AN EXAMINATION OF THE CONCURRENT VALIDITY, UNIQUE CONTRIBUTIONS, AND TEACHER ACCEPTABILITY OF UNIVERSAL SCREENING ASSESSMENTS

by

STACY-ANN A. JANUARY

(Under the Direction of Scott P. Ardoin)

ABSTRACT

Schools employ curriculum-based measurement of oral reading (CBM-R) and the Measures of Academic Progress (MAP) in their universal screening efforts to identify struggling readers. Among other things, universal screening assessments must be technically adequate and acceptable to teachers. A large body of research exists supporting the technical adequacy of CBM-R, but the degree to which it is acceptable to teachers is unclear. Regarding the MAP, only one empirical study examined its technical adequacy and no research exists evaluating its acceptability. The first study in this dissertation examined the concurrent validity of MAP and the relationship between CBM-R and reading achievement by administering CBM-R and MAP to 802 students in Grades 1-5. The second study examined teachers' knowledge, use, and acceptability of MAP and CBM-R with a survey completed by 86 first through fifth-grade teachers. Together, results from these two studies provide strong support for the technical adequacy of MAP for universal screening and the acceptability of both MAP and CBM-R.

INDEX WORDS: Curriculum-based measurement, Reading achievement, Computer Adaptive Tests, School-wide assessments, Universal screening

AN EXAMINATION OF THE CONCURRENT VALIDY, UNIQUE CONTRIBUTIONS, AND TEACHER ACCEPTABILITY OF UNIVERSAL SCREENING ASSESSMENTS

by

STACY-ANN A. JANUARY

B.A., University of Richmond, 2006

M.A., University of Georgia, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

AN EXAMINATION OF THE CONCURRENT VALIDITY, UNIQUE CONTRIBUTIONS, AND TEACHER ACCEPTABILITY OF UNIVERSAL SCREENING ASSESSMENTS

by

STACY-ANN A. JANUARY

Major Professor:	Scott P. Ardoin
Committee:	Amy L. Reschly A. Michele Lease Martha M. Carr

Electronic Version Approved:

Julie Coffield
Interim Dean of the Graduate School
The University of Georgia
August 2014

DEDICATION

This dissertation is dedicated to my grandmother. I have always admired your strength and determination. Thank you for all your love, support, and the encouragement to pursue my dreams. I hope that you are proud of me. I miss you and I love you.

ACKNOWLEDGEMENTS

I would like to extend my sincerest gratitude to the family, faculty, and friends who have supported me throughout this process. To my major professor, Scott Ardoin, thank you for guiding me academically, supporting me, and challenging me intellectually. I would also like to thank Drs. Amy Reschly, Michele Lease, and Martha Carr for their valuable insight and feedback. I would like to thank the graduate students who assisted me throughout my work on this dissertation. Finally, I owe the deepest appreciation to my mother, Charmain, and husband, Allen, for their unfailing emotional, intellectual, and financial support as well as their unconditional love.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER 1

GENERAL INTRODUCTION

Recent passage of federal laws such as No Child Left Behind (U.S. Department of Education, 2001) and the subsequent widespread adoption of a Response to Intervention framework led to an increase in assessment within schools. Schools across the country now utilize universal screening assessments to identify students who may not meet academic goals (Deno et al., 2009). Curriculum-based measurement (CBM) and computer adaptive tests (CATs) are widely used by districts across the country for universal screening (Clemens, Hilt-Panahon, Shapiro, & Yoon, 2012; Northwest Evaluation Association, 2011), often in combination. Hosp and Ardoin (2008) and Glover and Albers (2007) recommend that universal screening instruments should be short, easy to administer, cost effective, technically adequate, acceptable to users, and demonstrate decision-making utility. The extent to which multiple universal screeners predict achievement and the degree to which CBM and CATs meet criteria as universal screeners is unknown.

CBM is a set of empirically-valid procedures for measuring students' competencies in reading, among other subject areas (Hosp & Hosp, 2003). CBM of oral reading (CBM-R) requires that students read a grade-level passage aloud for 1 min while the examiner records the number of words read correctly per minute (WRCM). Research indicates that CBM-R is reliable (Betts, Pickard, & Heistad, 2009; Graney, Martinez, Missall, & Aricak, 2010; McGlinchey & Hixson, 2004) and valid as a strong predictor of students' reading achievement (Deno, Mirkin, & Chiang, 1982; Reschly, Busch, Betts, Deno, & Long, 2009; Wayman, Wallace, Wiley, Ticha, &

Espin, 2007). Despite research suggesting that CBM-R requires the integration of several component reading skills (Fuchs, Fuchs, Hosp, & Jenkins, 2001), prior research suggests its relationship with specific reading skills (e.g., word decoding, comprehension) changes with skill level and age (Ardoin et al., 2013; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003). For example, findings from Jenkins et al. (2003) indicate that reading comprehension explains more variance in fourth-graders' CBM-R performance than word reading. However, for younger, less skilled readers, word reading accounts for more variance in CBM-R than reading comprehension (Ardoin et al., 2013).

CATs are individually-administered multiple-choice computerized assessments in which each test taker receives a unique set of items that are near his or her achievement level. The Measures of Academic Progress (MAP) is a CAT that assesses achievement in reading (among other academic areas) for students in kindergarten through 12th grade (Northwest Evaluation Association, 2009). Despite its benefits, MAP is more expensive than CBM and to date there is no research examining teacher acceptability of MAP. Furthermore, besides evidence of its technical adequacy reported by the test publisher (Northwest Evaluation Association, 2009), only one published peer-refereed study (Merino & Beckman, 2010) exists related to the MAP's technical adequacy. Findings from that study were, however, promising with results indicating that CBM-R strongly predicts MAP performance and MAP moderately predicts CBM-R performance.

In regard to established criteria for universal screening, there is ample evidence that CBM-R procedures are relatively short, easy to administer, inexpensive, and have adequate reliability and validity (Ball & Christ, 2012; Marston, 1989). Unfortunately, research examining teachers' understanding and acceptability of CBM-R is less positive. The research that does

exist suggests that teachers doubt CBM-R as a valid estimate of reading comprehension (Foegen et al., 2001). This is especially troubling considering research indicates that teachers are not likely to employ CBM-R when making decisions regarding a student's curriculum if they do not rate CBM-R as acceptable (Allinder & Oats, 1997). These results are, however, derived from studies conducted prior to (a) the National Reading Panel identifying reading fluency as a key component of effective reading instruction (National Institute of Child Health and Human Development, 2000) and (b) the passing of federal legistation that allows for students' response to instruction to be used in the determination of students' special education eligibility (U.S. Department of Education, 2001). These two changes have increased the use and relevance of CBM-R data within schools and thus possibly teachers' knowledge, use, and acceptability of CBM-R.

Overall Purpose

Schools across the country employ universal screening measures such as CBM-R and MAP to identify students "at-risk" for academic underachievement (Mellard, McKnight, & Woods, 2009). Some schools in fact spend the time and money to administer both CBM-R and MAP for universal screening purposes. Although there may be obvious benefits to administering either test alone, researchers have yet to evaluate the unique contribution of these measures in predicting students' reading achievement and whether there is any benefit to administering both measures. Furthermore, the degree to which each assessment meets all of the criteria for universal screeners is yet to be determined. For instance, although acceptability is an important criterion of universal screeners (Glover & Albers, 2007), there is a paucity of research examining teachers' acceptability of both CBM-R and MAP. Therefore, the overall purpose of this dissertation is to examine the utility of CBM-R and the MAP as universal screening assessments.

Specifically, the first study in this dissertation will examine the concurrent relationship between CBM-R, MAP, and a group-administered norm-referenced achievement test and the incremental benefits of administering both CBM-R and MAP for universal screening. The second study will explore teachers' knowledge, use, and acceptability of CBM-R and MAP for universal screening.

References

Allinder, R. M., & Oats, R. G. (1997). Effects of acceptability on teachers' implementation of curriculum-based measurement and student achievement in mathematics computation. *Remedial and Special Education, 18*(2), 113-120. doi: 10.1177/074193259701800205

Ardoin, S. P., Eckert, T. L., Christ, T. J., Morena, L. S., January, S.-A. A., & Hine, J. F. (2013). Examining variance in reading comprehension among developing readers: Words in context (curriculum-based measurement in reading) versus words out of context (word lists). *School Psychology Review, 42*(3), 243-261.

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. doi: 10.1002/pits.21592

Betts, J., Pickard, M., & Heistad, D. (2009). An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating the alternate forms. *Journal of School Psychology, 47*, 1-17. doi: 10.1016/j.jsp.2008.09.001

Clemens, N. H., Hilt-Panahon, A., Shapiro, E. S., & Yoon, M. (2012). Tracing student responsiveness to intervention with early literacy skills indicators: Do they reflect growth toward text reading outcomes? *Reading Psychology, 33*(1/2), 47-77. doi: 10.1080/02702711.2011.630608

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*(1), 36-45.

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callender, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools, 46*(1), 44-55. doi: 10.1002/pits.20353

Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education, 34*(4), 226-236. doi: 10.1177/002246690103400405

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5*(3), 239-256. doi: 10.1207/s1532799xssr0503_3

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117-135. doi: 10.1016/j.jsp.2006.05.005

Graney, S. B., Martinez, R. S., Missall, K. N., & Aricak, O. T. (2010). Universal screening of reading in late elementary school: R-CBM versus CBM maze. *Remedial and Special Education, 31*(5), 368-377. doi: 10.1177/0741932509338371

Hosp, J. L., & Ardoin, S. P. (2008). Assessment for instructional planning. *Assessment for Effective Intervention, 33*(2), 69-77. doi: 10.1177/1534508407311428

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10-17.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729. doi: 10.1037/0022-0663.95.4.719

Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it? In M. R. Shinn (Ed.), *Curriculum-Based Measurement: Assessing Special Children* (pp. 18-78). New York: Guilford.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 193-203.

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24*(4), 186-195. doi: 10.1111/j.1540-5826.2009.00292.x

Merino, K., & Beckman, T. O. (2010). Using reading curriculum-based measurements as predictors for the Measures of Academic Progress (MAP) standardized test in Nebraska. *International Journal of Psychology: A Biopsychosocial Approach, 6*, 85-98.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. *NIH Publication No. 00-4769*. Washington, DC: U.S. Government Printing Office.

Northwest Evaluation Association. (2009). Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades. Lake Oswego, OR: Author.

Northwest Evaluation Association. (2011). 2011 Normative Data. Retrieved 9/13/12, from http://www.nwea.org/support/article/normative-data-2011

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*(6), 427-469. doi: 10.1016/j.jsp.2009.07.001

U.S. Department of Education. (2001). No Child Left Behind Act of 2001. Retrieved 9/8/2012, from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

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.

CHAPTER 2

EXAMINING THE CONCURRENT VALDITY AND UNIQUE CONTRIBUTIONS OF CURRICULUM-BASED MEASUREMENT IN READING AND THE MEASURES OF ACADEMIC PROGRESS IN PREDICTING READING ACHIEVEMENT[1]

[1] January, S.-A. A. & Ardoin, S. P. To be submitted for publication.

Abstract

Curriculum-based measurement of oral reading (CBM-R) and the Measures of Academic Progress (MAP) are assessment tools widely employed for universal screening in schools. Although a large body of research exists supporting the validity of CBM-R, limited empirical evidence exists supporting the technical adequacy of MAP for universal screening. Purposes of the current study were to replicate and extend prior research by (a) evaluating the concurrent validity of MAP; (b) determining what benefit there is, if any, to administering MAP with CBM-R for universal screening; and (c) examining the unique contribution of MAP subtests in explaining variance in CBM-R. Participants included 802 students in Grades 1-5 who were administered three CBM-R probes and the MAP during the schools' winter universal screening period. Results extend the literature by demonstrating grade-level differences in the relationship between CBM-R and reading skills and support the use of MAP for universal screening.

Introduction

Within the last 10 years, greater emphasis on accountability and the widespread adoption of a Response to Intervention (RtI) framework has revolutionized the nature of assessment within schools. RtI models include the universal screening of all students several times per year. The purposes of universal screening within schools are to (a) identify those who are "at-risk" for failing to meet academic goals (Mellard, McKnight, & Woods, 2009), (b) predict which students are likely to pass the high-stakes achievement test (Wiley & Deno, 2005), (c) provide a benchmark from which to evaluate growth (Deno et al., 2009), (d) evaluate the effectiveness of the general curriculum within a school (Reschly & Bergstrom, 2009), (e) evaluate the performance rates of schools within a district (Hosp & Ardoin, 2008), and (f) develop school-wide and district-level norms (Hosp & Ardoin, 2008). By identifying students early within RtI

models, schools are able to intervene quickly, maximizing the likelihood that students will benefit from intervention (Barnett, 1995) and reducing the referral rates for special education services (Fuchs, Mock, Morgan, & Young, 2003; VanDerHeyden, Witt, & Gilbertson, 2007). Two measurement procedures often employed for conducting such screenings are curriculum-based measurement (CBM) and computer adaptive tests (CATs).

