could be discussed and exemplified. Either approach should earn points.

APPROACH #1:

Typically, teachers collect short-term or mastery measurement type data to monitor progress . **(some mention of this overall issue is worth 1 point)**

Problems with mastery measurement include:
(1 point available for each reason, up to 3 points)

1. Measurement material is changing constantly according to the curriculum skills sequence ("measurement shifts"). No opportunity to evaluate alternative skill sequences (1).
2. Measurement is often idiosyncratic with unknown reliability and validity (1).
3. Not useful for answering broad questions about student growth over time or comparing benefits of available instructional strategies (1).
4. May encourage teaching to the test so that students can progress through the curriculum (1).
5. No ongoing assessment of skill retention and generalization (1).
   May discourage teachers from reviewing skills materials sufficiently to promote skill maintenance (1)

OR

APPROACH #2

Long-term measurement is a more desireable approach to monitoring student progress
(some mention of this overall issue is worth 1 point)

Advantages of long-term measurement include:
(1 point available for each reason, up to 3 points)
1. Measurement correspnds more closely to the teminal behavior or ultimate desired outcome.
2. Measurement material remains consistent throughout the year. Little time creating testing material during progress monitoring.
3. Has greater concurrent validity with respect to other measures of achievement.
4. Growth on long-term measures corresponds to growth on other achievement measures.
5. Measurement system does not determine instructional content; measurement and instruction no confounded.
6. Assesses retention and generalization.
7. Sensitive to instructional effects which can be used for instructional decision-making.
   a. Effectiveness of alternative instructional programs implemented with a student at different times during the year
   b. Rates of improvement under diffeerent service delivery arrangements
   c. Program efficacy for different students

10a. Review the following data and fill in the blanks below: (7 points)

Results of CBM Survey-Level Assessment for Kasey (Grade 3)

Level of Curriculum	Median WRC	Peer Fall Norms	Peer Spring Norms
3.2	32	88	108
3.1	36	–	–
2.2	37	74	105
2.1	44	–	–
1.2	58	6	62

*WRC=Words Read Correctly

Weeks until annual review of IEP:	34
Curriculum	ABC Reading Series
Current instructional placement:	3.1
Appropriate instructional placement:	**2.1**
Measurement material selected:	**3.1 or 3.2**
Criterion for success: **140 WRC per minute with 6 or fewer errors**	anywhere in the range of **100 to**

IEP goal for Kasey:

In **34 (+ or - 4)** weeks, when given a randomly selected passage from **Level 3.1 (or 3.2)** of the ABC Reading Series, Kasey will read aloud at **108 (or in range of 100-140)** words read correct per minute with **6** or fewer errors.

10b. Write your rationale for selection of the long-term or annual goal measurement material and criterion for success. (2 points)

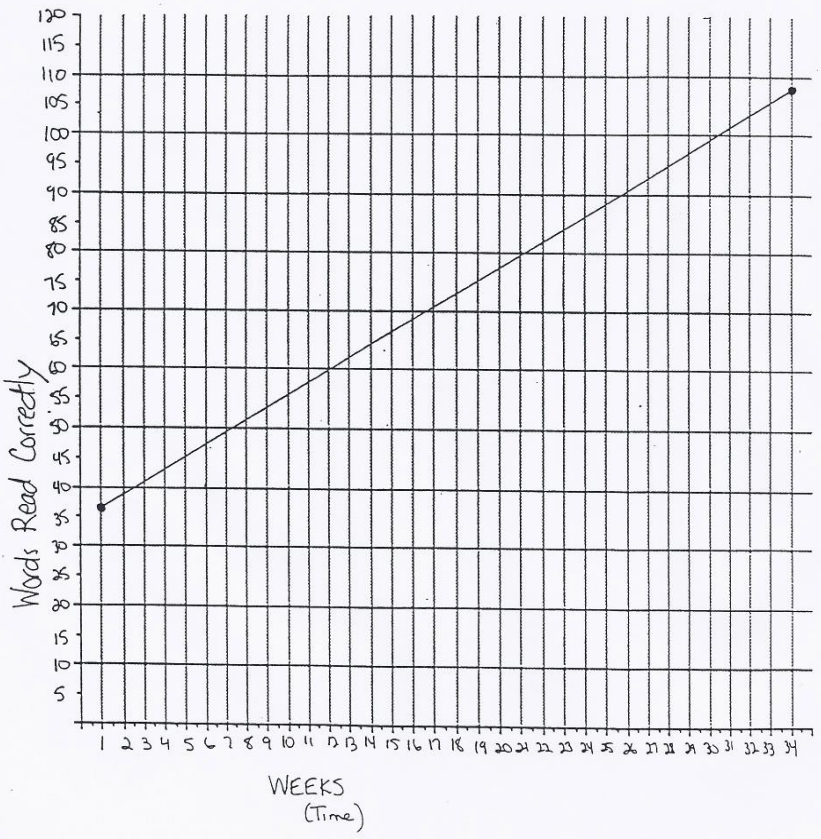
(1 point for rationale about measurement material)
The measurement material of 3.1 (or 3.2) was selected because it is at least one curriculum level above Kasey's instructional placement (expert judgment strategy) (1) and because it corresponds to the curriculum currently used in Kasey's general education classroom (same-grade norms used to set criteria) (1).

(1 point for rationale about criterion)
The criterion for success of 108 WRC per minute corresponds with the spring normative score of Kasey's peers (same-grade norms strategy) (1). The criterion of 6 or fewer errors corresponds to instructional placement standards for grades 3-6 material .
OR
Goal selected represents a 2-3 word per week increase (i.e., 2 word per week increase: 2 x 34 = 68 + current rate of 32 or 36 = 100 or 104 WRC as goal) (1)
OR
100 WRC goal because it is the top of the range of WRC suggested for instructional placement in the goal material (1).

11. **Using the IEP goal above, develop a graph for monitoring Kasey's reading progress. (5 points)**

a. Label the horizontal axis. (1)
b. Label the vertical axis. (1)
c. Plot the initial data from the long-term goal material. (1)
d. Plot the long-term goal at the appropriate point on the graph. (1)
e. Draw the aimline. (1)

If the graph is done correctly and represents the goals they set, it is given credit--even if the goals set were not correct.



12. At a student's annual review, data from the student's rates of progress towards the IEP annual goal and whether the discrepancy from grade-level peers is reduced are discussed. Four outcome combinations are possible. (For example, Cell A means that the student made progress towards the annual goal and reduced the discrepancy from peers.) Write about the decision that would be made given the outcomes from each cell with respect to whether the program was effective or ineffective and whether it required modification. (8 points)

Reduced Discrepancy

Progress Towards IEP Annual Goal

	Yes	No
Yes	Cell A	Cell B
No	Cell C	Cell D

Cell A:
Program Effective? Yes (1)
Implication? Reconsider this student's needs. Does this student continue to need additional resources? If so, is the present level of service appropriate? (1)

Cell B:
Program Effective? Uncertain (1)
Implication? The goals for this student may not be ambitious enough (1).

