7-23-2012

Relations Among Brief Measures of Mathematics, Reading, and Processing Speed: A Construct Validity Study

Jennifer Leigh Maynard

RELATIONS AMONG BRIEF MEASURES OF MATHEMATICS, READING, AND PROCESSING SPEED: A CONSTRUCT VALIDITY STUDY

by

Jennifer Leigh Maynard

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Major: Psychology

The University of Memphis

August, 2012

DEDICATION

This dissertation represents the culmination of many years of hard work, dedication, stress, challenges, and successes. I am at a loss to find the words to express my love for my parents Bob and Candy and how awestruck I am with their unwavering love, support, and pride in what I have accomplished. They motivate me every day to work hard, do good, and be grateful for all the blessings of this life. They are my twin pillars, without whom I cannot stand. For everything my parents have given me, this is dedicated to them.

ACKNOWLEDGEMENTS

There are more people who have helped me through this journey than I can number, but s few of those individuals deserve more appreciation than I can adequately express. Overwhelming appreciation is due to Dr. Randy Floyd for his role as major professor, committee chair, teacher, mentor, and supporter. This part of my journey could not have been as successful without his support and gentle, but effective, firmness. Appreciation is also owed to my dissertation committee members Dr. Charles Blaha, Dr. Elizabeth Meisinger, and Dr. Gilbert Parra for their support and valuable feedback.

I am grateful for all of the individuals who participated in this study. I wish to express my gratitude to the University of Memphis Campus School for their willingness to allow their students to participate in this study and for their cooperation and patience during the testing process. I would also like to thank the parents and guardians who gave their children permission to participate. Most importantly, I would like to thank all of the children who participated in this study and put forth their best effort even when they were tired of seeing math problems. I am very grateful for the efforts of my fellow students who assisted with the preparation for this study and data collection. Overwhelming thanks are due to Sarah Irby for her assistance with the inter-scorer reliability part of this study.

I offer thanks to the friends who have always been there to support me and for my sister Stephanie and brother Ethan who provided me with encouragement, laughter, and love. For Jonathan, who will never know just how important his support, understanding, patience, and love throughout the most difficult parts of this process have been, I am eternally grateful.

ABSTRACT

Maynard, Jennifer Leigh. Ph.D. The University of Memphis. August/2012. Relations Among Brief Measures of Mathematics, Reading, and Processing Speed: A Construct Validity Study. Major Professor: Randy G. Floyd, Ph.

Emphasis on regular mathematics skill assessment, intervention, and progress monitoring under the RTI model has created a need for the development of assessment instruments that are psychometrically sound, reliable, universal, and brief. Important factors to consider when developing or selecting assessments for the school environment include what skills are assessed; mathematics curriculums typically include computation and applications as separate skills taught in sequence. It is also important to consider what additional factors may potentially influence performance on such tests due to the nature of test administration and characteristics of the test items. The current study investigated construct validity of established, widely-used curriculum-based measurement (CBM) tests and standardized, norm-referenced tests of mathematics as well as the potential confounding influence of processing speed and reading abilities. Construct validity of the tests administered was assessed through an investigation of convergent and discriminant validity, using confirmatory factor analysis (CFA). Numerous prespecified, theoretical models were tested to replicate previous studies suggesting specific models of mathematics ability (convergent validity) and to identify construct-irrelevant variance (discriminant validity) imposed on tests of computation and applications by processing speed and reading.

The current study extended previous work in the area of mathematics providing additional evidence for a two-factor structure of mathematics with Computation and Applications as distinct, yet related constructs and investigated the relations between mathematics constructs and processing speed and reading. Results of the current study indicated all constructs were significantly correlated with each other while mathematics constructs were more highly correlated with each other than with unrelated constructs, with the exception of Applications and Reading. Four *a priori* models of mathematics ranging from including a single factor to including four factors were tested using CFA. Results indicated that a four-factor model of mathematics including Computation, Applications, Processing Speed, and Reading as factors was the best-fitting model. The four-factor model was extended to test the construct-irrelevant variance imposed by Processing Speed on fluency-based tests as well as variance imposed by Reading on applications tests. Results indicated that in all but one case, no significant influence was contributed to fluency-based tests by Processing Speed or applications tests by Reading.