CBM is a set of individually- or group-administered fluency-based assessment procedures in the academic areas of reading, mathematics computation, spelling, and written expression (Hosp & Hosp, 2003). CBM procedures, particularly those in reading, are validated as strong estimates of students' global academic achievement (Foegen, Jiban, & Deno, 2007; Fuchs, Fuchs, & Maxwell, 1988; Mercer, Martínez, Faust, & Mitchell, 2012; Schilling, Carlise, Scott, & Zeng, 2007), making it useful for screening, benchmarking, and making data-based decisions. CATs are individually administered computer-based achievement measures, and unlike CBM, which estimates students' global academic achievement, CATs generally are designed to measure academic competencies (e.g., phonics, vocabulary, comprehension). Unfortunately, despite the potential benefits of such data for teachers, little published research exists examining their utility for universal screening or benchmarking purposes. Furthermore, the expense of CATs may prevent many schools from adopting them in place of cheaper and empirically-validated procedures such as CBM.

Interestingly, some schools conduct universal screenings using both CBM and CATs. Employing CBM and CATs for universal screening may have utility to the extent that the combined resultant data provide more information about students' reading achievement than each assessment independently. However, the degree to which these multiple sources of data provide unique information to the prediction of academic achievement is unknown. The

remainder of this introduction will present relevant background information, review pertinent assessment literature, and detail the purposes of the current study.

Curriculum-Based Measurement of Oral Reading

The most common CBM procedure used in schools and examined within the research literature is CBM of oral reading (CBM-R; Busch & Reschly, 2007). CBM-R is a timed, individually-administered assessment of oral reading rate with accuracy. Students read a passage aloud for 1 min while an examiner records errors. The number of words read correctly per minute (WRCM) is the dependent score. Although researchers originally developed CBM-R to assist special education teachers in developing individual education goals and evaluating individual students' progress toward their goals (Deno, 2003), its use has expanded. Schools now utilize CBM-R for universal screening purposes largely due to its strong relationship with measures of academic achievement (Fuchs & Fuchs, 1992; Shinn, 1989) and its ability to reliably identify students "at-risk" for later failing high-stakes achievement tests (Hintze & Silberglitt, 2005; McGlinchey & Hixson, 2004). Over 30 years of research provides evidence of the reliability and validity of CBM-R. CBM-R has excellent test-retest and alternate forms reliability in the elementary (Betts, Pickard, & Heistad, 2009; Graney, Martinez, Missall, & Aricak, 2010; Hosp & Fuchs, 2005; McGlinchey & Hixson, 2004) and secondary grades (Espin, Wallace, Lembke, Campbell, & Long, 2010; McMaster, Wayman, & Cao, 2006; Ticha, Espin, & Wayman, 2009). Additionally, evidence suggests that CBM-R has adequate concurrent and predictive validity for estimating reading achievement at the elementary (Deno, Mirkin, & Chiang, 1982; Fuchs et al., 1988; Jenkins & Jewell, 1993; McGlinchey & Hixson, 2004) and secondary levels (Espin et al., 2010; Ticha et al., 2009).

Despite a long history of support demonstrating its technical adequacy and its wide use within schools for universal screening and progress monitoring purposes, there are limitations associated with CBM-R. For example, a recent meta-analysis indicated that the strength of the relationship between CBM-R and reading achievement differs by type of criterion assessment and the length of time between administrations of measures (Reschly, Busch, Betts, Deno, & Long, 2009). Specifically, CBM-R demonstrates higher correlations with individually- versus group-administered tests, for national norm-referenced tests versus state achievement tests, and within the academic year as opposed to across academic years (Reschly et al., 2009). These differences are important when considering the utility of CBM-R for universal screening. A second limitation associated with CBM-R is that it lacks face validity among educators as an estimate of students' reading comprehension (Faykus & McCurdy, 1998; Hamilton & Shinn, 2003), even after educators have received training on its utility and validity (Foegen, Espin, Allinder, & Markell, 2001). Although CBM-R is established as a global outcome measure (Fuchs & Deno, 1991), teachers reportedly often perceive it merely as a measure of word reading skills (Hamilton & Shinn, 2003).

Despite teachers' beliefs, research suggests that CBM-R is more than simply a measure of students' decoding skills. For example, Jenkins, Fuchs, van de Brock, Espin, and Deno (2003a, 2003b) examined the relationship between word reading, CBM-R, and a group-administered norm-referenced assessment of reading comprehension. Results indicated that fourth-grade students read more words in context (CBM-R) versus out of context (word lists) and that reading comprehension explained more variance in CBM-R than word lists (Jenkins et al., 2003b). However, for less skilled readers, single-word decoding accounted for more variance in CBM-R than for more skilled readers (Jenkins et al., 2003a). The aforementioned studies were

replicated and extended in recent research employing younger, less skilled readers (Ardoin et al., 2013). Consistent with Jenkins et al. (2003b), first- and second-grade students' oral reading rate with accuracy was greater for connected text (i.e., CBM-R) than sight-word lists. However, in contrast to Jenkins et al. (2003b), for the first- and second-grade students in Ardoin et al. (2013), the high-frequency sight-word lists explained more variance in CBM-R than did the measure of reading comprehension (Ardoin et al., 2013). These results, together with other research (Jenkins & Jewell, 1993), suggest that the relationship between CBM-R and component reading skills (such as decoding and comprehension) may change with age and skill level (Ardoin et al., 2013; Jenkins et al., 2003a, 2003b). However, this conclusion was drawn across several studies employing different aged participants. Thus far, no studies have examined the unique contribution of decoding and reading comprehension skills to CBM-R using a single sample of students across first through fifth grades.

Computer Adaptive Tests

Computer Adaptive Tests (CATs) are uniquely advantageous because they provide students with a personalized set of multiple-choice test items, presented one at a time, tailored to the students' achievement level. The first test item on a CAT is of moderate difficulty, targeted at the level of an average test-taker in that grade. Subsequent items are selected for students based on the accuracy of their response to the previous item(s). Therefore, unlike a traditional test in which every student receives the same items, a CAT provides assessment items that adapt to each student's performance, presenting items that are near the student's achievement level instead of items that might be too easy or too difficult. The purpose of the adaptive nature of the assessment is to obtain a precise measurement of a student's achievement level and provide information about how that student performs in that particular academic domain. The adaptive

nature of CATs also results in a lower standard error of measurement when compared to traditional group administered norm-referenced tests. Decreased standard error equates to a more accurate assessment and potentially greater confidence in the scores gleaned from such tests. Other advantages of CATs are that they typically have a lower floor and higher ceiling than traditional tests, due to the large number of available items at all levels of achievement, and because it is computer-based, test administration procedures are reliable across students. Although CATs may provide schools with more detailed information regarding students' skills than does CBM, drawbacks of CATs include the expense per student, resources required (e.g., computers), and limited validity evidence.

The Measures of Academic Progress (MAP) is a CAT that is based on the Rasch model of Item Response Theory. The Rasch model considers item difficulty and test-taker ability to estimate the probability that a student will be successful on a given item. Items are administered to each test taker, one at a time, adapting to the students' responses until the probability of the student responding accurately to each item is estimated to be 50%. MAP scores are reported on an equal interval Rasch Unit (RIT) scale.

MAP provides assessments in the subject areas of reading, language use, mathematics, and science in Grades 2-12 and reading and math for students in Grades K-2 (Northwest Evaluation Association, 2009). Within the subject area of reading (for Grade 2 and above), the MAP provides information regarding students' competencies from the following goal areas (a) foundational skills and vocabulary; (b) informational texts; and (c) literature texts. In Grades K-2, schools have the option of administering the MAP Primary Grades (MPG), which is similar to the MAP, but measures early literacy skills. The MPG Reading includes three goal areas that measure: (a) foundational skills; (b) literature and informational texts; and (c) vocabulary use

and functions. Schools administer the MAP multiple times per year for screening and informing instruction. According to the publisher, teachers can use students' MAP scores to provide differentiated instruction based upon each student's demonstrated strengths and weaknesses (Northwest Evaluation Association, 2012). In addition to the aforementioned advantages, RIT scores fall along a vertical scale for Grades 2-12, allowing for comparisons of student growth within and across grade levels. For kindergarten through second grade, the MPG has a separate RIT vertical scale.

Despite the fact that over 13,000 schools across all 50 states utilize the MAP (Northwest Evaluation Association, 2011), currently no peer-refereed study exists examining the criterion or predictive validity of MAP with a norm-referenced reading achievement test. Merino and Beckman (2010) did, however, examine the relationship between students' performance on CBM-R and the MAP Reading Survey with Goals. The authors examined the predictive validity of CBM-R with MAP as the criterion measure for 376 students in Grades 2-5. The MAP and CBM-R probes were both administered in the spring of one academic year and in the fall of the following year. Results indicated that spring CBM-R performance demonstrated moderately strong predictive validity with fall MAP scores (*r*s = .67 to .73), spring MAP scores demonstrated moderate predictive validity with fall CBM-R performance (*r*s = .62 to .66), and spring CBM-R scores in Grade 2 accounted for more variance in fall MAP scores than in any other grade. Despite having administered both measures at each assessment period, the researchers failed to report the concurrent validity between the two measures at a given time point.

Purpose of the Study

Schools conduct universal screenings several times per year employing CBM-R and MAP (Mellard et al., 2009; Northwest Evaluation Association, 2011). Despite wide use of the MAP, there are currently no studies examining the combined and unique contributions of CBM-R and the MAP in explaining students' overall reading achievement or the concurrent validity of the MAP with a norm-referenced achievement test. Therefore, the overall purpose of the current study was to determine the concurrent relationship among CBM-R, MAP, and reading achievement. First, the current study examined the relationship between CBM-R and MAP across Grades 1-5. Although Merino and Beckman (2010) previously examined the predictive relationship between CBM-R and MAP, that study failed to examine that relationship with first-grade students. Furthermore, Merino and Beckman (2010) administered grade-level probes from AIMSweb (Howe & Shinn, 2002), a universal screening and progress monitoring assessment system that provides CBM probes in reading. AIMSweb probes are not equated vertically across grade levels, resulting in the inability to make comparisons by collapsing across grade levels. Therefore, the current study extended prior research by employing first-grade students and a different set of probes [Formative Assessment Instrumentation and Procedures for Reading (FAIP-R)], which are equated both vertically and horizontally, allowing for comparisons across grade levels.

A second purpose of the current study was to extend upon the research by Jenkins et al. (2003a, 2003b) and Ardoin et al. (2013) to examine the unique contribution of MAP subtests in explaining CBM-R performance within Grades 1-5. These analyses further extended knowledge regarding differences in the skills assessed by CBM-R across Grades 2-5. A third purpose of the current study was to examine concurrent validity of CBM-R and MAP with the Iowa Test of

Basic Skills (ITBS). Finally, the current study aimed to examine the unique benefit, if any, to administering both CBM-R and MAP when conducting universal screenings.

The following research questions were addressed by the current study:

(1) What is the concurrent validity of the MAP and CBM-R?

(2) What is the unique contribution of MAP subtests in explaining variance in CBM-R performance within and across grades?

(3) What is the concurrent validity between CBM-R, MAP, and ITBS for students in third grade?

(4) What is the unique contribution of the MAP and CBM-R to third-grade students' performance on the ITBS?

Method

Participants and Setting

Students. Participants were 802 students in first (n = 158), second (n = 158), third (n = 147), fourth (n = 159) and fifth (n = 180) grade from two public elementary schools in one suburban district located in the Southeast. All students in first through fifth grade in the participating schools who were present during the winter universal screening sessions were included as participants in the current study. Students participated in the current study anonymously, therefore their individual demographic data was not obtained.

Schools. Data were collected at two elementary schools. Elementary school A enrolled approximately 480 students in Grades K-5 and elementary school B, a Title 1 school, enrolled approximately 500 students in Grades K-5. The most current demographic information available for school A indicated that approximately 81% of the students are White, 6% Hispanic, 5% Asian, 4% Black, and 4% Multiracial. School-wide, 20% of students qualified for free or

reduced-priced meals and 11% were enrolled in special education. Demographic data reported for school B indicated that 88% of the student population is White, 5% Hispanic, 3% Black, 3% Multiracial, and 1% Asian. Approximately 26% of students in school B qualified for free or reduced-priced meals and 9% were enrolled in special education.

Measures

Curriculum-Based Measurement of Oral Reading (CBM-R). CBM-R probes from the Formative Assessment Instrumentation and Procedures for Reading (FAIP-R; Christ, Ardoin, & Eckert, 2010) passage set were utilized for the current study. FAIP-R consists of 18 CBM-R probes, 3 for universal screening and 15 for progress monitoring for students in Grades 2-5, and 3 universal screening and 10 progress monitoring probes in Grade 1. This study employed the three universal screening probes at each grade level. The median and transformed WRCM for the three probes were utilized as the dependent scores.