Cell C:
Program Effective? May be effective (1)
Implication? The goals for this student may be too ambitious (1).
OR
Reference group may be deteriorating--something wrong in classroom (1)

Cell D:
Program Effective? No (1)
Implication? A change in program is warranted (1).

---

13. Identify 3 criticisms of CBM and write a brief response to each of them. (6 points)

Maximum of 3 points available for identifications and maximum of 3 points available for logical responses. There are many more correct responses...

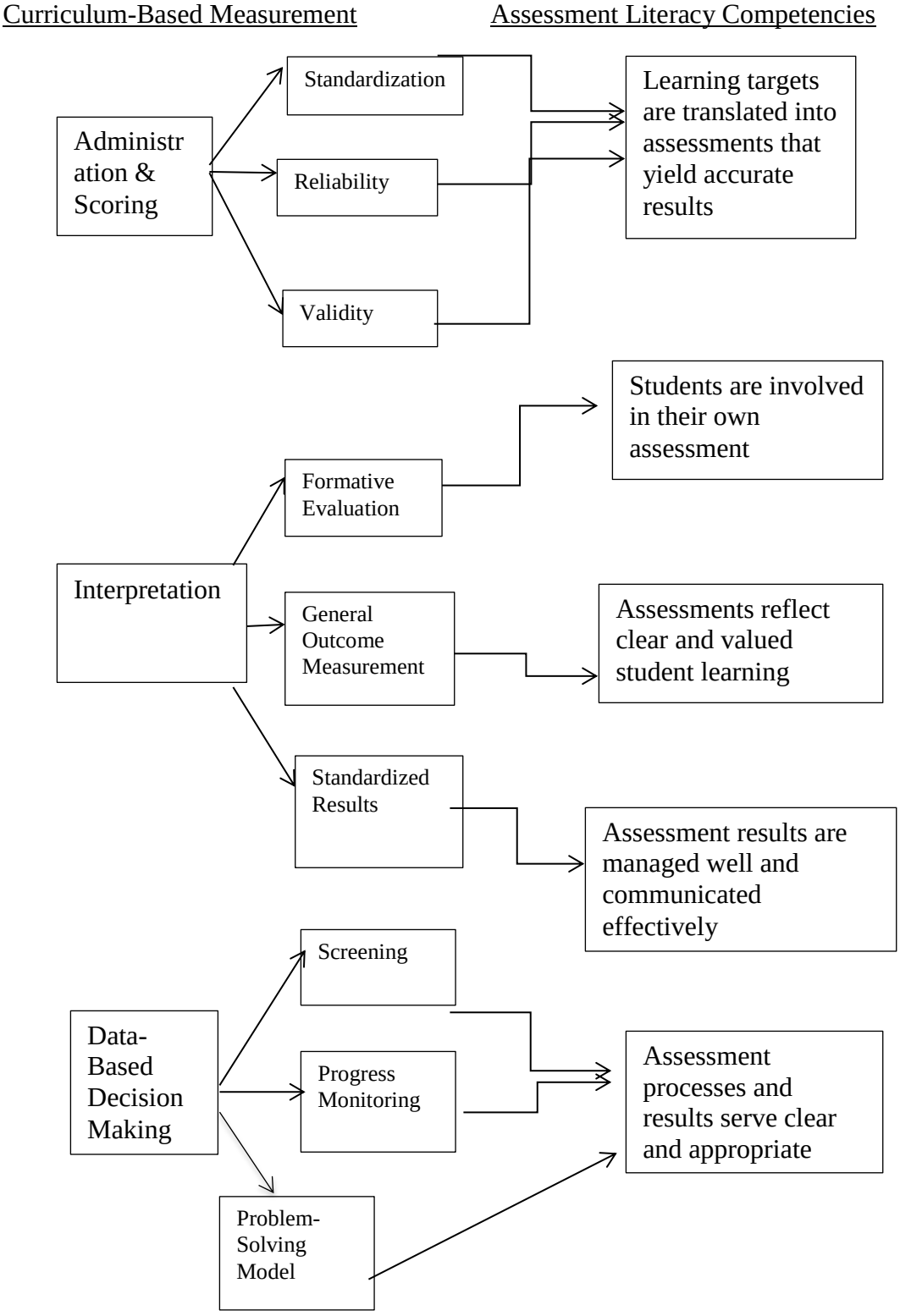
1. ID: Takes too much time to develop norms or monitor progress of individual students 2x/week. (1)
   RESPONSE: (1)

2. ID: Too simplistic (1)
   RESPONSE: (1)

3. ID: Just a reading rate--no comprehension (1)
   RESPONSE: (1)

4. ID: Limited sample of behavior (1)
   RESPONSE: (1)

Appendix B: Framework for Conceptualizing Subdomains of CBM



Appendix C: Final Adapted CBM Instrument

Below is an assessment instrument consisting of multiple choice questions. Please read each question carefully, some are single response questions, some are multiple response (check all that apply), and some require you to use provided information to answer the proceeding questions. This survey is focused on academic curriculum-based measurement that teachers may use in the classroom. Curriculum-Based Measurement (referred to as CBM) may include, but is not limited to AIMSWEB, DIBELS, FAIP-R, easyCBM, Renaissance Star, and teacher-made CBMs . Please answer the following questions based on your knowledge of CBM acquired from courses, trainings, and experience in the field, and refrain from searching for the answers on the internet or discussions with colleagues. This instrument is for general research purposes and results will not be distributed to anyone in your district. Some of the questions were designed to be challenging, so do not be concerned and please give your best possible response. Thank you in advance for your contribution and assistance in this study.

To help schools make effective data-based decisions, a systematic problem-solving process is used. The following questions on this page will ask you to think about this problem-solving process.

Which is the most appropriate order of stages in the problem solving process?

- ❍ Problem Identification, Problem analysis, Plan Development, Plan implementation, Plan evaluation
- ❍ Plan evaluation, Problem Identification, Problem Analysis, Plan implementation, Plan Development
- ❍ Problem Identification, Problem Analysis, Plan Evaluation, Plan development, Plan implementation
- ❍ Problem Analysis, Problem Identification, Plan development, Problem Evaluation, Plan implementation

What is the purpose of the problem identification stage?

- ❍ To determine if items for assessment are drawn from the local curriculum
- ❍ To set goals and plan intervention
- ❍ To decide whether a student's performance on academic tasks is discrepant enough from expectations to indicate a problem exists
- ❍ To determine the magnitude of a problem

What is the purpose of evaluating the effects of a plan (i.e. intervention or change in instruction)?

- To monitor student progress toward goals and determine if growth is occurring
- To tie results to local norms
- To determine whether a problem continues to exist and if support services are still required
- To decide whether a student's performance on academic tasks is discrepant enough from expectations to indicate a problem exists

What is the purpose of the plan development stage in the problem-solving model?

- To monitor student progress toward goals and determine effectiveness of an intervention
- To make decisions within a problem-solving model
- Describe how CBM is standardized
- To set goals and plan intervention

What is the purpose of the problem analysis stage in the problem-solving model?