TABLE OF CONTENTS

LIST OF TABLES

Relations Among Brief Measures of Mathematics, Reading, and Processing Speed: A Construct Validity Study

According to the National Mathematics Advisory Panel, competency in mathematics is a national education goal (see http://www.ed.gov/about/bdscomm/list/mathpanel/index.html for information; Kelley, Hosp, & Howell, 2008). Skills deficits in mathematics fluency, knowledge, and problem-solving are evident in grades kindergarten through 12 despite national attention and educational reform. Skills deficits in math extend beyond the group of students who exhibit a learning disability in math (Kelley et al., 2008). Changes in school-based assessment and determination of academic deficiencies require giving more attention to the skills involved in mathematics and how they are evaluated. Historically, students exhibiting difficulty in some academic area were referred for evaluation including an assessment of cognitive and academic skills. Presence of a learning disability was determined by a sufficient discrepancy between a student's cognitive ability score and academic achievement score. It was understood that a student's scores in both areas should fall within a similar range and that low scores in an area of academic achievement without cognitive deficits would indicate a specific learning problem (Shapiro, 2004). The reauthorization of the Individuals with Disabilities Education Improvement Act (IDEIA) in 2004 (IDEIA, 2004) included a provision for the use of the Response-to-Intervention (RTI) model as a means for identifying students with learning difficulties (Fuchs, 2004; Fuchs, Fuchs, & Zumeta, 2008b). Using the RTI model, all students are screened at set intervals during the school year. Students showing the least progress are immediately provided with academic interventions designed to enhance learning and are frequently monitored for progress. Students who do not respond adequately to interventions are given more intensive interventions including special education placement.

Curriculum-based Measurement

Curriculum-based measurement (CBM) is a formative assessment method that is used to monitor progress of student academic achievement and make decisions regarding instruction methods, program modification and termination, and remedial and special education placement (Hintze, Christ, & Keller, 2002; Kelley et al., 2008; Fuchs, Fuchs & Zumeta, 2008a). CBM is a validated assessment system that serves as an alternative to traditional tests and is closely tied to the curriculum (Fuchs, 2004; Fuchs, Fuchs, & Courey, 2005). Since the inception of No Child Left Behind (NCLB), increased emphasis has been placed on monitoring student progress in several academic areas, in particular reading and mathematics. Progress monitoring under the RTI model includes the use of CBM probes to assess growth and learning over time. Traditional tests focus on mastery of individual or small subsets of skills and, therefore, typically measure mastery of a specific skill or, in the case of math, problem type. Annual high-stakes academic evaluations examine an entire curriculum of skills, but do so only at the end of the academic year. CBM represents the spectrum of skills within an academic area to be learned throughout the year (Fuchs, Fuchs, & Courey, 2005) but allows for frequent assessment of skills. These measures represent level of performance or skill development desired at the end of the year based on specific curriculum goals, therefore, CBM scores should increase as students learn required aspects of the curriculum. The slope of scores obtained from CBM probes can be used to quantify the amount of learning that is taking place, determine responsiveness to interventions, and predict trajectory throughout the year (Deno, 1985). CBM can be used by teachers to assess progress towards local and national standards, monitor growth over time, set goals for growth, and design interventions for individual students.

Unlike traditional methods of assessment, which often assess mastery of multiple skills within one assessment, CBM probes are developed to assess progress towards skills taught by an accepted curriculum. CBM probes represent the level of performance desired at the end of the year, and scores from frequent progress monitoring can be used to assess a student's actual progress as well as hypothesized trajectory toward mastery (Fuchs, 2004). Traditional tests are often administered under some form of time limit; however, those restrictions are often conservative and allow ample time for students to complete the assessment (e.g., students having a class period to complete a test given at the end of a unit). Because of the adequate amount of time given to complete traditional assessments, these assessments are a measure of *accuracy*—the number or percentage of items answered correctly (Connell, 2005). CBM probes are typically administered under strict time limits, often 1 to 8 minutes in length. Such strict time limits result in the measurement of *fluency*—how quickly problems can be accurately answered (Binder, 1996).

Whereas research in the area of reading has dominated the literature and the use of CBM measures for assessing reading ability and monitoring progress in literacy development is well established (Deno, Mirkin, & Chiang, 1982; Shinn, Tindal, & Stein, 1988), research in mathematics-based CBM measures has lagged behind (Thurber, Shinn, & Smolkowski, 2002). Math *computation* was the first area of math CBM to enjoy research and is typically referred to as M-CBM. More recently, mathematics achievement in the area of *applications* has been studied.