FAIP-R CBM-R probes are unique in that as opposed to relying on readability formulas for passage equating, passage equivalence was based upon the field testing of passages with over 500 students reading all passages within specified difficulty levels. Passages were equated both horizontally (within grades) and vertically (across grades). FAIP-R probes, which are approximately 220-240 words, follow a specific structure. Each narrative probe tells a story in three to four paragraphs in which characters encounter a problem or establish a goal in the first paragraph, take action to resolve that problem or meet the goal in the second or third paragraph, and the problem is resolved or the goal is attained in the final paragraph(s). Probes do not include dialogue, slang/jargon, or controversial topics (e.g., divorce).

Measures of Academic Progress (MAP). The MAP is a multiple-choice computer adaptive test that is published by Northwest Evaluation Association. In reading, MAP has two

test options, Survey and Survey with Goals (Northwest Evaluation Association, 2009). The Survey test consists of the administration of less than 20 items with an approximate administration time of 35 min. The Reading Survey assessment provides an overall RIT score and is designed to be a quick measure of general reading proficiency, but not to measure student growth. In contrast, the Reading Survey with Goals assessment is longer (approximately 42 items with a 65-min administration time) and includes a more detailed assessment of each subscale. Therefore, the MAP Reading Survey with Goals produces an overall RIT score and RIT scores for each goal area.

In addition to the Survey with Goals tests, the MAP Primary Grades (MPG) includes a Screening test and Skills Checklist, which are diagnostic assessments designed to provide teachers with information regarding students' competencies in basic reading skills such as identification of initial consonants, final consonants, and vowels. In contrast to the Survey with Goals test, results of the Screening test and Skills Checklist are reported as the percent or number of items correct, not a RIT score. Therefore, the Screening test and Skills Checklist are not beneficial for universal screening, but may be useful in other applications such as for planning instruction or as a summative assessment.

For this study, the MAP (Grades 2-5) and MPG (Grade 1) Reading Survey with Goals tests were administered. The MAP Reading Survey with Goals provides an overall composite score and scores for each goal area: (a) foundational skills and vocabulary, (b) informational texts, and (c) literature texts. MPG Reading Survey with Goals includes an overall score and subscores for performance in the following goal areas: (a) foundational skills, (b) vocabulary use and functions, and (c) literature and informational texts. MAP tests were administered via the

computer with adult supervision, in each school's computer lab. Students' RIT scores, which ranged from 120 to 250 for the MAP and the MPG, were utilized as the dependent scores.

According to the technical manual, the reliability and validity of the MAP is adequate. The marginal reliability of the MAP ranges from .94 to .95 for students in second through fifth grade and .94 to .97 for the MPG for students in kindergarten through second grade (Northwest Evaluation Association, 2009). Test-retest reliability ranges from .70 to .85 for students in Grades 2-10 and .71 to .86 for Grade 1. Internal consistency for Grades 1-9 ranges from .70 to .85 from spring to fall administration. The concurrent validity between the MAP Reading Survey with Goals and various state achievement tests is moderate to strong, ranging from .57 to .79 for students in Grades 2-5 (Northwest Evaluation Association, 2009). Similarly, the predictive validity of the MAP Reading Survey with Goals and state achievement tests ranges from .63 to .70, indicating a moderate relationship for Grades 2-5 from fall to spring (Northwest Evaluation Association, 2009).

Iowa Test of Basic Skills (ITBS). The ITBS is a group-administered norm-referenced achievement test for students in kindergarten through eighth grade that is published by Riverside Publishing. The participating district administers the ITBS to third grade students only. Third-grade students in the current study were administered Form A of the ITBS. For the purposes of this study, the ITBS-Total Reading (ITBS-TR) composite, which includes two subtests that measure students' reading comprehension and vocabulary skills, were employed. According to the ITBS technical manual, the Total Reading composite has a Kuder Richardson-20 internal consistency of .94 (University of Iowa, 2005). Students' standard scores for the ITBS-TR composite, which ranged from 120 to 250, were utilized as the dependent score.

Procedures

All data were collected across three sessions. First, classroom teachers following standardized procedures administered the ITBS to all third graders during the second week of the fourth month of the school year. Next, the MAP was administered via the computer during the last week of the fourth month through the first two weeks of the fifth month of the year. Finally, all first- through fifth-grade students were individually administered three grade-level universal screening CBM-R probes by graduate and undergraduate student examiners, during the second week of the fifth month of the year, in the hallways adjacent to their classrooms.

Procedural Integrity and Inter-Scorer Agreement

Third-grade teachers who administered the ITBS were trained in standardized administration procedures by their district. CBM-R probes were administered by graduate students enrolled in a school psychology doctoral program and undergraduate research assistants who had extensive experience with CBM-R standardized administration procedures. During one half-hour training session led by the primary investigator, examiners were provided with an overview of the study and a review of administration and scoring procedures. Each CBM-R administration was audio recorded. Approximately 20% of these audio recordings were randomly selected for the evaluation of inter-scorer agreement and procedural integrity. Inter-scorer agreement was calculated by dividing the number of word by word agreements by the number of agreements plus disagreements multiplied by 100%. Inter-scorer agreement averaged 99% (range, 93 to 100%). Procedural integrity was calculated by dividing the number of accurately completed steps by the number of total steps (17) multiplied by 100%. Procedural integrity averaged 98% (range, 94 to 100%) across examiners.

Data Analysis

To examine the relationship between CBM-R, MAP, and ITBS-TR within and across grade levels, Pearson product-moment correlations were employed. Hierarchical multiple regression analyses were utilized to determine the shared and unique contribution of MAP subtests in explaining CBM-R performance and the unique contribution of CBM-R and MAP scores in explaining variance in ITBS-TR performance. For the purposes of making statistical comparisons across grade levels, students' median WRCM were transformed onto the FAIP-R vertical scale, yielding a transformed WRCM score. To control for Type 1 error, a Bonferroni correction was applied and an alpha level of .006 was used across all eight regression models.

Results

Descriptive Statistics

All variables were examined for normality, kurtosis, and skewness; and all values were found to be within normal limits. Descriptive statistics for CBM-R and MAP are presented in Table 2.1. Correlational analyses yielded statistically significant correlations among all experimental variables. Within each grade, the magnitude of the correlations between CBM-R and MAP overall RIT scores were high (range, .72 to .79). Fisher's Z comparisons yielded no statistically significant differences between grade-level correlational coefficients, suggesting that the relationship between CBM-R and MAP did not differ significantly between grade levels. Across Grades 2-5, CBM-R was highly related to the MAP overall RIT score (r = .83) and goal areas of foundational skills and vocabulary (r = .79), informational texts (r = .80), and literature texts (r = .78; see Table 2.2).

MAP Goal Areas as Components of CBM-R

Hierarchical regression analyses were utilized to examine the unique contribution of MAP goal areas to CBM-R within each grade (see Table 2.3). For first graders, the foundational skills and vocabulary use and functions subtests were entered in the first block, and the literature/informational texts subtest was entered in the second block. Results indicated that foundational skills and vocabulary accounted for 45% of the variance in CBM-R scores and the additional variance explained by literature/informational texts (4%) approached significance (p = .035). Due to differences in MAP goal areas, hierarchical regression analyses for second through fifth grade were slightly different than for first grade. For this set of regressions, CBM-R performance was predicted by entering foundational skills/vocabulary in the first block and the literature texts and informational texts in the second block. Across Grades 2-5, foundational skills and vocabulary explained a significant portion of the variance in CBM-R (63%) and together, informational and literature texts accounted for an additional 6% ($p < .001$) of the variance in CBM-R (see Table 2.4). Within grade level findings were consistent with across grade level results and are presented in Table 2.3.

MAP and CBM-R as Components of Reading Achievement among Third-Grade Students

For students in third grade, both MAP overall RIT scores and median CBM-R scores were highly related to the ITBS-TR composite. A statistically significant difference between the CBM-R and ITBS-TR (r = .74) and MAP and ITBS-TR (r = .87) correlations was observed, z = -3.25, p = .001. Hierarchical multiple regression analyses were conducted to evaluate the unique contribution of CBM-R and MAP overall RIT scores in third-grade students' ITBS-TR performance (see Table 2.5). Together, CBM-R and MAP explained 76% of the variance in the ITBS-TR composite. In the first hierarchical regression, CBM-R scores were entered first and

results indicated that CBM-R accounted for 55% of the variance in students' ITBS-TR scores and MAP accounted for an additional 21% of the variance ($p < .001$). For the second hierarchical regression, when entered first, MAP accounted for 75% of the variance in ITBS-TR, but CBM-R did not explain any significant unique variance.

Discussion

Schools across the country employ CBM-R and MAP in their universal screening efforts as a part of the implementation of an RtI framework. By screening all students, educators can, among other things, identify those students who are "at-risk" and predict students' performance on standardized tests (Mellard, McKnight, & Woods, 2009; Wiley & Deno, 2005). Some schools administer both the MAP and CBM-R to students, despite the fact that the benefit of administering both instruments for universal screening is unknown. The technical adequacy of CBM-R is well established (Reschly et al., 2009; Wayman, Wallace, Wiley, Ticha, & Espin, 2007); however, to date, only one published study (Merino & Beckman, 2010) has evaluated the predictive validity of MAP with CBM-R for students in second through fifth grade. The current study extends Merino and Beckman (2010) by including students in first grade and aggregating results across second through fifth grade. In general, findings from the current study replicate and extend the CBM-R literature and provide strong support for the use of the MAP for universal screening.

Findings from the current study add to the wealth of CBM-R literature supporting its use as a universal screener for elementary school students (e.g., Ardoin et al., 2004; Graney, Martinez, Missal, & Aricak, 2010). Results support previous research demonstrating a strong relationship between CBM-R and measures of reading achievement as results indicated a strong relationship between CBM-R and MAP overall RIT scores as well as with specific MAP goal

areas (foundational skills and vocabulary, informational texts, and literature texts). Furthermore, regression analyses suggested that CBM-R explained a large portion of the variance in third-grade students' ITBS-TR performance. This study also extends existing CBM-R research suggesting that CBM-R is not simply a measure of students' word reading rates. Consistent with prior research (Jenkins et al., 2003b), after controlling for foundational reading skills and vocabulary, the two MAP subtests designed to measure components of reading comprehension (i.e., literature texts and informational texts) explained a significant amount of variance in CBM-R both across and within Grades 2-5. Despite supporting previous research, the amount of variance explained by reading comprehension in Jenkins et al. is greater in magnitude than in current findings, which is likely at least partially due to differences in the instruments employed. Whereas Jenkins et al. utilized the ITBS subtest that was developed to be a primary measure of reading comprehension, the MAP goal areas employed in the current study measure students' understanding of the craft and structure of texts in addition to measuring reading comprehension. When examining data for the first grade sample, results were similar to those reported by Ardoin et al. (2013), in that after controlling for basic reading skills, the MAP's measure of comprehension (i.e., literature and informational texts goal area) failed to explain unique variance in CBM-R.

By employing one sample of students in Grades 1-5, the current study also extends prior research suggesting that the relationship between CBM-R and reading comprehension varies by grade level (Jenkins & Jewell, 1993; Jenkins et al., 2003b; Ardoin et al., 2013). As compared with data reported in previous research, findings suggest that MAP's measure of reading comprehension explained more variance in CBM-R for students in third through fifth grade than for students in first and second grade. Results are also consistent with prior research suggesting

that reading comprehension explains less variance in CBM-R for students in the lower elementary grades (Ardoin et al., 2013). Grade level differences may be due, in part, to the likelihood that students in the upper elementary grades receive less instruction (if any) in foundational reading skills, because it is assumed they have mastered those skills by third grade.

Technical Adequacy of MAP for Universal Screening

With the exception of the present study, only one empirical study (Merino & Beckman, 2010) evaluated the technical adequacy of the MAP for universal screening. Current findings provide further evidence of the technical adequacy of MAP for universal screening in Grades 2 through 5 and provide initial evidence of MAP's concurrent validity in first grade. The current study also addressed a gap in the literature by examining what benefit there is, if any, to administering both MAP and CBM-R for universal screening. Results indicated that MAP scores explained variance in third graders' ITBS-TR scores above and beyond CBM-R. CBM-R, however, failed to explain variance above and beyond MAP. These findings are not surprising, given that both the MAP and ITBS are comprehensive measures of reading achievement, assessing basic reading skills, vocabulary, and reading comprehension. Furthermore, MAP and ITBS share method variance, in that the manner in which they measure these skills (i.e., multiple choice format) are similar. In contrast, despite CBM-R being described as a global outcome measure (Fuchs & Deno, 1991), it only directly measures students' oral reading rate with accuracy.