- To determine the magnitude of the discrepancy
- To determine whether a problem continues to exist and if support services are still required
- To facilitate continuity across special education decisions
- To validate the assessment measure

What is the purpose of the plan implementation stage in the problem-solving model?

- To monitor student progress toward goals and make changes when appropriate
- To determine the reliability and validity of the solution
- To determine if items for assessment are drawn from the local curriculum
- To begin using a new curriculum with the class

What question or questions does the problem analysis stage address?

- Is the intervention tied to the student's curriculum?
- What are appropriate and effective interventions?
- Did the intervention work (i.e. has the student made progress toward their goal)?
- Why is the problem occurring?

What question or questions does the plan implementation stage address?

- Is the intervention working (i.e. is the student making progress toward their goal)?
- Is the intervention valid and reliable?
- Does a problem exist?
- Is the intervention tied to the local curriculum?

What question or questions does the problem identification stage answer?

- ❍ Is there a discrepancy between the student's actual and expected performance?
- ❍ What materials are necessary to help the student benefit from their education?
- ❍ What is an appropriate long-term goal for the student?
- ❍ Is the measure reliable and valid?

Which question does the plan development stage first address?

- ❍ What materials are required to help the student benefit from their classroom education?
- ❍ What are appropriate and effective interventions?
- ❍ Is there a discrepancy between the student's actual and expected performance?
- ❍ Can local norms be established?

What question or questions does the plan evaluation stage address?

- ❍ Was the plan implemented correctly?
- ❍ Does a problem still exist?
- ❍ What is the content of the intervention?
- ❍ What is an appropriate long-term goal for the student?

CBM was developed to be a simple, efficient method for assessing student achievement in the basic skill areas. What are the advantages of this method? CHECK ALL THAT APPLY.

- ❑ Tied to a problem-solving model of decision making
- ❑ Measures mastery of specific skills
- ❑ Performance based assessment
- ❑ Accurate predictor if student will graduate high school
- ❑ Tied to national norms
- ❑ Valid and reliable indicators of academic performance
- ❑ Results used to diagnose specific skill deficits
- ❑ Cost efficient in terms of time and money
- ❑ Designed for simple, repeatable administration
- ❑ Can be used to determine the effectiveness of instruction and monitor student progress.
- ❑ Allows instructor to make inferences about student behavior beyond the behavior measured
- ❑ Use in different content areas (i.e social studies, science)
- ❑ Used to develop local norms
- ❑ Measures student comprehension (i.e. higher order thinking)
- ❑ Standardized administration

Note: Some of the questions were designed to be challenging, so do not be concerned and please give your best possible response.

CBM is tied to a problem-solving model of decision making. This means:
- ❍ CBM is designed to facilitate consistency across instructional decisions
- ❍ All students will receive the same intervention
- ❍ All students have problems and instructors are required to fix them
- ❍ Instructors are guided to make the right decisions

CBMs employ production-type responses. This means:
- ❍ Students choose correct answer from a list of responses
- ❍ The tasks students are asked to perform are similar to tasks in the curriculum
- ❍ Students must actually perform the skill of concern
- ❍ Students are asked to discuss the skill of concern

CBMs should be reliable measures of student achievement. What is the definition of reliability?
- ❍ If a student takes the assessment more than once their score will be consistent
- ❍ The assessment measures what it intends to
- ❍ The measure systematically samples the year-long curriculum
- ❍ The assessment is given at the end of the school year

CBMs should be valid measures of student achievement. What is the definition of validity?
- ❍ If a student takes the assessment more than once their score will be consistent
- ❍ The assessment measures what it intends to
- ❍ The measure systematically samples the year-long curriculum
- ❍ The assessment is given at the end of the school year

CBM is cost efficient in terms of time and money. This means:
- ❍ CBM involves many materials that are funded by the state
- ❍ Less students need to be assessed
- ❍ Instructors need less training to administer and score CBMs
- ❍ Less money is spent on test materials and less time is spent on administration than published norm-referenced tests (i.e. standardized tests such as SAT and NWEA)

CBM was designed for simple, frequent administration. This means:

- ❍ CBMs are quick to administer and results can be easily graphed to monitor progress
- ❍ CBMs are difficult to administer and score
- ❍ Frequent administration allows CBM to be used for screening students
- ❍ Frequent administration allows for comparison between students in the nation

CBM can be used to determine the effectiveness of instruction and monitor student progress. This means:

- ❍ The tasks students are asked to perform are similar to tasks in the curriculum
- ❍ CBM graphs demonstrate relationship between student achievement and instructional interventions
- ❍ CBMs are short duration, multiple form measures.
- ❍ CBM is the only data needed to make accurate decisions regarding student achievement and interventions

Standardized administration means:

- ❍ Administration procedures differ for progress monitoring and screening purposes
- ❍ Administration procedures are specified so teachers do not have to be involved in the administration
- ❍ Administration procedures are specified so you can change the instructions based on student needs
- ❍ Administration procedures are consistent across all settings

CBMs can be used to develop local norms. What are local norms?

- ❍ Local norms define guidelines for diagnosing skill deficits
- ❍ Local norms define a score that represents a desired level of performance
- ❍ Local norms define expectations for student achievement and reflect the student's learning culture and community
- ❍ Local norms define expectations for student achievement so students' scores can be compared to other students in the nation

At the classroom level, local norms can be used for which of the following. CHECK ALL THAT APPLY.

- ❑ Identifying if a student score is below expected achievement/performance
- ❑ Goal setting
- ❑ Diagnose skill deficits
- ❑ Determining if an instructional technique or program is effective
- ❑ To discipline students
- ❑ Determining if a problem still exists
- ❑ Assisting teachers and multi-disciplinary teams in prioritizing students' needs
- ❑ Choose which students should receive rewards
- ❑ Instructional planning

- ❑ Progress monitoring
- ❑ Compare student achievement to same-age peers in the nation

What is the definition of fluency?

- ❍ A measure that reflects speed/automaticity and indicates several elements of proficiency
- ❍ Assessments during instruction to inform or assessment for learning
- ❍ A measure that is based on a systematic sample the year-long curriculum so that each skill is represented
- ❍ Assessment that focuses on a particular set of skills

The short duration of CBM facilitates

- ❍ Instructional planning
- ❍ Creating national norms
- ❍ Monitoring student progress
- ❍ Reliability and validity

A general outcome measure, or long-term measurement approach, is:

- ❍ Degree to which an assessment measures what it intended to measure
- ❍ Assessment that reflects proficiency of specific skill
- ❍ Assessment that reflects overall competence of a basic skill
- ❍ Assessment conducted on a frequent basis to inform instruction

Mastery Measurement is

- ❍ A type of assessment that is based on a systematic sample of the year-long curriculum so that each skill is represented
- ❍ A type of assessment that reflects accuracy of a students level of proficiency
- ❍ A type of assessment that reflects speed/automaticity and indicates several elements of proficiency
- ❍ A type of assessment to identify which students are at-risk for academic failure

CBM is able to detect small amounts of growth over time. This is an advantage for:

- ❍ Identifying skill deficits
- ❍ Mastery measurement
- ❍ Obtaining validity of student scores
- ❍ Monitoring student progress

Progress monitoring, or formative assessment, is:

- An established process with consistency from context to context
- Assessment of learning, after instruction or a summation of what the students know and are able to do after instruction
- Assessment conducted during instruction or an intervention to inform instructional decisions
- Scientifically demonstrated consistency of a measure

In relation to assessment, the purpose of screening is to:

- Identify which students are at-risk for academic failure
- Assess on a frequency basis to inform instruction
- Determine if an intervention will be effective before it is tried
- Assess student achievement for graduation

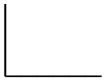
Curriculum-Based Measurement is:

- A well-researched type of general outcome measurement
- A well-researched type of mastery measurement
- A poorly researched type of mastery measurement
- A poorly researched type of general outcome measurement

What is the purpose of using CBM to systematically measure student performance at grade level and the grade levels sequentially below until the student has reached the benchmark?