Mathematics CBM probes originated as assessments of single-digit basic math facts (i.e., computation skills; Jiban & Deno, 2007). Computation involves solving basic operations problems (i.e., addition, subtraction, multiplication, and division) involving single- and multiple-digit numbers, fractions, and decimals. To solve computation problems, knowledge of basic math operations is required. Because computation-only assessment requires knowledge of only one component of a comprehensive mathematics curriculum and many instructional programs focus on computation that involves conceptual concepts, assessment in more complex areas of math is necessary (Shapiro, Edwards, & Zigmond, 2005). More recent mathematics CBM probes have been developed to assess skills in applying mathematics concepts and applications (Fuchs et al., 1994). Concepts and applications require the understanding and application of math concepts to solve problems; probes include problems requiring estimation, measurement, charts and graphs, money, problem solving, and the application of more advanced mathematics concepts such as geometry and algebra (Fuchs et al., 2008a). The development of mathematics CBM probes for computation as well as concepts and applications allows for a more accurate assessment of student progress toward skills determined to be appropriate for a particular grade level.

A number of CBM systems that include probes of math ability have been developed and evaluated for technical adequacy and are widely used in school districts across the United States. Such systems include AIMSweb (PsycCorp/Pearson, 2004), Monitoring Basic Skills Progress (MBSP; Fuchs, Hamlett, & Fuchs, 1999), and easyCBM (University of Oregon, 2009). These systems provide teachers with numerous brief assessment probes for computation and probes for concepts and applications that can be used to assess student progress toward annual mathematics curriculum goals. Consistent with the focus on fluency measures in the assessment of academic skills, several standardized, norm-referenced achievement test batteries include fluency-based assessments of mathematics ability. The Woodcock-Johnson III (WJ III; Woodcock, McGrew, & Mather, 2001) includes Math Fluency, a 3-minute test of single-digit addition, subtraction, and multiplication facts. The Wechsler Individual Achievement Test, Third Edition (WIAT-III; Pearson, 2009) added three 45-second, single-skill math subtests for addition, subtraction, and multiplication.

Construct Validity

Mathematics CBM probes are useful in assessing student skills and progress and have many years of research behind them. While studies have demonstrated adequate to better validity of these probes, they generally assess the criterion validity and other psychometric properties of single or small-groups of measures. These studies have demonstrated the concurrent relations between mathematics CBM probe scores and locally used criterion assessments, such as annual state tests (Connell, 2005; Fuchs et al., 1994; Jiban & Deno, 2007; VanDerHeyden & Burns, 2008) and standardized norm-referenced achievement measures (Betts, Pickart, & Heistad, 2008; Clarke & Shinn, 2004), and have explored technical adequacy of mathematics CBM probe scores as potential indicators of growth (Foegen & Deno, 2001; Leh, Jittendra, Caskie, & Griffin, 2007). The review of the literature found here regarding construct validity is concerned with more theoretically-driven studies, rather than criterion-related validity evidence that supports the use and interpretation of single instruments or scores from particular methods. The patterns of relations across a variety of different types of mathematics measures are of interest to the current study.

Construct validity refers to the extent to which an assessment instrument or test measures a specified characteristic that is presumed to have an important influence on behavior in a given situation (Aiken, 1998). Campbell and Fiske (1959) presented the multitrait–multimethod assessment approach to examine construct validity. They proposed the assessment of construct validity by measuring multiple traits (e.g., skills or abilities) using multiple methods (e.g., tests, mean of assessment). The multitrait–multimethod (MTMM) approach asserts *convergent validity* investigates the extent to which measures of the same construct that are measured differently demonstrate high correlations. *Discriminant validity* investigates the extent to which measures of different constructs, measured in similar ways, demonstrate low correlations. Applied to measures of mathematics, construct validity is observed when individuals who are presumed to have a more highly developed understanding and knowledge of basic math facts receive higher scores on a computation test and individuals with presumed to have deficient knowledge of basic math facts receive lower scores. Convergent validity in mathematics is evidenced by measures of the same construct being more highly correlated than measures of varying constructs or constructs that appear to be unrelated. For example, it would be expected that two measures of basic math facts would be more highly correlated than a measure of basic math facts and a measure of word problem solving ability. Additionally, it would be expected that discriminant validity exists between a measure of basic math facts and a reading measure as the correlation between these two measures would be lower than two measures of mathematics ability, regardless of whether the mathematics measures targeted the same skill (e.g., two measures of basic math facts) or varied, but seemingly related, skills (e.g., a basic facts measure and a word problems measure).