Results suggest that, if schools are deciding between MAP and CBM-R for universal screening, it might be best to administer the MAP. However, given the expense of the MAP, it may not be feasible for some districts to purchase it for use in their schools; therefore, CBM-R remains a technically adequate choice for universal screening. In addition to CBM-R's cost

advantage over MAP, administering CBM-R during universal screening is necessary for those schools that are monitoring students' progress with CBM-R because the resultant data provide teachers a baseline from which to develop CBM-R reading goals.

Limitations

Findings from the current study significantly add to extant CBM-R and MAP literature; however, multiple limitations must be considered. First, the current sample was a fairly homogeneous group of participants enrolled in two schools within the same high performing district in the southeast. As such, both schools utilized the same curriculum. Therefore, the current study should be replicated with a sample more representative of the United States with regard to race, ethnicity, and socioeconomic background and with lower performing students.

A second limitation is that the MAP assessment employed in this study was designed specifically for Georgia, as it is reportedly developed for each state in which it is administered (Northwest Education Association, 2013a). Starting in fall 2012, the publishers of the MAP modified the goal areas of the MAP to align with the Common Core Standards. The district in which the MAP was administered switched to this version of the MAP at the start of the 2012-2013 academic year. Although Georgia has adopted the Common Core Standards, publishers of the MAP contend that it may vary state to state, because states were able to adopt up to 15% of their own standards when adopting the Common Core (Northwest Evaluation Association, 2013a). Therefore, replication of this study with students from states other than Georgia is necessary to determine if current findings hold with MAP tests aligned to slightly different standards.

A third limitation of the study is that the ITBS was only administered to third-grade students. Therefore, although the current study provides initial evidence of the concurrent

validity of MAP with ITBS, these results can only be generalized to third-grade students. Given that the MAP is designed to be administered to students in kindergarten through 12th grade, additional research is necessary to establish the concurrent and predictive validity of MAP with nationally-normed assessments of reading achievement. A final limitation is that prior to the MAP's alignment to the common core standards, it included a reading comprehension goal area. With the revisions made to MAP in order to align to the Common Core Standards, there are two goal areas (literature texts and informational texts) which not only measure students' reading comprehension, but also their understanding of the craft and structure of informational and narrative texts as well (Northwest Education Association, 2013b). Therefore, comparisons between the current and previous studies employing MAP or other assessments of reading comprehension only may be limited.

Implications and Future Directions

Current findings provide strong support for the use of MAP for universal screening within schools and have important implications for future research. If a school's purpose for universal screening is to identify the lowest performing students, it is essential to determine the extent to which a universal screener accurately identifies students who are "at-risk." Research indicates that CBM-R is technically adequate for identifying "at-risk" students (Ball & Christ, 2012); however, this has yet to be evaluated with MAP. For a universal screener to be appropriate, it must be technically adequate and demonstrate classification accuracy (Glover & Albers, 2007). Thus far, only one study in addition to the current study provides evidence of the technical adequacy of the MAP and none have evaluated its ability to accurately identify "at-risk" students. Therefore, future research is necessary to examine the classification accuracy of MAP data for universal screening. In addition to investigations of the classification accuracy of

MAP scores, research is needed to evaluate the treatment utility of MAP data with the online instructional decision-making manual and the degree to which instructional changes lead to increases in student achievement.

References

Ardoin, S. P., Eckert, T. L., Christ, T. J., Morena, L. S., January, S.-A. A., & Hine, J. F. (2013). Examining variance in reading comprehension among developing readers: Words in context (curriculum-based measurement in reading) versus words out of context (word lists). *School Psychology Review, 42*(3), 243-261.

Ardoin, S. P., Witt, J. C., Suldo, S. M., Connel, J. E., Koenig, J. L., Resetar, J. L., & Slider, N. J. (2004). Examining the incremental benefits of administering a maze and three versus one curriculum-based measurement reading probe when conducting universal screening. *School Psychology Review, 33*, 218-233.

Baker, S.K., Smoklowski, K., Katz, R., Fien, H., Seely, J.R., Kame'enui, E.J., & Beck, C.T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review, 37* (1), 18-37.

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. doi: 10.1002/pits.21592

Barnett, W. S. (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *Future of Children, 5*(3), 25-50. doi: 10.2307/1602366

Betts, J., Pickard, M., & Heistad, D. (2009). An investigation of the psychometric evidence of CBM-R passage equivalence: Utility of readability statistics and equating the alternate forms. *Journal of School Psychology, 47*, 1-17. doi: 10.1016/j.jsp.2008.09.001

Busch, T. W., & Reschly, A. L. (2007). Progress monitoring in reading: Using curriculum-based measurement in a response-to-intervention model. *Assessment for Effective Intervention, 32*(4), 223-230. doi: 10.1177/15345084070320040401

Christ, T. J., Ardoin, S. P., & Eckert, T. L. (2010). *Formative Assessment Instrumentation and Procedures for Reading (FAIP-R): Oral reading passages*. University of Minnesota. Retrieved from www.fast.cehd.umn.edu

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184-192. doi: 10.1177/00224669030370030801

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 49*(1), 36-45.

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callendar, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools, 46*(1), 44-55. doi: 10.1002/pits.20353

Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long, J. D. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research & Practice, 25*(2), 60-75. doi: 10.1111/j.1540-5826.2010.00304.x

Faykus, S. P., & McCurdy, B. L. (1998). Evaluating the sensitivity of the maze as an index of reading proficiency for students who are severely deficient in reading. *Education and Treatment of Children, 21*(1), 1-21.

Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education, 34*(4), 226-236. doi: 10.1177/002246690103400405

Foegen, A., Jiban, C., & Deno, S. (2007). Progress monitoring measures in mathematics: A review of the literature. *Journal of Special Education, 41*(2), 121-139. doi: 10.1177/00224669070410020101

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-Intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities, Research and Practice, 18*(3), 157-171. doi: 10.1111/1540-5826.00072

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488-501.

Fuchs, L. S., & Fuchs, D. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21*(1), 45-58.

Fuchs, L. S., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9*(2), 20-28. doi: 10.1177/074193258800900206

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117-135. doi: 10.1016/j.jsp.2006.05.005

Graney, S. B., Martinez, R. S., Missall, K. N., & Aricak, O. T. (2010). Universal screening of reading in late elementary school: R-CBM versus CBM maze. *Remedial and Special Education, 31*(5), 368-377. doi: 10.1177/0741932509338371

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*(2), 228-240.

Hintze, J. M., Shapiro, E. S., Conte, K. L., & Basile, I. M. (1997). Oral reading fluency and authentic reading material: Criterion validity of the technical features of CBM survey-level assessment. *School Psychology Review, 26*(4), 535-553.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*(3), 372-386.

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*(1), 9-26.

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10-17.

Howe, K. B., & Shinn, M. M. (2002). *Standard reading assessment passages (RAPs) for use in general outcome measurement: A manual describing development and technical features.* Retrieved 5/2/2013 from https://aimswebtraining.pearson.com/downloads/PassagesTechnicalManual08252002.pdf

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003a). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice, 18*(4), 237-245. doi: 10.1111/1540-5826.00078

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003b). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, *95*(4), 719-729. doi: 10.1037/0022-0663.95.4.719

Jenkins, J. R., & Jewell, M. (1993). Examining the validity of two measures for formative teaching: Reading aloud and maze. *Exceptional Children*, *59*(5), 421-432.

McGlinchey, M. T., & Hixson, M. D. (2004). Using curriculum-based measurement to predict performance on state assessments in reading. *School Psychology Review, 33*(2), 193-203.

McMaster, K. L., Wayman, M. M., & Cao, M. (2006). Monitoring the reading progress of secondary-level English learners: Technical features of oral reading and maze tasks. *Assessment for Effective Intervention, 31*(4), 17-31. doi: 10.1177/073724770603100402

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24*(4), 186-195. doi: 10.1111/j.1540-5826.2009.00292.x

Mercer, S. H., Martínez, R. S., Faust, D., & Mitchell, R. R. (2012). Criterion-related validity of curriculum-based measurement in writing with narrative and expository prompts relative to passage copying speed in 10th grade students. *School Psychology Quarterly, 27*(2), 85-95. doi: 10.1037/a0029123

Merino, K., & Beckman, T. O. (2010). Using reading curriculum-based measurements as predictors for the Measures of Academic Progress (MAP) standardized test in Nebraska. *International Journal of Psychology: A Biopsychosocial Approach, 6*, 85-98.

Northwest Evaluation Association. (2009). Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades. Lake Oswego, OR: Author.

Northwest Evaluation Association. (2011). 2011 Normative Data. Retrieved 9/13/12, from http://www.nwea.org/support/article/normative-data-2011

Northwest Evaluation Association. (2012). DesCartes: A continuum of learning. Retrieved 9/19/12, from http://www.nwea.org/products-services/classroom-resources/descartes

Northwest Evaluation Association. (2013a). Common Core MAP® and MAP® for Primary Grades. Retrieved 2/6/13, from

http://www.nwea.org/sites/www.nwea.org/files/resources/Common%20Core%20MAP%20MPG%20FAQ%20July2012.pdf

Northwest Evaluation Association. (2013b). Measures of Academic Progress (MAP) Common Core Aligned Version 4. Retrieved 2/6/13, from http://www.nwea.org/sites/www.nwea.org/files/resources/MAP%20CC%20V4%20Goal%20Structure.pdf

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*(6), 427-469. doi: 10.1016/j.jsp.2009.07.001

Reschly, D. J., & Bergstrom, M. (2009). Response to Intervention. In T. B. Gutkin & C. R. Reynolds (Eds.), *The Handbook of School Psychology* (4th ed., pp. 434-460): Wiley.

Schilling, S. G., Carlisle, J. F., Scott, S. E., & Zeng, J. (2007). Are fluency measures accurate predictors of reading achievement? *Elementary School Journal, 107*(5), 429-448. doi: 10.1086/518622

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

Ticha, R., Espin, C. A., & Wayman, M. M. (2009). Reading progress monitoring for secondary-school students: Reliability, validity, and sensitivity to growth of reading-aloud and maze-selection measures. *Learning Disabilities Research & Practice, 24*(3), 132-142. doi: 10.1111/j.1540-5826.2009.00287.x

University of Iowa. (2005). *Iowa Test of Basic Skills, Form A*. Rolling Meadows, IL: Riverside Publishing.

U.S. Department of Education. (2001). No Child Left Behind Act of 2001. Retrieved 9/8/2012, from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

VanDerHeyden, A.M., Witt, J.C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Intervention (RTI) model on identification of children for special education. *Journal of School Psychology, 45*(2), 225-256. doi: 10.1016/j.jsp.2006.11.004

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.

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*(4), 207-214.

Table 2.1

Means and Standard Deviations Among Study Variables for Grades 1-5

	First Grade[a]		Second Grade[b]		Third Grade[c]		Fourth Grade[d]		Fifth Grade[e]	
Variables	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*	Mean	*(SD)*
Median CBM-R	75.75	(39.32)	109.68	(40.93)	128.59	(36.07)	155.06	(35.48)	148.98	(35.80)
Transformed CBM-R	35.34	(19.17)	92.21	(37.16)	131.57	(38.09)	146.17	(31.99)	159.42	(33.28)
MPG Overall	176.87	(13.56)								
MPG Foundational Skills	176.77	(16.59)								
MPG Vocabulary Use and Functions	177.91	(13.62)								
MPG Literature & Informational Texts	174.30	(16.95)								
MAP Overall			190.63	(15.09)	202.41	(14.88)	210.94	(11.20)	214.78	(13.73)
MAP Foundational Skills & Vocabulary			190.41	(14.89)	201.76	(15.21)	210.70	(11.36)	214.05	(13.93)
MAP Literature Texts			191.58	(16.61)	202.96	(15.26)	210.94	(13.26)	214.80	(15.01)
MAP Informational Texts			189.80	(16.59)	202.25	(16.87)	211.23	(12.27)	215.39	(15.06)
ITBS-TR					70.58	(26.05)				

Note. CBM-R = Curriculum-based measurement of oral reading; MPG = Measures of Academic Progress Primary Grades; MAP = Measures of Academic Progress; ITBS-TR = Iowa Test of Basic Skills-Total Reading Standard Score.