- To identify students who are at risk for academic failure
- To identify students' instructional level
- To diagnose skill deficits
- To use at pre-post testing

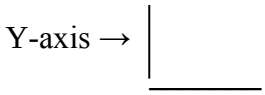
When graphing progress monitoring data for Oral Reading Fluency, what should be the title of the horizontal (X) axis?



^ X-axis

- Words read correctly
- Grade level
- Number of weeks
- Type of curriculum

When graphing progress monitoring data for Oral Reading Fluency, what should be the title of the vertical (Y) axis?



- ❍ Words read correctly
- ❍ Grade level
- ❍ Number of weeks
- ❍ Type of curriculum

What is benchmarking?

- ❍ Marking a student score on a graph
- ❍ Comparing student score to other students of the same age
- ❍ Comparing student score to a standard representing a desired level of performance
- ❍ Frequently collecting data to determine if a student is making growth

How is the median score calculated?

- ❍ Add up scores on all probes and divide by the number of probes
- ❍ Using your own judgment about which score best fits the students ability
- ❍ Put all scores in order lowest to highest and pick the middle score
- ❍ Add up scores and divide by the student's grade level

Hannah is a third-grade student who transferred to Red Elementary School late in the fall. Her teacher has noticed that she seems to struggle with many independent reading assignments. When the teacher administered the mid-year universal screening measure, she was not surprised to see that Hannah's score had fallen below the grade-level benchmark. Consequently, the teacher gives Hannah additional instruction in a small group two times per week for 30 minutes and monitors her reading performance once per week for seven weeks using a measure of reading fluency. After seven weeks Hannah's scores show steady growth in her performance, but she still has not met the mid-year benchmark. It can be concluded that:

- ❍ The small group instruction is ineffective
- ❍ The small group instruction requires a modification
- ❍ The instructional strategies used in the small group are not valid
- ❍ The small group instruction is effective and should be continued

Larry is a third-grade student at Purple Elementary School. His teacher administered a universal screening measure a few weeks after school began. Larry's score indicated that he may be struggling in reading. As a result, the school provides him with an intervention two times per week for 30 minutes and monitors his reading performance once per week for eight weeks using a measure of reading fluency. He is expected to improve by one word per minute each week. After eight weeks Larry's score

has increased by 4 words per minute. His teacher notices that Larry scored 55 words read correct for three weeks straight. It can be concluded that:

- ❍ The program is effective for Larry
- ❍ The program is ineffective for Larry
- ❍ The program should continue to be used with Larry
- ❍ The program is not valid

Use the provided data to answer the following questions. The data below was collected for Diane, a 4th-grade student who was referred by her teacher for performing "far below grade level" in reading.

Results of CBM Survey-level Assessment in Reading

Level Reading	Median WRC*	Diane's % Rank**	Fall Norms***	Spring Norms
Grade 4	44	3	94	123
Grade 3	40	14	71	107
Grade 2	65	64	51	89
Grade 1	92	99	-	53

*Words Read Correctly **District Norms *** 50th Percentile

Is Diana reading at grade-level?

- ❍ Yes
- ❍ No

Is additional information needed to plan an intervention?

- ❍ Yes
- ❍ No

On grade 3 material, Diane scored in the 14th percentile. What does this mean?

- ❍ Diane scored lower than 14% students in grade 3.
- ❍ Diane scored higher than 14% of the students in grade 3.
- ❍ Diane answered 14% of questions correctly.
- ❍ Diane received a total score of 14 out of 100

What is the fall benchmark for grade 4?

- ❍ 44 WRC
- ❍ 3 WRC
- ❍ 94 WRC
- ❍ 123 WRC

At which grade level does Diane's score meet the benchmark?

- ❍ Grade 4
- ❍ Grade 3
- ❍ Grade 2
- ❍ Grade 1

Use the provided data to answer the following questions. The data below was collected for Kasey, a 4th-grade student who was referred by her teacher for performing "below grade level" in reading.

Results of CBM Survey-Level Assessment in Oral Reading Fluency for Kasey conducted in the Fall (Grade 4)

Grade Level	Median *WRC	Kasey's % Rank**	Fall Norms***	Spring Norms***
4	45	10	94	123
3	68	48	71	107
2	79	75	51	89
1	110	99	-	
53				

*Words read correctly **District Norms ***50th percentile

Is Kasey reading at grade-level?

- ❍ Yes
- ❍ No

Is additional information needed to plan an intervention?

- ❍ Yes
- ❍ No

On grade 2 material, Kasey scored in the 75th percentile. What does this mean?

- ❍ Kasey scored lower than 75% of students in grade 2
- ❍ Kasey answered 75% of the questions correctly
- ❍ Kasey received a total score of 75 out of 100
- ❍ Kasey scored higher than 75% of students in grade 2

What is the spring benchmark for grade 2

- ❍ 79 WRC
- ❍ 75 WRC
- ❍ 51 WRC
- ❍ 89 WRC

At which grade level does Kasey's score meet the benchmark?

- ❍ Grade 4
- ❍ Grade 3
- ❍ Grade 2

Grade 1

Appendix D: Recruitment Emails

Superintendent Request to Conduct Research

Dear

My name is Paige Hamilton, and I am currently a doctoral student in school psychology at the University of Rhode Island. The research I wish to conduct for my Master's Thesis involves the exploration of teachers' understanding of assessment. This project will be conducted under the supervision of Gary Stoner, PhD (University of Rhode Island).

I am hereby seeking your consent to approach a number of elementary schools in your district to solicit teachers as participants for this project. **Participation in this project is anonymous, and individual teacher participation will be voluntary.** Teachers who choose to participate will be asked to complete two 20-25 minute online assessment instruments. The second instrument will be sent out two weeks following completion of the first instrument.

I have provided you with a copy of my thesis proposal, which includes a copy of the consent form to be used in the research process. In addition, I have attached a copy of both measures to be used, as well as a copy of the approval letter which I received from the University of Rhode Island Institutional Review Board.