The guidelines of the MTMM approach are based on correlations among observed variables rather than correlations among latent constructs. Another criticism suggests the MTMM model is lacking in accounting for differences in measures and other methodologies, such as confirmatory factor analysis (CFA) better address these issues (Marsh, 1993). Latent variable modeling, including CFA, has been utilized to investigate constructs within the area of mathematics. Recently, research has provided evidence of two specific variables within tests of mathematics skills: (a) Computation and (b) Applications. Computation involves knowledge and skill in applying basic mathematics concepts of addition, subtraction, multiplication, and division to problems including single digit, multiple digits, decimals, and fractions (Fuchs et al., 1994; Jiban & Deno, 2007). Skills in simple computation are generally taught to be memorized for automaticity in solving more complex computation problems (i.e., multiple-digit problems); automaticity is the ability to perform a task without significant demands on attention or conscious thought (LaBerge & Samuels, 1974). Additionally, assessments of computation assess computation skills in isolation and do not focus on the application of computation skills to problem solving (Fuchs et al., 2008a). Computation, or M-CBM, probes are fluency-based and assess student skills in completing computation tasks sampled from local curriculum standards and were developed as a general measure of math achievement, not specifically as a measure of computation or applications (Thurber et al., 2002). Applications involves knowledge and skills in applying math concepts to word problem solving, number concepts, money, graphs and charts, measurement, and applied concepts (Fuchs et al., 2008a). Whereas computation requires understanding of math concepts, strategies, and facts (Howell, Fox, & Morehead, 1993), applications requires knowing how to apply the skills of computation to solving problems (Salvia & Ysseldyke, 1991).

The most prominent study examining the distinctions and relations between the constructs of computation and concepts and applications was Thurber et al. (2002). In this study, the authors suggested a two-factor model of mathematics with (a) Computation and (b) Applications as distinct but related constructs. Thurber and colleagues evaluated models of math skills utilizing CFA to determine constructs measured by traditional M-CBM measures within the context of other measures of mathematics. In this study, fourth-grade students were administered M-CBM probes sampled from the local curriculum, basic math fact probes, the Stanford Diagnostic Mathematics Test (SDMT; Beatty, Gardener, Madden, & Karlsen, 1985), the California Achievement Test (CAT; McGraw Hill, 1992), and items from the National Assessment of Educational Progress (NAEP). Correlations among the math measures were greater than .50, but some general patterns among them were evident. Measures traditionally considered to be assessments of Computation were more highly correlated with each other and had lower correlations with measures assessing Applications skills. Measures conceptualized as assessments of Applications skills exhibited the same pattern of correlating higher with each other and lower with Computation measures. M-CBM probes, traditionally considered to be assessments of general math ability, were most highly correlated with the skills tested by the basic math fact probes (displaying a median correlation of .82) and less highly with commercially produced measures of Computation (i.e., CAT and SDMT). M-CBM probes had lower correlations with measures of accuracy-based Applications skills such as SDMT Applications, CAT Applications, and the NAEP (displaying a median correlation of .44).

Thurber and colleagues (2002) conducted analyses to examine the fit of prespecified models of mathematics to the correlation matrix including several measures of Computation and Applications as part of their study. The first model was a single-factor model in which Computation and Applications form a general math factor that is measured accurately by M-CBM; this model was nested within the other models. The second model had a two-factor structure in which Computation and Applications were separate constructs and M-CBM was a measure of Computation. The third model also was a two-factor model with Computation and Applications as separate constructs, but it included M-CBM was a measure of Applications. Results of the factor analysis suggested that the single-factor model provided a poor to marginal fit to the data. In contrast, both two-factor models were significantly better fitting than the single-factor model. In the first two-factor model Computations and Applications were highly related (r = .83). Factor loadings on Computation ranged from .60 to .93, and factor loadings on Applications ranged from .89 to .90 for applications measures. In the second two-factor model Computations and Applications were again highly related (r = .88). Factor loadings on Computation ranged from .54 to .91, and factor loadings on Applications ranged from .38 to .90 for Applications measures. Comparison of the fit indices suggested that the first two-factor model, with M-CBM probes loading on the Computation factor, was the best-fitting of the models.

Confounding Influences

The work of Thurber and colleagues (2002) demonstrated computation and applications are distinct but related constructs in mathematics. Providing further support for the distinction between the two constructs, CBM measures are generally divided into probes assessing computation skills and probes assessing applications skills. Additionally, CBM probes are often a measure of fluency of skills. Important questions remain, however, regarding the construct validity of math fluency and accuracy probes and the degree to which other constructs may influence scores on these assessments and introduce construct-irrelevant variance, variance accounted for by factors unrelated to a particular construct, within their scores. The sections that follow will address the constructs of processing speed and reading as potential factors related to mathematics.