[a]n = 158, [b]n = 158, [c]n = 147, [d]n = 159, [e]n = 180

Table 2.2
Intercorrelations Among Study Variables for Grade 1 and Across Grades 2-5

Variables	1.	2.	3.	4.	5.
Grade 1					
1. Median CBM-R	-				
2. MPG Overall	.73**	-			
3. MPG Foundational Skills	.65**	.92**	-		
4. MPG Vocabulary Use and Functions	.60**	.90**	.75**	-	
5. MPG Literature & Informational Texts	.64**	.93**	.81**	.79**	-
Grades 2-5					
1. Transformed CBM-R	-				
2. MAP Overall	.83*	-			
3. MAP Foundational Skills & Vocabulary	.79*	.94*	-		
4. MAP Literature Texts	.78*	.96*	.85*	-	
5. MAP Informational Texts	.80*	.96*	.86*	.89*	-

Note. CBM-R = Curriculum-based measurement of oral reading; MPG = Measures of Academic Progress Primary Grades; MAP = Measures of Academic Progress.
* $p < .01$
** $p < .001$

Table 2.3
Summary of Hierarchical Regression Analyses Using MAP Goal Areas to Predict CBM-R Performance for Grades 1-5

Grade	Predictor and Steps		R^2	ΔR^2
1	1	Foundational Skills Vocabulary	.45*	
	2	Literature and Informational Texts	.47	.04
2	1	Foundational Skills and Vocabulary	.56*	
	2	Literature Texts Informational Texts	.60*	.04*
3	1	Foundational Skills and Vocabulary	.51*	
	2	Literature Texts Informational Texts	.61*	.10*
4	1	Foundational Skills and Vocabulary	.39*	
	2	Literature Texts Informational Texts	.52*	.13*
5	1	Foundational Skills and Vocabulary	.46*	
	2	Literature Texts Informational Texts	.55*	.09*

Note. CBM-R = Curriculum-based measurement of oral reading; MAP = Measures of Academic Progress.
*$p < .006$.

Table 2.4

Summary of Hierarchical Regression Analysis Using MAP Goal Areas to Predict CBM-R Performance Across Grades 2-5

Predictor and Steps		R^2	ΔR^2
1	Foundational Skills and Vocabulary	.63*	
2	Literature Texts Informational Texts	.69*	.06*

Note. CBM-R = Curriculum-based measurement of oral reading; MAP = Measures of Academic Progress.
*$p < .001$.

Table 2.5
Summary of Hierarchical Regression Analyses Using MAP and CBM-R to Predict ITBS-Total Reading in Third Grade

Model	Predictor and Steps	R^2	ΔR^2
A.	1 CBM-R	.55*	
	2 MAP	.76*	.21*
B.	1 MAP	.75*	
	2 CBM-R	.76	.01

Note. CBM-R = Curriculum-based measurement of oral reading; MAP = Measures of Academic Progress; ITBS = Iowa Test of Basic Skills

*$p < .001$.

CHAPTER 3

TEACHERS' KNOWLEDGE, USE, AND ACCEPTABILITY OF CURRICULUM-BASED MEASUREMENT IN READING AND THE MEASURES OF ACADEMIC PROGRESS[2]

[2] January, S.-A. A. & Ardoin, S. P. To be submitted for publication.

Abstract

Universal screening assessments such as curriculum-based measurement of oral reading (CBM-R) and the Measures of Academic Progress (MAP) are utilized by thousands of elementary schools across the country. Although research indicates that acceptability is an important criterion for universal screening assessments, evidence of the acceptability of CBM-R and MAP is either outdated or nonexistent. Therefore, the current study examined the extent of teacher's knowledge, use, and acceptability of CBM-R and MAP for universal screening. Participants were 86 first- through fifth-grade teachers who completed a 55-item universal screening assessments questionnaire. Results provide initial evidence regarding the acceptability of MAP for universal screening and extend outdated information regarding teacher's perceptions of CBM-R as a universal screener and as an indicator of students' reading comprehension.

Introduction

A recent report estimates that states spend $1.7 billion on assessments each year (Chingos, 2012) and, in light of federal laws such as No Child Left Behind (U.S. Department of Education, 2001), assessment of students' academic skills is an important task for educators. Purposes of assessment include classification, diagnosis, intervention, evaluating the effectiveness of interventions, and screening for those who are "at-risk" for adverse outcomes (Sattler, 2002). To accomplish the aforementioned tasks, school use various types of assessments such as state achievement tests, group norm-referenced achievement tests, and universal screening measures. Universal screening consists of the administration of assessments to all students within a school three to four times yearly. Schools employ universal screening measures to (a) identify which students are "at-risk" for not meeting academic standards (Mellard, McKnight, & Woods, 2009), (b) predict which students may not achieve proficiency

on state end-of-year tests (Hintze & Silberglitt, 2005; Wiley & Deno, 2005), (c) plan and modify instruction/intervention (Reschly & Bergstrom, 2009), and (d) evaluate the effectiveness of the general curriculum (Reschly & Bergstrom, 2009). Two assessments often employed for universal screening are Computer Adaptive Tests (CATs) and curriculum-based measurement of oral reading (CBM-R). The current study focused on teachers' acceptability of universal screening procedures within schools.

Universal Screening Assessments

Research indicates that universal screening measures should meet certain criteria to be useful. Hosp and Ardoin (2008) suggest that to demonstrate utility, universal screening assessments must be short, easy to administer by teachers and other school staff, inexpensive, and reliable and valid. Glover and Albers (2007) recommend that universal screeners must also demonstrate utility for making instructional decisions and be acceptable to teachers and school personnel. The extent to which commonly employed universal screeners meet these characteristics varies.

Measures of Academic Progress. The Measures of Academic Progress (MAP), one type of CAT, is an individually administered multiple-choice computer-based assessment that presents a unique set of test items to each test taker. The MAP provides assessments in reading (among other academic areas) for students in kindergarten through 12th grade (Northwest Evaluation Association, 2009). One advantage of the MAP is that as a CAT, it presents more items that are near a student's performance level than traditional standardized achievement tests. Therefore, the standard error of measurement is typically lower than that of traditional tests, which results in increased measurement precision and greater confidence that resultant scores accurately estimate students' reading achievement. Another reported advantage of the MAP

over some other universal screeners (e.g., CBM) is that it provides information regarding students' academic strengths and weaknesses. For example, the MAP Reading Survey with Goals assessment provides teachers with estimates of students' competencies in (a) foundational skills and vocabulary; (b) literature texts; and (c) informational texts. The publisher of the MAP also provides teachers with a manual (i.e., DesCartes; Northwest Evaluation Association, 2012) that provides instructional recommendations based upon students' performance on each of the subtests. The validity of those recommendations have, however, yet to be tested. Unfortunately, recent research suggests that utilizing MAP for making instructional decisions may not actually improve student achievement (Cordray, Pion, Brandt, Molefe, & Toby, 2012).

In regard to its utility as a universal screener, the MAP only meets two of the seven criteria suggested by Hosp and Ardoin (2008) and Glover and Albers (2007). MAP administration meets the criterion of being short, as the time required for students to complete the reading portion of the MAP is estimated at 65 min, which is approximately the time required to individually administer CBM-R to a classroom of students. The MAP also meets the criterion of being easy to for teachers to administer as teachers simply have to log students into a web-based system. Computer administration also minimizes and potentially eliminates administration errors and the need for administration training. MAP does not, however, meet the universal screening criterion of being inexpensive. As compared to other universal screeners that cost only $2-$5 per student screened, the MAP is costly (i.e., $13 per student). Despite the expense of MAP, it is widely used, even though it lacks extensive research supporting its reliability and validity. However, growing evidence suggests that MAP meets the criteria of being a reliable and valid measure of reading achievement (January & Ardoin, 2013; Merino & Beckman, 2010). The sixth criterion suggested by researchers is that universal screeners must demonstrate

decision-making utility. Although publishers of the MAP suggest assessment results can be used in conjunction with an online resource (i.e., DesCartes) to make instructional decisions, a recent report indicates MAP may not actually be useful for instructional decision-making (Cordray, Pion, Brandt, Molefe, & Toby, 2012). Whether MAP meets the final criterion of teacher acceptability has yet to be empirically examined. There is also a lack of evidence regarding teachers' knowledge and understanding of the skills measured by MAP.

Curriculum-Based Measurement of Oral Reading. CBM-R is a timed, individually administered, fluency-based assessment of oral reading (Fuchs, 2004; Reschly et al., 2009). With CBM-R, students read a passage aloud for 1 min while the examiner records reading errors and then calculates the number of words read correctly per minute (WRCM; Deno, Fuchs, Marston, & Shin, 2001; Hosp & Hosp, 2003). Although CBM-R was developed for special education teachers to monitor their students' progress (Deno, 1985, 2003), it is widely employed by elementary schools for universal screening purposes (Busch & Espin, 2003; Mellard et al., 2009). A primary feature of CBM-R is that it is a general outcome measure; as such, it assesses global proficiency within a curriculum across the year, instead of assessing mastery of a set of hierarchically-organized subskills (Fuchs & Deno, 1991).

When considering the properties of universal screeners suggested by researchers (Glover & Albers, 2007; Hosp & Ardoin, 2008), CBM-R meets 6 of 7 criteria. CBM-R is quick to administer, easy to use by teachers and school staff, inexpensive and the resultant data are reliable and valid (Ball & Christ, 2012; Deno, 1985; Wayman, Wallace, Wiley, Ticha, & Espin, 2007). Regarding decision-making utility, research suggests that CBM-R generally meets this criterion. Despite a lack of evidence for the rules used to make instructional decisions with CBM-R (Ardoin, Christ, Morena, Cormier & Klingbeil, 2013), extant research indicates that

teachers make more frequent instructional changes and, subsequently, students make greater achievement gains when CBM-R is used for progress monitoring (Stecker, Fuchs, & Fuchs, 2005). Although teachers use CBM-R for instructional decision-making, the degree to which it meets the final criterion of teacher acceptability is questionable. Despite ample evidence that CBM-R is a strong predictor of reading comprehension (Ardoin et al., 2013; Jenkins et al., 2003), researchers speculate that teachers think CBM-R is simply a measure of students' word reading skills (Hamilton & Shinn, 2003), thus diminishing its acceptability to teachers. Unfortunately, even with training regarding CBM-R's utility and validity, teachers reportedly still doubt its value as a predictor of reading comprehension (Foegen, Espin, Allinder, & Markell, 2001). The use of CBM-R and the emphasis on reading fluency has greatly increased in recent years, potentially altering teachers' acceptability and knowledge regarding CBM-R.

Acceptability of Universal Screening Assessments

Prior research on treatment acceptability, or the degree to which the person implementing an intervention views it as attractive, fair, reasonable, and appropriate (Kazdin, 2000), indicates that acceptability, among other factors, influences treatment fidelity (Perepletchikova & Kazdin, 2005). Similarly, assessment acceptability, which is studied less frequently than treatment acceptability, may be defined as "the degree to which an assessment method is appropriate for analyzing a given problem, the fairness of the method, and its intrusiveness into teachers' and students' instructional time... [and] the degree to which the measures help in developing treatment plans" (Shapiro & Eckert, 1994 pp. 168-169). Prior research indicates that teachers who find CBM more acceptable are more likely to utilize the resultant data (Allinder & Oats, 1997).

Despite the importance of acceptability (Allinder & Oats, 1997) and an increase in the utilization of CBM-R for universal screening (Clemens, Hilt-Panahon, Shapiro, & Yoon, 2012; Mellard et al., 2009; Northwest Evaluation Association, 2011), there are only two studies that provide relevant information regarding CBM-R's acceptability. The first study, conducted in 1995 by Eckert and Shapiro, examined teachers' acceptability of two assessment methods, curriculum-based assessment (CBA) and norm-referenced tests (NRTs). CBA is a set of assessment methods that measure students' achievement level in a school's curriculum (Tucker, 1985). CBM-R is one type of CBA. In this study, teachers read a case summary with assessment data featuring either CBA or NRTs. After reading the case summary, teachers completed a questionnaire regarding their acceptability of CBA or NRTs. Findings indicated that teachers rated CBA as more acceptable than published NRTs.

The second study, conducted by Foegen, Espin, Allinder, and Markell (2001), examined pre-service teachers' beliefs regarding CBM-R. Participants received training about the validity and utility of CBM-R then completed a 15-item Teacher Belief Survey. Results indicated that overall, teachers believed in the utility of CBM-R but doubted its validity as a measure of reading comprehension. Unfortunately, these two studies (Eckert & Shapiro, 1995; Foegen et al., 2001) are somewhat outdated. Both were published over 10 years ago and thus pre-dated the wide use of CBM-R as part of Response to Intervention (RtI) models.

Purpose of the Study

Universal screeners such as MAP and CBM-R are employed by schools across the country several times per year to identify students who may not meet reading benchmark goals and to inform reading instruction (Reschly & Bergstrom, 2009). Not only must universal screening measures be technically sound (Glover & Albers, 2007; Hosp & Ardoin, 2008), they

must also be acceptable to teachers (Glover & Albers, 2007); however, no published studies examining teacher acceptability of the MAP exist. Evidence indicates that high levels of assessment acceptability leads to improved decision making and student outcomes (Allinder & Oats, 1997); however, the most recent research questions the degree to which teachers believe CBM-R is a valid indicator of reading comprehension (Foegen et al., 2001). Therefore, the purpose of the current study was to examine teachers' knowledge, use, and acceptability of MAP and CBM-R for universal screening.