With your permission, I would like to proceed as follows:

1. Obtain permission from the Superintendent to contact elementary school principals
2. Seek permission from principals to distribute information to teachers
3. Distribute information to teachers for their consideration to participate

If you require any further information, please do not hesitate to contact me at (914) 804-7621 and phamilton@my.uri.edu. Thank you for your time and consideration in this matter.

Sincerely,

Paige Hamilton, B.S.

University of Rhode Island

Principal Request to Conduct Research

Dear

My name is Paige Hamilton, and I am a doctoral student in school psychology at the University of Rhode Island. The research I wish to conduct for my Master's Thesis involves the exploration of teachers' understanding of assessment. This project will be conducted under the supervision of Gary Stoner, PhD (University of Rhode Island).

The Superintendent or Research person of your district has given permission for this research to be carried out. I am hereby seeking your consent to approach a number of teachers in your school to solicit as participants for this project. More specifically, I would like to approach general education teachers teaching Kindergarten through fifth grade, special education teachers, and specialists for this study. **Participation in this project is anonymous and individual teacher participation is voluntary.** Teachers who choose to participate will be asked to complete two online assessment instruments with a two week interval between the first and second.

I have provided you with a copy of my thesis proposal, which includes a copy of the consent form to be used in the research process. In addition, I have attached a copy of both measures to be used, as well as a copy of the approval letter which I received from the University of Rhode Island Institutional Review Board.

With your permission, I would like to proceed as follows:

1. Obtain permission from the Principal to distribute information to teachers
2. Distribute information to teachers for their consideration to participate

If you require any further information, please do not hesitate to contact me at (914) 804-7621 and phamilton@my.uri.edu. Thank you for your time and consideration in this matter.

Sincerely,

Paige Hamilton, B.S.
University of Rhode Island

Appendix E: Consent Form

Dear Teacher,

You are being invited to take part in a research study exploring teachers' understanding of assessment. In particular the study will examine teachers' understanding of assessment in general and teachers' understanding of a specific type of assessment known as curriculum-based measurement. The title of my research project is 'Toward a better understanding of teachers' knowledge of assessment: Examining the concurrent validity of tests of Assessment Literacy and Curriculum-Based Measurement'. The study is described in more detail below. If you have any questions, please feel free to call Paige Hamilton, graduate student, at 914-804-7621 or Gary Stoner, PhD at 401-874-4234, the people mainly responsible for this study.

You must be at least 18 years old and currently teaching in an elementary school to participate in this research project.

The purpose of the study is to explore the extent to which teachers' knowledge and understanding of general assessment is related to their knowledge and understanding of curriculum-based measurement.

If you decide to take part in this study, you will complete two online assessment instruments containing multiple choice questions. One assessment consists of 35 questions and will take approximately 15-20 minutes to complete. The other assessment consists of 48 questions and will take approximately 20-25 minutes to complete. After completion of the first assessment instrument, there will be a two week period before the second assessment instrument is sent to you.

The possible risks or discomforts of participating in this study are minimal, although you may feel some frustration while answering challenging questions. Following the completion of the study the answer key to each questionnaire will be sent to all participants to allow for reviewing the questions asked along with the answers to the questions.

Although there are no direct benefits of the study to you, it may prompt you to think and learn more about the use of assessments in the classroom. Your answers will help improve the general understanding of how general assessment knowledge may be related to the understanding and knowledge of curriculum-based measurement. This information could potentially lead to better methods for teaching teachers about the purpose and use of assessment in the schools.
Your participation in the study is anonymous. That means that your answers to all questions are private. I have taken measures to ensure that your answers will not be linked with your name or email. No one else will know if you participated in this study and no one else will find out what your answers were. You will not have to identify the school district or name of your school at any time. All information will be kept private, in a password-protected file, in case someone could look at your demographics and identify you. To link your responses to both instruments, you will

automatically receive a randomly generated 7 digit number from the survey website after completing the first assessment instrument. You will enter the 7 digit number at the beginning of the second assessment instrument. You are the only person that will know the number that has been assigned to you. Any scientific reports generated will be based on group data and will not identify you or any individual as being a participant in this project.

The decision to participate in the research project is up to you. You do not have to participate and if you do participate you can refuse to answer any question.

Participation in the study is not expected to be harmful or injurious to you. However, if this study causes you any injury you should email or call Paige Hamilton (914)-804-7621 or phamilton@my.uri.edu. You may also contact Gary Stoner, PhD 401-874-4234 or gstoner@uri.edu if you choose.

If you have any more questions or concerns about this study, you may contact University of Rhode Island's Vice President for Research, 70 Lower College Road, Suite 2, URI, Kingston, RI, (401)-874-4328.

By clicking 'Yes' you are implying that you are at least 18 years old, have read the consent, your questions have been answered to your satisfaction, and that you consent to participate in this study.

Thank you,

Paige Hamilton, B.S.
Doctoral Student in School Psychology

Appendix F: Correlation Matrix of CBM Items

CBM Instrument Inter-Item Correlation Matrix

	Q1	Q4	Q5	Q7	Q8	Q9	Q11	Q13
Q1	1.000	.135	.459	.153	-.007	-.064	.149	.070
Q4	.135	1.000	-.131	.060	-.053	-.209	-.316	-.147
Q5	.459	-.131	1.000	.232	.196	.401	.265	.402
Q7	.153	.060	.232	1.000	-.147	.229	-.529	.229
Q8	-.007	-.053	.196	-.147	1.000	-.319	.320	-.134
Q9	-.064	-.209	.401	.229	-.319	1.000	.151	.467
Q11	.149	-.316	.265	-.529	.320	.151	1.000	.047
Q13	.070	-.147	.402	.229	-.134	.467	.047	1.000
Q19	.070	.184	-.015	-.247	.079	-.247	.047	-.174
Q20	-.235	-.079	.265	.151	.169	.151	.100	.466
Q31	-.199	.184	-.015	-.247	.079	-.247	.047	-.174
Q32	-.101	.029	-.024	-.199	.122	-.014	.073	.186
Q33	-.025	-.189	-.020	.090	-.081	-.113	.060	.028
Q37	-.093	-.069	.189	-.116	-.163	.331	.219	.470
Q38	.411	-.069	.189	-.116	.237	-.116	.219	-.082
Q39	-.309	.029	-.186	.171	-.208	.171	-.254	.186
Q16	.149	.158	.265	.151	.017	.321	.100	.256
Q17	.270	.000	.210	.120	-.267	.478	.158	.147
Q18	.135	.250	.105	.060	.426	-.209	.158	-.147
Q21	.411	.555	.189	.331	.237	-.116	-.175	-.082
Q22	.270	.000	.210	.299	.213	-.060	.000	.147
Q24	-.135	-.100	-.010	-.167	-.235	.155	.316	.280
Q25	.411	-.069	.189	-.116	.237	-.116	.219	-.082
Q35	.304	.131	.187	.276	-.045	.276	-.265	.433
Q40	-.199	-.147	-.015	-.009	.291	.229	.256	.413
Q41	.153	.060	.063	-.157	.025	.036	.151	-.009
Q42	-.201	-.053	-.106	-.147	.080	.197	.169	.291
Q43	-.093	-.069	.189	-.116	.237	-.116	.219	-.082
Q44	.229	-.100	-.010	.155	.053	-.167	.032	.280
Q45	.264	-.169	.459	.153	.187	.153	.149	.606
Q46	.270	.000	.210	.478	.213	-.060	-.158	.147
Q47	-.078	-.341	-.258	.106	-.045	-.063	.033	.015
Q27	.007	.293	-.196	.319	-.233	-.025	-.472	-.079
Q48	-.135	-.100	-.010	.155	-.235	.155	-.253	.280
Q49	.018	-.169	.269	.371	-.201	.371	-.043	.606
Q50	.229	-.100	.273	-.167	.053	.155	.316	.280
Q51	.338	-.147	.402	.467	.079	-.009	-.163	.413