Processing Speed. Processing Speed is a well-replicated factor in the intelligence literature. Processing Speed represents individual differences in the general speed of cognitive performance; it involves the speed at which stimuli are identified and decisions are made (Carroll, 1993). Although brief, fluency-based assessments of mathematics, including CBM probes, are a valuable tool for assessing progress toward skills determined appropriate by an accepted curriculum, there is some concern about the construct-irrelevant influences of Processing Speed on their scores. For example, CBM mathematics probes are administered within a brief period of time—almost always less than 8 minutes and frequently less than 3 minutes. Therefore, one question worth answering is the degree to which results from fluency-based measures of mathematics (such as CBM probes) are comparable to results from assessments of math skills that are not administered with brief time limits.

Research has illustrated the relation between measures of fluency in performing academic tasks and measures of processing speed, suggesting that math fluency measures are influenced by factors other than those targeted by the measure (e.g., math computation skills; McGrew & Woodcock, 2001; Thurber et al., 2002). For example, CFA was used to determine the relation between select tests from the WJ III. Tests of processing speed from the WJ III, including Visual Matching, Decision Speed, and Pair Cancellation, were shown to have moderate to strong factor loadings on the Processing Speed factor (.71, .71, and .68, respectively). However, somewhat surprisingly, Math Fluency from the WJ III, which requires completing simple computations problems, was also shown to have a lower but moderate factor loadings on the same factor (.44; McGrew & Woodcock, 2001). Thus, there is evidence that mathematics fluency tasks may measure the same general speed-related construct.

As part of the study previously described, Thurber and colleagues (2002) considered method variance from fluency-based measures in their models because 8 of the 13 tests used in their study were fluency-based. For each of the three-factor models previously described, they investigated the effects of a Timed Tests factor on each of these measures. The best-fitting model included Computation, Applications, and Reading as first-order factors and included the Timed Tests factor as well. Within this model, tests of computation skill loaded on the Computation factor, tests of applications skills loaded on the Applications factor, and reading measures loaded on the Reading factor. All timed measures loaded on the Timed Tests factor as well as the content-related factor (i.e., Reading or Computation; no Applications measures were timed). Factor loadings for traditional M-CBM measures on the Timed Tests factor were strong (.70, .71, and .73). Factor loadings were moderate for Basic Facts measures (.54 and .51) and weak to moderate for reading CBM Maze measures (.34, .26, and .33). Reading maze measures are fluency-based measures that include a passage that typically has every seventh word removed; the word is replaced by a set of three words and the respondent chooses the word that best fits the sentence or passage. Factor loadings on the Timed Tests factor were higher for all math measures than for the reading measures. Within measures of mathematics computation skills factor loadings were higher for the M-CBM measures that included more complex (i.e., multi-digit) problems to be completed in a 5-minute period and were lower for Basic Facts measures that included only single-digit problems to be completed within a 2-minute time period. Results from the study conducted by Thurber and colleagues indicate that traditional measures of computation (M-CBM and Basic Facts measures) are highly correlated with speed, suggesting importance of automaticity in math computation knowledge.

Reading. Previous research has demonstrated significant relations between mathematics and reading abilities (Fuchs & Fuchs, 2003) and revealed that they, too, are distinct but related constructs (Betts et al., 2008; Thurber et al., 2002). Reading is a central aspect of contemporary mathematics curriculum and assessments with particular regard to applications (Jiban & Deno, 2007). Computation tasks involve memory for facts and counting strategies and require children to look at patterns of numbers in problem form, determine the operation required, and calculate the answer without reading words. Applications tasks often require reading for accurate understanding of the procedures required to complete presented problems. Therefore, it is logical that reading ability may interfere with the completion of mathematics problems. An important question is whether reading is a confounding factor when assessing mathematics applications skills. Many studies have discussed how reading ability may interfere with students' ability to complete math CBM probes measuring computation as well as concepts and applications (Betts et al., 2008; Jiban & Deno, 2007; Thurber et al., 2002).