This study aimed to address the following research questions:

(a) What is the current state of teachers' knowledge, use, and acceptability of MAP and CBM-R?

(b) Do teachers' knowledge, use, and acceptability of universal screeners differ as a function of the type of assessment (MAP vs. CBM-R)?

(c) Does greater knowledge and acceptability lead to increased use of MAP and CBM-R?

Method

Participants and Setting

Teachers. A total of 86 elementary school teachers from five elementary schools located within one district participated in the current study. Teachers self-reported their age (n = 77) and gender (n = 85). According to self-report, teachers ranged in age from 23 to 63 years (M = 44 years, SD = 9.5 years), with 89.5% self-reporting as female. Remaining demographic information is presented in the Results section.

Schools. As of 2011-2012, the school district in which the five schools were located enrolled 6,680 students in pre-kindergarten to 12th grade. Across the district, 2,445 students across six schools were enrolled in Grades 1 (492 students), 2 (485 students), 3 (441 students), 4

(506 students), and 5 (521 students). The teachers in one elementary school in the district did not participate as their principal did not respond to the request by the experimenter to administer the survey to teachers.

Measure

Teacher Questionnaire. An investigator-developed questionnaire was used to measure the extent of teachers' knowledge, use, and acceptability of CBM-R and the MAP. The questionnaire was adapted from the Teacher Belief Survey (Foegen, et al., 2001) and the CBM-Acceptability Scale (Allinder & Oats, 1997) employed in prior research. The questionnaire included 55-items (provided in Appendix A) developed to assess teachers' knowledge, use, and acceptability of CBM-R and the MAP. Of the 55 items, 15 items were True/False, 36 items were Likert-like items, 2 items were multiple-choice questions, and 2 items were open-ended questions. The first 15 items aimed to assess teachers' knowledge of MAP and CBM-R. Participants then responded to 18 items regarding their use of CBM-R and MAP using a 6-point Likert-like scale. Responses ranged from never (1) to daily (6). Teachers rated another 18 questions regarding their acceptability of CBM-R and MAP on a 6-point Likert-like scale. Responses ranged from strongly disagree (1) to strongly agree (6). Each completed survey yielded a percent accuracy of knowledge regarding CBM-R and MAP and summed scores for CBM-R and MAP in the domains of use and acceptability.

Procedures

The researcher first emailed the principals of six elementary schools in the participating district, inviting them to participate in the study. After principals from 5 of the 6 schools agreed, a day and time for data collection was scheduled for each school. Potential participants were recruited during the after school faculty meetings at the participating elementary schools. At

these meetings, the details of the study were explained to participants. Teachers who agreed to participate received the consent documents and a copy of the survey. Participants completed the survey in 10 to 15 min. All participants were entered into a drawing to win one of two $25 e-gift cards.

Data Analysis

This study employed several analytical methods to investigate the research questions. First, descriptive statistics and frequencies were obtained for each survey question and the sum of responses for each domain. Then, three paired samples *t*-tests were used to examine differences in participants' summed knowledge, use, and acceptability scores for MAP and CBM-R. Next, six multiple regression analyses were conducted to determine the extent to which teachers' knowledge and acceptability predicted their use of CBM-R and MAP. In an effort to minimize Type 1 error, a Bonferroni connection was applied and an alpha level of .008 was employed across all six regression models. For the multiple-choice items, the percentage of respondents that selected each response was calculated. Finally, for the open-ended items, participants' responses were recorded and examined for common themes. Next, each response was assigned a code (1-9) that corresponded with an identified theme. Lastly, the number of responses for each theme was summed and the percent of respondents identifying each theme was calculated.

Results

The average percent accuracy for the questions on the knowledge domain are presented in Table 3.1 and means, standard deviations, and medians for the Likert-like questions on the use and acceptability domains are presented in Table 3.2. Internal consistency of the survey, as estimated by Cronbach's Alpha, was .92. Three one-way ANOVAs were conducted to examine

differences in participant responses based on self-reported demographic variables. Results indicated that respondents differed in their responses to the knowledge and use items with regard to their training on the MAP. This section presents descriptive results first, then inferential results.

Descriptive Results

Participant Characteristics. All participants self-reported their demographic information, but not all respondents responded to each question. Overall, 25.6% (n = 22) held a Bachelor's degree, 45.3% (n = 39) earned a Master's degree, 22.1% (n = 19) held an Educational Specialist degree, and 3.5% (n = 3) held Doctorate degrees. There were 71 general education teachers that taught first (n = 15), second (n = 13), third (n = 16), fourth (n = 11), fifth (n = 14), or multiple (n = 2) grades. Ten respondents reported teaching special education and three reported teaching both general and special education. On average, teachers had 17 years' experience (range, 1 to 39) and used CBM-R for 4.9 years (SD = 3.9 years, range, 1 to 20 years) and the MAP for 2.1 years (SD = 0.7 years, range, 1 to 4 years). Teachers reported receiving an average of 1.2 hours (SD = 1.7 hours, range, 0 to 8 hours) of training on CBM-R and 5 hours (SD = 6 hours, range, 0 to 40 hours) of training on MAP, with 45.3% of teachers being trained in CBM-R by the district in which they were currently working.

Knowledge of Universal Screening Assessments. Teachers responded to 15 True/False items regarding their knowledge of MAP and CBM-R. Participants were more accurate with MAP items (M = 92.6%; SD = 14.82) than with CBM-R items (M = 58.1%; SD = 21.67), t = 12.46, $p < .001$. Responses to the most frequently missed items indicated that most respondents did not know that CBM-R is a good predictor of reading comprehension or that students can

make greater gains in reading if CBM-R is used to evaluate the impact of instruction on student performance.

Use of Universal Screening Assessments. Participants responded to 18 Likert-like items on a scale of never (1) to daily (6) regarding their use of MAP and CBM-R. Overall, teachers reported using MAP (M = 31.25; SD = 7.69) significantly more than CBM-R (M = 24.92; SD = 9.41), t = 5.88, $p < .001$. With regard to the use of MAP, teachers most frequently indicated using the MAP for providing differentiated instruction to individual students (M = 4.28; SD = 1.24) and planning their classroom reading instruction (M = 3.86; SD = 1.26). In contrast, teachers reported using results from CBM-R universal screening most often to identify struggling readers (M = 3.31; SD = 1.18) and to determine which students were in need of additional instruction/intervention in reading (M = 3.19; SD = 1.15). Additionally, most respondents reported rarely using CBM-R to obtain an estimate of their students' reading comprehension skills (M = 1.70; SD = 1.23).

Acceptability of Universal Screening Assessments. Teachers responded to 18 Likert-like items on a scale of strongly disagree (1) to strongly agree (6) regarding the extent to which they found MAP and CBM-R acceptable. Results indicated high degrees of acceptability with MAP (M = 37.51; SD = 6.95) and CBM-R (M = 37.39; SD = 8.38), which did not differ significantly, t = .198, p = .844. Most respondents like having MAP data for their students (M = 5.15; SD = .89), would suggest that other teachers use MAP data to make instructional decisions (M = 4.74; SD = 1.03), think that collecting MAP data is beneficial for students (M = 4.88; SD = .93), and believe that it is appropriate for evaluating their students' academic problems (M = 4.76; SD = .94), reading skills (M = 4.72; SD = 1.06) and reading progress (M = 4.72; SD = .99). Regarding CBM-R, most teachers like having CBM-R data for their students (M = 4.48; SD =

1.05), think it is practical in its administration time (M = 4.67; SD = .86), and that collecting CBM-R data is beneficial for students (M = 4.36; SD = 1.06). Interestingly, teachers believe CBM-R should have a comprehension question at the end (M = 2.30; SD = 1.28) and that it is simply a measure of students' decoding skills (M = 2.91; SD = 1.12).

Multiple-Choice and Open-Ended Items. When asked to select one screening instrument (MAP or CBM-R) that best measures comprehension and reading achievement, most participants thought that MAP best measures students' reading comprehension (95%) and global reading achievement (82%). Participants also responded to two open-ended questions, asking what they liked best about MAP and CBM-R. With regard to MAP, the majority of respondents (70%) indicated that they most liked the detailed information that it provides about students reading skills. Regarding CBM-R, most respondents (60%) indicated that they liked that it is a quick assessment.

Inferential Results

Multiple regression analyses were conducted to determine if participants' knowledge and acceptability predicted their use of MAP and CBM-R. Results indicated that alone, MAP knowledge accounted for only 9% of the variance in MAP use and that acceptability accounted for 27% of the variance in MAP use. When knowledge and acceptability were entered as predictors, together they accounted for 28% of the variance in MAP usage; however, knowledge did not contribute significantly to the model (p = .284). Regression analyses with CBM-R were similar to MAP results. Knowledge of CBM-R accounted for 15% of the variance in participants' use of CBM-R and acceptability accounted for 30% of the variance in CBM-R usage. Together, knowledge and acceptability of CBM-R accounted for 30% of the variance in

CBM-R use, and, similar to MAP results, CBM-R knowledge was not a significant predictor of CBM-R use (p = .303).

Discussion

Universal screening assessments such as MAP and CBM-R are utilized by schools to identify students who are "at-risk" (Mellard et al., 2009) and are in need of additional instruction/intervention (Reschly & Bergstrom, 2009). Despite recommendations that universal screeners must be technically adequate, short and easy to administer, inexpensive, acceptable, and demonstrate decision-making utility (Glover & Albers, 2007; Hosp & Ardoin, 2008), the degree to which MAP and CBM-R meet all of these criteria varies. Although there is some evidence of CBM-R's acceptability (Eckert & Shapiro, 1995; Foegen et al., 2001), this research is outdated and no research exists regarding the acceptability of MAP for universal screening. The lack of acceptability research is also problematic because evidence suggests that teachers with higher degrees of acceptability use assessments more frequently (Allinder & Oats, 1997). Overall, findings from the current study replicate and extend the CBM-R acceptability literature and provide initial evidence regarding the acceptability of MAP.

In general, current findings indicate that teachers found both MAP and CBM-R highly acceptable, providing evidence that they meet one of the criteria suggested by Glover and Albers (2007). Present findings are also consistent with prior research indicating that a high degree of acceptability significantly predicts teachers' use of an assessment (Allinder & Oats, 1997). However, teachers demonstrated greater knowledge and use of MAP than CBM-R, which is not surprising given that the teachers in this study reported receiving more training on MAP than CBM-R, despite having used CBM-R for a longer period of time. This interesting finding highlights the importance of teachers receiving initial as well as ongoing training with feedback

regarding effective use of assessment data (Codding, Skowron & Pace, 2005; Stecker, Fuchs, & Fuchs, 2005).

Acceptability of Curriculum-Based Measurement in Reading

Findings from the current study replicate and extend previous research as teachers reported high degrees acceptability of CBM-R, which significantly predicted their use of CBM-R (Allinder & Oats, 1997; Eckert & Shapiro, 1995; Foegen et al., 2001). Consistent with prior research (Allinder & Oats, 1997; Foegen et al., 2001), teachers in the current study thought that CBM-R was practical and beneficial for assessing students; however, they doubted its validity as an estimate of students' reading comprehension. It is unlikely that one training session is enough to remedy CBM-R's lack of face validity, (Foegen et al., 2001); as such, ongoing training may be necessary. Teachers in the current study indicated that CBM-R should have a comprehension question at the end, perhaps this is a solution to CBM-R's face validity problem. However, one publisher of CBM-R probes (i.e., DIBELS) currently employs a passage retell after each probe as a measure of comprehension, but the reliability and validity of that data is questionable (Bellinger & DiPerna, 2011; Marcotte & Hintze, 2009).

The present study is the first to provide information regarding teachers' knowledge and use of CBM-R. Unfortunately, respondents in the current study indicated that they thought CBM-R was simply a measure of students' decoding skills. Although prior research speculated this (Hamilton & Shinn, 2003), the current study is the first to directly investigate this claim. Despite the fact that research indicates that using CBM-R to evaluate and modify teachers' instruction leads to greater achievement gains (see Stecker, Fuchs, & Fuchs, 2005 for a review) and that CBM-R is a strong estimate of reading comprehension (Ardoin et al., 2013; Hosp & Fuchs, 2005), teachers in the current study lacked this knowledge. It is possible that teachers'

lack of knowledge regarding CBM-R may be due in part to their reported lack of training. Despite their lower knowledge of CBM-R, teachers reported using of CBM-R for identifying struggling readers and students in need of additional instruction/intervention in reading was consistent with research-based recommendations (Deno, 2003; Mellard et al., 2009; Reschly & Bergstrom, 2009).