Inter-Item Correlation Matrix

	Q19	Q20	Q31	Q32	Q33	Q37	Q38	Q39
Q1	.070	-.235	-.199	-.101	-.025	-.093	.411	-.309
Q4	.184	-.079	.184	.029	-.189	-.069	-.069	.029
Q5	-.015	.265	-.015	-.024	-.020	.189	.189	-.186
Q7	-.247	.151	-.247	-.199	.090	-.116	-.116	.171
Q8	.079	.169	.079	.122	-.081	-.163	.237	-.208
Q9	-.247	.151	-.247	-.014	-.113	.331	-.116	.171
Q11	.047	.100	.047	.073	.060	.219	.219	-.254
Q13	-.174	.466	-.174	.186	.028	.470	-.082	.186
Q19	1.000	.047	.120	.186	.028	-.082	.470	-.042
Q20	.047	1.000	.256	.073	-.120	.219	-.175	.236
Q31	.120	.256	1.000	-.271	.028	-.082	-.082	.186
Q32	.186	.073	-.271	1.000	-.152	-.127	.302	-.066
Q33	.028	-.120	.028	-.152	1.000	-.105	-.105	.238
Q37	-.082	.219	-.082	-.127	-.105	1.000	-.038	.302
Q38	.470	-.175	-.082	.302	-.105	-.038	1.000	-.127
Q39	-.042	.236	.186	-.066	.238	.302	-.127	1.000
Q16	-.373	-.200	.047	-.091	.418	.219	-.175	.236
Q17	-.074	-.158	-.074	-.115	.189	.277	.277	.057
Q18	-.147	.158	.184	.029	.094	-.069	-.069	.029
Q21	-.082	-.175	-.082	-.127	-.105	-.038	-.038	-.127
Q22	-.295	.000	-.074	.057	.189	-.139	-.139	.057
Q24	-.118	.032	-.118	.126	-.151	.693	-.055	.126
Q25	.470	-.175	-.082	.302	-.105	-.038	1.000	-.127
Q35	.015	-.116	-.193	.024	-.158	.204	.204	.024
Q40	-.174	.256	-.174	.186	.279	-.082	-.082	-.042
Q41	-.009	-.189	-.247	.171	-.316	.331	.331	-.199
Q42	.291	.169	.291	-.043	.101	.237	.237	.122
Q43	.470	.219	.470	-.127	-.105	-.038	-.038	-.127
Q44	-.118	.316	-.118	-.184	-.151	-.055	-.055	-.184
Q45	.070	.341	-.199	.108	-.255	.411	.411	-.101
Q46	-.074	.316	.147	.057	.000	-.139	.277	.229
Q47	.015	-.116	.015	-.138	.376	-.189	-.189	.186
Q27	-.079	.135	.134	-.452	-.101	.163	-.237	.373
Q48	.280	.032	-.118	.126	.529	-.055	-.055	.126
Q49	-.199	.341	.070	-.101	.204	.411	-.093	.526
Q50	.280	.032	-.118	.126	-.151	.693	.693	.126
Q51	.120	.256	-.174	.186	.028	-.082	.470	-.042

Inter-Item Correlation Matrix

	Q16	Q17	Q18	Q21	Q22	Q24	Q25	Q35
Q1	.149	.270	.135	.411	.270	-.135	.411	.304
Q4	.158	.000	.250	.555	.000	-.100	-.069	.131
Q5	.265	.210	.105	.189	.210	-.010	.189	.187
Q7	.151	.120	.060	.331	.299	-.167	-.116	.276
Q8	.017	-.267	.426	.237	.213	-.235	.237	-.045
Q9	.321	.478	-.209	-.116	-.060	.155	-.116	.276
Q11	.100	.158	.158	-.175	.000	.316	.219	-.265
Q13	.256	.147	-.147	-.082	.147	.280	-.082	.433
Q19	-.373	-.074	-.147	-.082	-.295	-.118	.470	.015
Q20	-.200	-.158	.158	-.175	.000	.032	-.175	-.116
Q31	.047	-.074	.184	-.082	-.074	-.118	-.082	-.193
Q32	-.091	-.115	.029	-.127	.057	.126	.302	.024
Q33	.418	.189	.094	-.105	.189	-.151	-.105	-.158
Q37	.219	.277	-.069	-.038	-.139	.693	-.038	.204
Q38	-.175	.277	-.069	-.038	-.139	-.055	1.000	.204
Q39	.236	.057	.029	-.127	.057	.126	-.127	.024
Q16	1.000	.474	.158	.219	.474	.032	-.175	.182
Q17	.474	1.000	-.250	-.139	-.167	.100	.277	.262
Q18	.158	-.250	1.000	.555	.500	-.100	-.069	-.105
Q21	.219	-.139	.555	1.000	.277	-.055	-.038	.204
Q22	.474	-.167	.500	.277	1.000	-.200	-.139	-.052
Q24	.032	.100	-.100	-.055	-.200	1.000	-.055	.010
Q25	-.175	.277	-.069	-.038	-.139	-.055	1.000	.204
Q35	.182	.262	-.105	.204	-.052	.010	.204	1.000
Q40	.256	.147	-.147	-.082	-.074	-.118	-.082	.015
Q41	-.189	-.060	.329	.331	-.060	.155	.331	.106
Q42	.169	.373	-.293	-.163	-.267	.053	.237	.257
Q43	-.175	-.139	-.069	-.038	-.139	-.055	-.038	-.189
Q44	-.253	-.200	-.100	-.055	.100	-.080	-.055	.010
Q45	-.043	.067	-.169	-.093	.067	.229	.411	.495
Q46	.000	.000	.500	.277	.333	-.200	.277	.262
Q47	.033	-.052	.131	-.189	.105	.010	-.189	-.038
Q27	-.017	-.053	.293	.163	.107	-.053	-.237	.196
Q48	.032	.100	-.100	-.055	-.200	-.080	-.055	.294
Q49	.341	.270	.135	-.093	.270	.229	-.093	.304
Q50	.032	.400	-.100	-.055	-.200	.460	.693	.294
Q51	-.163	.147	-.147	-.082	-.074	-.118	.470	.433

Inter-Item Correlation Matrix

	Q40	Q41	Q42	Q43	Q44	Q45	Q46	Q47
Q1	-.199	.153	-.201	-.093	.229	.264	.270	-.078
Q4	-.147	.060	-.053	-.069	-.100	-.169	.000	-.341
Q5	-.015	.063	-.106	.189	-.010	.459	.210	-.258
Q7	-.009	-.157	-.147	-.116	.155	.153	.478	.106
Q8	.291	.025	.080	.237	.053	.187	.213	-.045
Q9	.229	.036	.197	-.116	-.167	.153	-.060	-.063
Q11	.256	.151	.169	.219	.032	.149	-.158	.033
Q13	.413	-.009	.291	-.082	.280	.606	.147	.015
Q19	-.174	-.009	.291	.470	-.118	.070	-.074	.015
Q20	.256	-.189	.169	.219	.316	.341	.316	-.116
Q31	-.174	-.247	.291	.470	-.118	-.199	.147	.015
Q32	.186	.171	-.043	-.127	-.184	.108	.057	-.138
Q33	.279	-.316	.101	-.105	-.151	-.255	.000	.376
Q37	-.082	.331	.237	-.038	-.055	.411	-.139	-.189
Q38	-.082	.331	.237	-.038	-.055	.411	.277	-.189
Q39	-.042	-.199	.122	-.127	-.184	-.101	.229	.186
Q16	.256	-.189	.169	-.175	-.253	-.043	.000	.033
Q17	.147	-.060	.373	-.139	-.200	.067	.000	-.052
Q18	-.147	.329	-.293	-.069	-.100	-.169	.500	.131
Q21	-.082	.331	-.163	-.038	-.055	-.093	.277	-.189
Q22	-.074	-.060	-.267	-.139	.100	.067	.333	.105
Q24	-.118	.155	.053	-.055	-.080	.229	-.200	.010
Q25	-.082	.331	.237	-.038	-.055	.411	.277	-.189
Q35	.015	.106	.257	-.189	.010	.495	.262	-.038
Q40	1.000	-.247	.503	-.082	.280	.070	-.074	.015
Q41	-.247	1.000	-.147	-.116	-.167	.153	.120	-.232
Q42	.503	-.147	1.000	.237	.053	.187	.053	-.045
Q43	-.082	-.116	.237	1.000	-.055	-.093	-.139	.204
Q44	.280	-.167	.053	-.055	1.000	.229	.100	.010
Q45	.070	.153	.187	-.093	.229	1.000	.270	-.269
Q46	-.074	.120	.053	-.139	.100	.270	1.000	.105
Q47	.015	-.232	-.045	.204	.010	-.269	.105	1.000
Q27	-.503	-.025	-.233	-.237	-.053	.007	.267	.196
Q48	.280	-.167	.341	-.055	-.080	-.135	.100	.294
Q49	.070	-.064	.187	-.093	.229	.264	.472	.304
Q50	-.118	.478	.341	-.055	-.080	.593	.100	-.273
Q51	.120	-.009	.079	-.082	.280	.606	.590	.015

Inter-Item Correlation Matrix

	Q27	Q48	Q49	Q50	Q51
Q1	.007	-.135	.018	.229	.338
Q4	.293	-.100	-.169	-.100	-.147
Q5	-.196	-.010	.269	.273	.402
Q7	.319	.155	.371	-.167	.467
Q8	-.233	-.235	-.201	.053	.079
Q9	-.025	.155	.371	.155	-.009
Q11	-.472	-.253	-.043	.316	-.163
Q13	-.079	.280	.606	.280	.413
Q19	-.079	.280	-.199	.280	.120
Q20	.135	.032	.341	.032	.256
Q31	.134	-.118	.070	-.118	-.174
Q32	-.452	.126	-.101	.126	.186
Q33	-.101	.529	.204	-.151	.028
Q37	.163	-.055	.411	.693	-.082
Q38	-.237	-.055	-.093	.693	.470
Q39	.373	.126	.526	.126	-.042
Q16	-.017	.032	.341	.032	-.163
Q17	-.053	.100	.270	.400	.147
Q18	.293	-.100	.135	-.100	-.147
Q21	.163	-.055	-.093	-.055	-.082
Q22	.107	-.200	.270	-.200	-.074
Q24	-.053	-.080	.229	.460	-.118
Q25	-.237	-.055	-.093	.693	.470
Q35	.196	.294	.304	.294	.433
Q40	-.503	.280	.070	-.118	.120
Q41	-.025	-.167	-.064	.478	-.009
Q42	-.233	.341	.187	.341	.079
Q43	-.237	-.055	-.093	-.055	-.082
Q44	-.053	-.080	.229	-.080	.280
Q45	.007	-.135	.264	.593	.606
Q46	.267	.100	.472	.100	.590
Q47	.196	.294	.304	-.273	.015
Q27	1.000	-.053	.201	-.053	-.079
Q48	-.053	1.000	.229	-.080	.280
Q49	.201	.229	1.000	.229	.338
Q50	-.053	-.080	.229	1.000	.280
Q51	-.079	.280	.338	.280	1.000

BIBLIOGRAPHY

Adams, M. (1990). Chapter 5. What needs to be taught? Hints from skilled readers. In *Beginning to Read*.

"AIMSweb National Norms Technical Documentation," (2012). Retrieved from: http://aimsweb.com/uploads/AIMSweb%20National%20Norms%20Technical%20Documentation.pdf

Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. *Asia-Pacific Journal of Teacher Education, 39*(2), 113-123.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competence in educational assessment of students. Educational Measurement: Issues in Practice, 9(4), 30-32.

American Psychological Association. (2005). *Policy statement on evidence-based practice in psychology*. Washington, DC: Author. Retrieved on September 19, 2012, from http://www.apa.org/practice/resources/evidence/index.aspx.

Ardoin, S. P., Christ, T. J., Morena, L. S., Cormier, D. C., & Klingbeil, D. A. (2013). A systematic review and summarization of the recommendations and research surrounding curriculum-based measurement or oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51*(1), 1-18.

Arter, J. (2006). Making use of data: what educators need to know and be able to do. In J. O'Reilly (Ed.) Beyond NCLB: From measuring status to informing improvement, *39-72*.

Ball, C. R., & Christ, T. J. (2012). Supporting valid decision making: Uses and misuses of assessment data within the context of RTI. *Psychology in the Schools, 49*(3), 231-244.

Betts, J., Pickart, M., & Heistad, D. (2009). An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating for alternate forms. *Journal of School Psychology, 47*, 1-17.

Bolt, D. M., Ysseldyke, J., Patterson, M. J. (2010). Students, teachers, and schools as sources of variability, integrity, and sustainability in implementing progress monitoring. *School Psychology Review, 39(4),* 612-630.

Brown, G. T., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education, 27*(1), 210-220.

Burns, M. K., and Gibbons, K. (2008). *Implementing response-to-intervention in elementary and secondary schools.* New York, NY: Taylor & Francis Group, LLC.

Chase, C. I. (1999). *Contemporary assessment for educators.* New York: Longman.