Acceptability of the Measures of Academic Progress

The current study is the first to provide evidence of the acceptability of MAP as a universal screening assessment. Provided that MAP is widely used, its high degree of acceptability supports, at least in part, its use as a universal screening assessment (Glover & Albers, 2007). Findings also extend prior research (Allinder & Oats, 1997), indicating that teachers' acceptability of an assessment significantly predicts their use of that measure. Interestingly, many teachers also reported using MAP to make instructional decisions, which is not surprising given that MAP publishers' advocate for the use of goal area RIT scores and provided resources (i.e., DesCartes) to determine which skills students need to be taught next. Unfortunately, there is no research to support this practice.

Limitations

Despite the significance of the current findings in supporting the acceptability of MAP and CBM-R, there are some limitations to acknowledge. The first limitation is that there were only 86 teachers from the same high-performing district that participated in the current study. It is possible that results may differ with teachers from other districts that also use both MAP and CBM-R. A second limitation is that the survey employed in this study was developed by the investigator. Nonetheless, the survey utilized in this study had strong internal consistency, as estimated by Cronbach's Alpha. A third limitation is that teachers self-reported their use of

CBM-R and MAP anecdotally. Results may differ if the study employed direct measures of teachers' actual use of data instead of their perceptions.

Implications and Future Directions

Findings from the current study have several implications for practice and research. Despite the fact that CBM-R has more empirical support than MAP, in the current study, teachers reported using CBM-R data less than MAP data. Assessing oral reading rate with accuracy is an important task for educators because fluent reading is necessary for adequate reading comprehension (LaBerge & Samuels, 1974) and fluency development is an important component of a comprehensive reading program (National Institute of Child Health and Human Development, 2000). Although MAP is strongly related to reading achievement (January & Ardoin, 2013), it does not provide an assessment of reading fluency. In addition to assessing fluency, CBM-R is a global outcome measure that requires the integration of many reading skills and thus provides a broad index of reading achievement (Fuchs & Deno, 1991; Reschly et al., 2009). Therefore, the continued use of CBM-R for universal screening and progress monitoring in schools is necessary.

Given that the use of MAP is widespread, it is likely that, similar to the report of teachers in the current study, MAP is used for making instructional decisions in schools across the country, despite the fact that this practice is not empirically supported. Although one report indicates that MAP data may lack decision-making utility (Cordray et al., 2012), future research is needed to determine the treatment utility of MAP data. Additionally, schools will benefit from knowing whether students make additional gains in reading based on the instructional decisions made using MAP data.

Findings from the current study also highlight the importance of providing teachers with adequate training regarding data-based decision making. One prerequisite of data based decision-making is to have professional development and support available for teachers (U.S. Department of Education, 2009). However, despite the increase in data collection by schools as a result of the adoption of an RtI framework and a focus on accountability, teachers report needing more professional development and training regarding the use of data (U.S. Department of Education, 2008). Extant CBM literature indicates that teachers who receive ongoing training and feedback regarding the use of CBM data for instructional decision making have students who make greater achievement gains (Stecker, Fuchs, & Fuchs, 2005). As such, schools should make an effort to provide more professional development and ongoing training and feedback for teachers regarding the use of assessment data in ways that are consistent with research.

References

Allinder, R. M., & Oats, R. G. (1997). Effects of acceptability on teachers' implementation of curriculum-based measurement and student achievement in mathematics computation. *Remedial and Special Education, 18*(2), 113-120. doi: 10.1177/074193259701800205

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 of oral reading fluency (CBM-R) decision rules. *Journal of School Psychology, 51*(1), 1-18. doi: 10.1016/j.jsp.2012.09.004

Ardoin, S. P., Eckert, T. L., Christ, T. J., Morena, L. S., January, S.-A. A., & Hine, J. F. (2013). Examining variance in reading comprehension among developing readers: Words in context (curriculum-based measurement in reading) versus words out of context (word lists). *School Psychology Review, 42*(3), 243-261.

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. doi: 10.1002/pits.21592

Bellinger, J. M., & DiPerna, J. C. (2011). Is fluency-based story retell a good indicator of reading comprehension? *Psychology in the Schools, 48*(4), 416-426.

Busch, T. W., & Espin, C. A. (2003). Using curriculum-based measurement to prevent failure and assess learning in the content areas. *Assessment for Effective Intervention, 28*(3-4), 49-58. doi: 10.1177/073724770302800306

Chingos, M. M. (2012). Strength in Numbers: State Spending on K-12 Assessment Systems. Retrieved from

http://www.brookings.edu/~/media/research/files/reports/2012/11/29%20cost%20of%20assessment%20chingos/11_assessment_chingos_final.pdf

Clemens, N. H., Hilt-Panahon, A., Shapiro, E. S., & Yoon, M. (2012). Tracing student responsiveness to intervention with early literacy skills indicators: Do they reflect growth toward text reading outcomes? *Reading Psychology, 33*(1/2), 47-77. doi: 10.1080/02702711.2011.630608

Codding, R. S., Skowron, J., & Pace, G. M. (2005). Back to basics: training teachers to interpret curriculum-based measurement data and create observable and measurable objectives. *Behavioral Interventions, 20*(3), 165-176. doi: 10.1002/bin.194

Cordray, D., Pion, G., Brandt, C., Molefe, A., & Toby, M. (2012). The impact of the Measures of Academic Progress (MAP) program on student reading achievement. Final Report. NCEE 2013-4000: National Center for Education Evaluation and Regional Assistance.

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

Deno, S. L. (2003). Developments in curriculum-based measurement. *The Journal of Special Education, 37*(3), 184-192. doi: 10.1177/00224669030370030801

Deno, S. L., Fuchs, L. S., Marston, D., & Shin, J. (2001). Using curriculum-based measurement to establish growth standards for students with learning disabilities. *School Psychology Review, 30*(4), 507-524.

Eckert, T. L., & Shapiro, E. S. (1995). Teachers' ratings of the acceptability of curriculum-based assessment methods. *School Psychology Review, 24*(3), 497.

Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education, 34*(4), 226-236. doi: 10.1177/002246690103400405

Fuchs, L. S. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188-192.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57*, 488-501.

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*(2), 117-135. doi: 10.1016/j.jsp.2006.05.005

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*(2), 228-240.

Hintze, J. M., & Silberglitt, B. (2005). A longitudinal examination of the diagnostic accuracy and predictive validity of R-CBM and high-stakes testing. *School Psychology Review, 34*(3), 372-386.

Hosp, J. L., & Ardoin, S. P. (2008). Assessment for instructional planning. *Assessment for Effective Intervention, 33*(2), 69-77. doi: 10.1177/1534508407311428

Hosp, M. K., & Fuchs, L. S. (2005). Using CBM as an indicator of decoding, word reading, and comprehension: Do the relations change with grade? *School Psychology Review, 34*(1), 9-26.

Hosp, M. K., & Hosp, J. L. (2003). Curriculum-based measurement for reading, spelling, and math: How to do it and why. *Preventing School Failure, 48*(1), 10-17.

January, S.-A. A. & Ardoin, S. P. (February, 2013). *CBM and Measures of Academic Progress in Instructional Decision Making*. Paper presented at the 2013 annual meeting of the National Association of School Psychologists, Seattle, WA.

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729. doi: 10.1037/0022-0663.95.4.719

Kazdin, A. E. (2000). Perceived barriers to treatment participation and treatment acceptability among antisocial children and their families. *Journal of Child and Family Studies, 9*(2), 157-174. doi: 10.1023/a:1009414904228

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293-323. doi: 10.1016/0010-0285(74)90015-2

Marcotte, A. M., & Hintze, J. M. (2009). Incremental and predictive utility of formative assessment methods of reading comprehension. *Journal of School Psychology, 47*(5), 315-335.

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24*(4), 186-195. doi: 10.1111/j.1540-5826.2009.00292.x

Merino, K., & Beckman, T. O. (2010). Using reading curriculum-based measurements as predictors for the Measures of Academic Progress (MAP) standardized test in Nebraska. *International Journal of Psychology: A Biopsychosocial Approach, 6*, 85-98.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific

research literature on reading and its implications for reading instruction. *NIH Publication No. 00-4769*. Washington, DC: U.S. Government Printing Office.

Northwest Evaluation Association. (2009). Technical manual for Measures of Academic Progress and Measures of Academic Progress for Primary Grades. Lake Oswego, OR: Author.

Northwest Evaluation Association. (2011). 2011 Normative Data. Retrieved 9/13/12, from http://www.nwea.org/support/article/normative-data-2011

Northwest Evaluation Association. (2012). DesCartes: A continuum of learning. Retrieved 9/19/12, from http://www.nwea.org/products-services/classroom-resources/descartes

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12*(4), 365-383. doi: 10.1093/clipsy/bpi045

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*(6), 427-469. doi: 10.1016/j.jsp.2009.07.001

Reschly, D. J., & Bergstrom, M. (2009). Response to Intervention. In T. B. Gutkin & C. R. Reynolds (Eds.), *The Handbook of School Psychology* (4th ed., pp. 434-460): Wiley.

Sattler, J. M. (2002). *Assessment of children: Behavioral and clinical applications* (4th ed.). San Diego, California: Author.

Shapiro, E. S., & Eckert, T. L. (1994). Acceptability of curriculum-based assessment by school psychologists. *Journal of School Psychology, 32*(2), 167-183. doi: 10.1016/0022-4405(94)90009-4

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. doi: 10.1002/pits.20113

Tucker, J. A. (1985). Curriculum-based assessment: An introduction. *Exceptional Children, 52*(3), 199-204.

U.S. Department of Education. (2001). No Child Left Behind Act of 2001. Retrieved 9/8/2012, from http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2008). *Teachers' use of student data systems to improve instruction: 2005 to 2007*. Washington, D.C.: Author.

U.S. Department of Education, Office of Planning, Evaluation and Policy Development. (2009). *Implementing data-informed decision making in schools: Teacher access, supports, and use*. Washington, D.C.: Author.

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.

Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education, 26*(4), 207-214.

Table 3.1

Teachers' Knowledge of CBM-R and MAP

Item	% Accuracy
ORF is not a good predictor of students' comprehension.	39.5
Using MAP data helps teachers to know if students are benefiting from their reading instruction.	93.0
ORF is useful for grouping students.	60.5
When teachers use MAP data to evaluate students' instruction, they often make greater achievement gains.	94.2
ORF provides me with information about students' general level of reading proficiency.	79.1
When teachers use MAP data, it helps them to know if students are making gains in reading achievement.	93.0
ORF provides me with information about how well students are able to decode words.	82.6
Assessing students' fluency on word lists is just as informative as assessing ORF.	76.7
MAP provides teachers with information about students' general level of reading proficiency.	89.5
Using ORF data helps teachers to know if students are benefiting from their reading instruction.	68.6
When teachers use ORF data, it helps them to know if students are making gains in reading achievement.	66.3
MAP provides teachers with information about how well students comprehend.	91.9
ORF provides me with information about how well a student comprehends.	9.3
MAP data is useful for grouping students.	94.2
When teachers use ORF data to evaluate students' instruction, they often make greater achievement gains.	40.7

Note. ORF = curriculum-based measurement of oral reading (CBM-R), MAP = Measures of Academic Progress.