Christ, T. J., & Ardoin, S. P. (2009). Curriculum-based measurement of oral reading: Passage equivalence and probe-set development. *Journal of School Psychology, 47,* 55-75.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy, & Practice, 17*(4), 419-438.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52 (3),* 219-232.

Deno, S. L. (2003). Development in curriculum0based measurement. *Remedial and Special Education, 37 (3),* 184-192.

Dzewaltowski, D. A., Glasgow, R. E., Klesges, L. M., Estabrooks, P. A., & Brock, E. (2004). RE-AIM: evidence-based standards and a web resource to improve translation of research into practice. *Annals of Behavioral Medicine, 28*(2), 75-80.

Fewster, S. & MacMillan, P. D. (2002). School-based evidence for the validity of curriculum-based measurement of reading and writing. *Remedial and Special Education. 23 (3).* 149-156.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*(6), 488-500.

Fuchs, D., & Fuchs, L. S. (2006). Introduction to response to intervention: What, why, and how valid is it? *Reading Research Quarterly, 41*(1), 93-99.

Fuchs, L. S., Fuchs, D., & Compton, D. L. (2004). Onitoring early reading development in first grade: Word identification fluency versus nonsense word fluency. *Exceptional Children, 71*(1), 7-21.

Fuchs, L.S., Fuchs, D., & Hamlett, C.L. (1989). Effects of instrumental use of curriculum-based measurement to enhance instructional programs. *Remedial and Special Education, 10*(2), 43-52.

Good, R. H., & Jefferson, G. (1998). Contemporary perspectives on curriculum-based measurement validity. In M. R. Shinn (Ed.), Advanced applications of curriculum-based measurements. New York, NY: Guildford Press.

Gotch, C. M., & French, B. F. (2013). Elementary teachers' knowledge and self-efficacy for measurement concepts. *The Teacher Educator, 48*, 46-57.

Green, S.K., & Johnson, R. L (2010). Assessment is essential. New York, NY: McGraw Hill.

Greenwood, C. R. (2001). Introduction to the topical issue: Bridging the gap between research and practice in special education: Issues and implications for teacher preparation. *Teacher Education and Special Education, 24 (4),* 273-275.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education, 24 (4),* 276-289.

Harlow, L. L. (2005). *The essence of multivariate thinking.* New Jersey: Mahwah.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

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

IBM Corp. (2013). IBM SPSS Statistics for Mac, Version 22.0. Armonk, NY: IBM Corp.

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425 (2002).

Individuals with Disabilities Education Improvement Act (2004). P.L. No. 108-446 614 [b][6][A]; 614 [b][2&3].

Koh, K. H. (2011). Improving teachers' assessment literacy through professional development. *Teaching Education, 22*, (3), 255-276.

McIntyre, L. L., Gresham, F. M., DiGennaro, F. D., & Reed, D. D. (2007). Treatment integrity of schoolbased interventions with children in the Journal of Applied Behavior Analysis: 1991–2005. *Journal of Applied Behavior Analysis, 40,* 659–672.

Mertler, C. A. (2003). Preservice versus inservice teachers' assessment literacy: Does classroom experience make a difference? Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, Ohio.

Mertler, C. A., & Campbell, C. (2005). Measuring teachers' knowledge & application of classroom assessment concepts: Development of the Assessment Literacy Inventory. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom assessment professional development. *Improving Schools, 12*(2), 101-113.

Nimon, K., Zientek, L. R., & Henson, R. K. The assumption of a reliable instrument and other pitfalls to avoid when considering the reliability of data. *Frontiers in Psychology, 3*, 1-20.

O'Sullivan, R., & Chalnick, M. (1991). Measurement related course work requirements for teacher certification and recertification. *Educational Measurement: Issues in Practice, 10*(1), 17-19.

Pallant, J. (2010). *SPSS Survival Manual*. Open University Press.

Pellegrino, J. (2002-03). Knowing what students know. *Issues in Science and Technology, 19*(2), 48-52.

Plake, B. S., Impara, J. C., & Fager J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement Issues and Practice, 12*(4), 10-12.

Quilter, S. M., & Gallini, J. K. (2000). Teachers' assessment literacy and attitudes. *The Teacher Educator, 36*(2), 115-131.

Reschly, A. L., Busch, T. W., Betts, J., Deno, S. L., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta-analysis of the correlational evidence. *Journal of School Psychology, 47,* 427-469.

Sanetti, L. H., & Kratochwill, T. R. (2009). Toward developing a science of treatment integrity: Introduction to the special series. *School Psychology Review, 38*, 445–459.

Schulte, A. C., Easton, J. E., & Parker, J. (2009). Advances in treatment integrity research: Multidisciplinary perspectives on the conceptualization,

measurement, and enhancement of treatment integrity. *School Psychology Review, 38,* 460–475.

Shapiro, E. S. (2013). Commentary on progress monitoring with CBM-R and decision making: Problems found and looking for solutions. *Journal of School Psychology, 51,* 59-66.

Shinn, M. R., & Bamonto, S. (1998). *Advanced applications of curriculum-based measurement: "Big ideas" and avoiding confusion*. In M. R. Shinn (Ed.), Advanced Applications of Curriculum-Based Measurement (pp.1–31). New York: Guilford Press.

Shinn, M. R., Shinn, M. M., Hamilton, C., & Clarke, B. (2002). *Using curriculum-based measurement in general education classrooms to promote reading success.* In M. R. Shinn, H. M. Walker & G. Stoner (Eds.) Interventions for Academic and Behavior Problems II: Preventative and Remedial Approaches (pp.113-142). Maryland: NASP Publications.

Stecker, P. M., Fuchs, L. S., & Fuchs, D. (2005). Using curriculum-based measurement to improve student achievement: Review of research. *Psychology in the Schools, 42(8),* 795-819.

Stiggins, R. J., & Conklin, N. F. (1992). In teachers' hands: Investigating the practices of classroom assessment. Albany: State University of New York Press.

Stiggins, R. (1991). Assessment literacy. Phi Delta Kappan, 72, 534-539.

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

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

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's Alpha. *International Journal of Medical Education, 2*, 53-55.

Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present.

University of Oregon, Center on Teaching and Learning (2012). *DIBELS Next Recommended Benchmark Goals: Technical Supplement* (Technical Report 1204). Eugene, OR: Author. Retrieved from https://dibels.uoregon.edu/docs/techreports/DDS2012TechnicalSupplement.pdf

Volante, L. & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749-770.

Wayman, M. M., Espin, C., Deno, S., McMaster, K., Mahlke, A., & Du, Xiaoqing. (2011). Teachers' understanding of curriculum-based measurement progress monitoring data.

Wayman, M. M., Wallace, T., Wiley, H. I., Ticha. R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*(2), 85-120.

Ysseldyke, J., & Bolt, D. M. (2007). Effect of technology-enhanced continuous progress monitoring on math achievement. *School Psychology Review, 36*(3), 453-467.

Ysseldyke, J., Burns, M. K., Scholin, S. E., & Parker, D. C. (2010). *Teaching Exceptional Children, 42*(4), 54-61.