Table 3.2

Teachers' Use and Acceptability of CBM-R and MAP

Item	Mean	Median	SD
I use (or have used) ORF data to plan my classroom reading instruction.	2.80	3.00	1.37
I use MAP data to provide differentiated instruction to individual students.	4.28	4.00	1.24
I use (or have used) ORF data to determine which students are struggling readers.	3.31	3.00	1.18
I use MAP data to evaluate my students' reading progress.	3.31	3.00	0.96
I use (or have used) ORF data to get a good estimate of students' overall reading achievement.	2.62	3.00	1.37
I use MAP data to plan my classroom reading instruction.	3.86	4.00	1.26
I use (or have used) ORF data to get a good estimate of students' reading comprehension.	1.70	1.00	1.23
I would choose to collect MAP data even if it was not required by my school.	3.04	3.00	1.20
I use (or have used) ORF data to evaluate my students' reading progress.	3.05	3.00	1.25
I use MAP data to get a good estimate of students' reading comprehension.	3.16	3.00	0.98
I use (or have used) ORF data to provide differentiated instruction to individual students.	2.93	3.00	1.45
I use MAP data to evaluate my students' strengths and weaknesses.	3.65	3.00	1.03
I use (or have used) ORF data to determine which students need additional instruction/intervention in reading	3.19	3.00	1.15
I use MAP data to determine which students are struggling readers.	3.33	3.00	1.14
I would choose to collect ORF data even if it was not required by my school.	2.95	3.00	1.23
I use MAP data to get a good estimate of students' overall reading achievement.	3.31	3.00	1.08
I use (or have used) ORF data to evaluate my students' strengths and weaknesses.	2.59	3.00	1.31

I use MAP data to determine which students need additional instruction/intervention in reading.	3.57	3.00	1.14
I think using ORF data is appropriate for evaluating students' academic problems.	3.46	4.00	1.31
Using MAP data is appropriate for evaluating the reading progress of children in the grade I teach.	4.71	5.00	0.99
I like the assessment procedures used in ORF.	4.08	4.00	1.13
Collecting MAP data is beneficial for children.	4.88	5.00	0.93
Using ORF data is appropriate for evaluating the reading skills of children in the grade I teach.	3.96	4.00	1.32
I like having MAP data for the students I teach.	5.15	5.00	0.89
I would suggest to other teachers that they use ORF data for making instructional decisions.	3.67	4.00	1.33
MAP is practical in the amount of time required for administration.	4.21	4.00	1.21
Using ORF data is appropriate for evaluating the reading progress of children in the grade I teach.	3.86	4.00	1.29
I think that the use of MAP data is appropriate for evaluating students' academic problems.	4.76	5.00	0.93
I like the assessment procedures used in MAP.	4.51	5.00	1.06
Collecting ORF data is beneficial for children.	4.36	5.00	1.06
I would suggest to other teachers that they use MAP data for making instructional decisions.	4.74	5.00	1.03
ORF is practical in the amount of time required for administration.	4.67	5.00	0.86
Using MAP data is appropriate for evaluating the reading skills of children in the grade I teach.	4.72	5.00	1.06
I like having ORF data for the students I teach.	4.48	5.00	1.05
I think that ORF is a simply a measure of students' decoding skills.	2.91	3.00	1.12
I think ORF should have a comprehension question at the end.	2.30	2.00	1.28

Note. ORF = curriculum-based measurement of oral reading (CBM-R), MAP = Measures of Academic Progress.

APPENDIX A: UNIVERSAL SCREENING ASSESSMENTS SURVEY

The purpose of the survey is to learn more about teachers' knowledge, use, and acceptability of universal screening assessments. This survey focuses on two commonly used procedures, curriculum-based measurement of oral reading fluency (ORF) and the Measures of Academic Progress (MAP) reading test.

ORF is a 1 minute assessment in which students read a grade-level passage aloud while the examiner records the number of errors. The number of words read correctly in 1 minute is used as the score.

The MAP is a computer-administered adaptive test that assesses students' competencies in several areas of reading; it produces a RIT score.

Part 1 Directions: Please write T (True) or F (False) for each of the following statements.

True/False:

____ ORF is not a good predictor of students' reading comprehension.
____ Using MAP data helps teachers to know if students are benefiting from their reading instruction.
____ ORF is useful for grouping students.
____ When teachers use MAP data to evaluate students' instruction, they often make greater achievement gains.
____ ORF provides me with information about students' general level of reading proficiency.
____ When teachers use MAP data, it helps them to know if students are making gains in reading achievement.
____ ORF provides me with information about how well students are able to decode words.
____ Assessing students' fluency on word list is just as informative as assessing ORF.
____ MAP provides teachers with information about students' general level of reading proficiency.
____ Using ORF data helps teachers to know if students are benefiting from their reading instruction.
____ When teachers use ORF data, it helps them to know if students are making gains in reading achievement.
____ MAP provides teachers with information about how well students comprehend.
____ ORF provides me with information about how well a student comprehends.
____ MAP data is useful for grouping students.
____ When teachers use ORF data to evaluate students' instruction, they often make greater achievement gains.

Part 2 Directions: Place a check (√) or an X in the column to indicate the how often you use each measure for the purpose listed.

Never	Yearly	3-4x Yearly	Monthly	Weekly	Daily	
						I use (or have used) ORF data to plan my classroom reading instruction.
						I use MAP data to provide differentiated instruction to individual students.
						I use (or have used) ORF data to determine which students are struggling readers.
						I use MAP data to evaluate my students' reading progress.
						I use (or have used) ORF data to get a good estimate of students' overall reading achievement.
						I use MAP data to plan my classroom reading instruction.
						I use (or have used) ORF data to get a good estimate of students' reading comprehension.
						I would choose to collect MAP data even if it was not required by my school.
						I use (or have used) ORF data to evaluate my students' reading progress.
						I use MAP data to get a good estimate of students' reading comprehension.
						I use (or have used) ORF data to provide differentiated instruction to individual students.
						I use MAP data to evaluate my students' strengths and weaknesses.
						I use (or have used) ORF data to determine which students need additional instruction/intervention in reading
						I use MAP data to determine which students are struggling readers.
						I would choose to collect ORF data even if it was not required by my school.
						I use MAP data to get a good estimate of students' overall reading achievement.
						I use (or have used) ORF data to evaluate my students' strengths and weaknesses.
						I use MAP data to determine which students need additional instruction/intervention in reading.

Part 3 Directions: Place a check (√) or an X in the column to indicate the extent to which you agree or disagree with each statement below.

Strongly Disagree	Disagree	Some-what Disagree	Some-what Agree	Agree	Strongly Agree	
						I think using ORF data is appropriate for evaluating students' academic problems.
						Using MAP data is appropriate for evaluating the reading progress of children in the grade I teach.
						I like the assessment procedures used in ORF.
						Collecting MAP data is beneficial for children.
						Using ORF data is appropriate for evaluating the reading skills of children in the grade I teach.
						I like having MAP data for the students I teach.
						I would suggest to other teachers that they use ORF data for making instructional decisions.
						MAP is practical in the amount of time required for administration.
						Using ORF data is appropriate for evaluating the reading progress of children in the grade I teach.
						I think that the use of MAP data is appropriate for evaluating students' academic problems.
						I like the assessment procedures used in MAP.
						Collecting ORF data is beneficial for children.
						I would suggest to other teachers that they use MAP data for making instructional decisions.
						ORF is practical in the amount of time required for administration.
						Using MAP data is appropriate for evaluating the reading skills of children in the grade I teach.
						I like having ORF data for the students I teach.
						I think that ORF is a simply a measure of students' decoding skills.
						I think ORF should have a comprehension question at the end.

Part 4 Directions: Please answer each question.

Which assessment best measures students' reading comprehension?

A. ORF
B. MAP

Which assessment best measures students' reading achievement?

A. ORF
B. MAP

What do you like best about ORF?

_____
_____
_____
_____
_____

What do you like best about the MAP?

_____
_____
_____
_____
_____

Age: _____

Gender: Male / Female

Number of years teaching: _____

Grade level currently teaching: 1 / 2 / 3 / 4 / 5

Number of years teaching at your current grade level: _____

Highest level of education attained: Bachelors / Masters / Ed.S. / Doctorate

I teach: General Education / Special Education

How many years have you used ORF: _____

How many years have you used MAP: _____

How many hours of training did you receive for: ORF ____ MAP ____

Did you receive your training in ORF from: This district / Another district / Both

My school currently uses: ORF / MAP / Both

CHAPTER 4

GENERAL CONCLUSION

Universal screening is an important task for schools that adopt a Response to Intervention framework. Many schools choose to administer curriculum-based measurement of oral reading (CBM-R) and/or the Measures of Academic Progress (MAP) to identify students who are "at-risk" for reading difficulties or who are in need of an intervention in reading (Deno et al., 2009; Mellard, McKnight, & Woods, 2009). Extant literature recommends that universal screeners be technically adequate, quick, easy to administer, cost effective, acceptable and have decision-making utility (Glover & Albers, 2007; Hosp & Ardoin, 2008). Although prior research indicates that CBM-R generally meets most of these criteria (Ball & Christ, 2012), due to outdated research, the degree to which CBM-R is acceptable to educators is unclear (Foegen et al., 2001). Conversely, only one empirical study exists supporting the technical adequacy of MAP (Merino & Beckman, 2010) and evidence of its acceptability is nonexistent. As such, the overall purpose of this dissertation was to evaluate the technical adequacy of MAP and the acceptability of MAP and CBM-R. Study one examined (a) the concurrent validity of MAP with CBM-R and the Iowa Test of Basic Skills-Total Reading composite; (b) how well MAP subtests predicted CBM-R performance; and (c) the benefit if any to administering MAP and CBM-R for universal screening. Study 2 evaluated (a) the acceptability of MAP for universal screening and (B) the acceptability of CBM-R for universal screening.

In general, findings from the two studies in this dissertation support the use of MAP and CBM-R for universal screening in schools. Together with prior research (Merino & Beckman,

2010) current findings support the technical adequacy of MAP for universal screening in first through fifth grade. In fact, for third graders, MAP is a better predictor of reading achievement than CBM-R. Despite MAP's demonstrated benefit over CBM-R, data from this dissertation provides more evidence that CBM-R is a strong predictor of reading achievement (Reschly et al., 2009). Therefore, schools must continue to administer CBM-R for universal screening if they plan to monitor students' reading progress with CBM-R or they cannot afford the expense of MAP (i.e., $13 per student, each year). Despite CBM-R's large body of research (Reschly et al., 2009; Wayman et al., 2007) and widespread use (Clemens, Hilt-Panahon, Shapiro, & Yoon, 2012), current results were consistent with prior research (Foegen et al., 2001) as teachers still doubt that CBM-R is a strong estimate of students' reading comprehension. Current findings added to a growing literature base (Ardoin et al., 2013; Jenkins et al., 2003; 2003b) indicating that the relationship between CBM-R and reading comprehension may be greater for students in Grades 3-5 than for students in Grades 1-2 and that CBM-R measures more than word reading skills. Unfortunately, teachers in study 2 lacked this knowledge, confirming speculation (Hamilton & Shinn, 2003) that teachers perceive CBM-R to simply be a measure of word reading skills. Although teachers demonstrated less knowledge and use of CBM-R than MAP, they reported high degrees of acceptability for both measures.

Results from this dissertation support the technical adequacy and acceptability of MAP and CBM-R for universal screening and have implications for its use within schools. It is now established that CBM-R meets all of the suggested criteria for universal screening assessments; however, there is a paucity of empirical research with the MAP. As such, future research is needed to examine the degree to which MAP meets the remaining criteria for universal screeners (e.g., decision-making utility).

References

Ardoin, S. P., Eckert, T. L., Christ, T. J., Morena, L. S., January, S.-A. A., & Hine, J. F. (2013). Examining variance in reading comprehension among developing readers: Words in context (curriculum-based measurement in reading) versus words out of context (word lists). *School Psychology Review*, *42*(3), 243-261.

Clemens, N. H., Hilt-Panahon, A., Shapiro, E. S., & Yoon, M. (2012). Tracing student responsiveness to intervention with early literacy skills indicators: Do they reflect growth toward text reading outcomes? *Reading Psychology*, *33*(1/2), 47-77. doi: 10.1080/02702711.2011.630608

Deno, S. L., Reschly, A. L., Lembke, E. S., Magnusson, D., Callendar, S. A., Windram, H., & Stachel, N. (2009). Developing a school-wide progress-monitoring system. *Psychology in the Schools*, *46*(1), 44-55. doi: 10.1002/pits.20353

Foegen, A., Espin, C. A., Allinder, R. M., & Markell, M. A. (2001). Translating research into practice: Preservice teachers' beliefs about curriculum-based measurement. *The Journal of Special Education*, *34*(4), 226-236. doi: 10.1177/002246690103400405

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology*, *45*(2), 117-135. doi: 10.1016/j.jsp.2006.05.005

Graney, S. B., Martinez, R. S., Missall, K. N., & Aricak, O. T. (2010). Universal screening of reading in late elementary school: R-CBM versus CBM maze. *Remedial and Special Education*, *31*(5), 368-377. doi: 10.1177/0741932509338371

Hamilton, C., & Shinn, M. R. (2003). Characteristics of word callers: An investigation of the accuracy of teachers' judgments of reading comprehension and oral reading skills. *School Psychology Review, 32*(2), 228-240.

Hosp, J. L., & Ardoin, S. P. (2008). Assessment for instructional planning. *Assessment for Effective Intervention, 33*(2), 69-77. doi: 10.1177/1534508407311428

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003a). Accuracy and fluency in list and context reading of skilled and RD groups: Absolute and relative performance levels. *Learning Disabilities Research & Practice, 18*(4), 237-245. doi: 10.1111/1540-5826.00078

Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003b). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology, 95*(4), 719-729. doi: 10.1037/0022-0663.95.4.719

Mellard, D. F., McKnight, M., & Woods, K. (2009). Response to intervention screening and progress-monitoring practices in 41 local schools. *Learning Disabilities Research & Practice, 24*(4), 186-195. doi: 10.1111/j.1540-5826.2009.00292.x

Merino, K., & Beckman, T. O. (2010). Using reading curriculum-based measurements as predictors for the Measures of Academic Progress (MAP) standardized test in Nebraska. *International Journal of Psychology: A Biopsychosocial Approach, 6*, 85-98.

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*(6), 427-469. doi: 10.1016/j.jsp.2009.07.001

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.