Studies using both basic correlations and more complex statistical analyses, such as CFA, have examined these potential construct-irrelevant influences of reading on math test performance. For example, Jiban and Deno (2007) investigated the predictive validity of a reading maze measure as a predictor of performance on the Minnesota Comprehensive Assessment in Mathematics (MCA-Math). The MCA-Math was administered as an accuracy-based criterion variable and includes areas of math such as Shape, Space, and Measurement; Number Sense and Chance and Data; Problem Solving; and Procedures and Concepts. Whereas other measures administered as part of this study included simple mathematics computation problems, problems on the MCA-Math require a varied set of skills including the application of mathematics skills beyond simple operations, reading instructions to understand how to complete problems involving measurement, and reading information to determine how to solve a problem. Correlations were negligible to weak between reading maze measures and both computations measures (r = .08 for Basic Facts measures, r = .22 for the cloze math measure). In contrast, the correlation between the reading maze test and MCA-Math was moderate (r = .44). Results of this study indicate that reading and mathematics are related, but distinct constructs regardless of type of math skill (i.e., computation or applications). Results also indicate that reading is more highly related to mathematics problems involving applications, which typically require children to engage in some degree of reading to complete presented problems, and suggest that reading serves as a confound for applications measures.

In the Thurber and colleagues (2002) study previously described, the authors also included three reading CBM maze (Shinn, 2002) tasks to measure reading ability in their investigation of construct validity of mathematics measures of computation and applications. A Reading factor, specified to influence the three maze tasks, was included in each of the models tested and was specified to correlate with the Computation and Applications factors. The Reading factor was strongly correlated with both the Computation (r = .76) and Applications (r = .77) factors. Results of this study indicated less discriminant validity regarding measures of reading and both math constructs administered as part of the research than would be expected.

Betts and colleagues (2009) investigated the relations between numeracy and literacy skills in their study of the Minneapolis Kindergarten Assessment (MKA; Minneapolis Public Schools, 2004; Pickart, Betts, Sheran, & Heistad, 2005), an accuracy-based assessment battery for literacy and numeracy skills administered to students in kindergarten, using correlational analyses and CFA methods. Mathematics tests from the NALT were moderately correlated with early literacy tests (r = .34 to .53) as well as early numeracy tests from the MKA (r = .37 to .53). Reading tests from the NALT were moderately correlated with early literacy tests (r = .40 to .56) and early numeracy tests (r = .42 to .56) from the MKA. The median intercorrelations among literacy subtests was .53, and median intercorrelation among subtests assessing early numeracy skills was .55.

The median cross-correlation between literacy and numeracy measures was estimated to be .44, and this result provided weak evidence for divergence between the two constructs. Confirmatory factor analysis was utilized to evaluate the fit of the data to three models of mathematics ability. Results indicate that that best fit the data was a two-factor model, with literacy and numeracy as correlated factors and correlated residuals (r = .88). Several studies have investigated the relationship between reading and mathematics abilities by examining correlations between the constructs. However, no studies were located as part of this review that used CFA to examine the potential influences of reading comprehension on performance on math applications tests.

Current Study

The current study was designed to investigate construct validity of brief fluency-based assessments of mathematics ability with children who are in the early stages of skill acquisition in mathematics and reading. These assessments included measures from curriculum-based measurement systems and standardized, norm-referenced achievement assessment batteries. Thus, the study drew upon the work of Thurber and colleagues (2002) and investigated multiple models of mathematics skills and the constructs measured by these brief assessments.

Its methods extended beyond simple correlations, utilizing confirmatory factor analysis to test a series of prespecified models. Model 1 was a general performance factor model with all indicators of mathematics, reading, and processing speed would load on the general factor. Model 2 was a two-factor model that specified an academic general performance factor and included processing speed as a distinct factor. Model 3 was a three-factor model specifying processing speed and reading as distinct factors in addition to a general math factor. Model 4 was a four-factor model that specified processing speed, reading, computation, and concepts and applications as distinct factors.

Based on previous research, hypotheses were formed. Regarding relations among the constructs of computation, concepts and applications, processing speed, and reading, it was hypothesized that all constructs would be significantly correlated to some degree. However, it was also hypothesized that the mathematics constructs would be more highly correlated with each other than with the other constructs. Regarding the models tested, it was hypothesized that Model 4, the four-factor model, would be the best fitting model to the data based on previous research (Thurber et al., 2002) indicating two distinct factors involved in mathematics.

Confirmatory factor analysis also allowed for models to be constructed that tested the strength of the influence of potential confounds on the fluency-based and applications-based mathematics tests. In a series of models based on Model 4, the addition of a single path from the Processing Speed factor to the individual fluency-based tests allowed for testing the potential influence of processing speed ability on timed tests. Similarly, the addition of a single path from the Reading factor to each of the applications tests allowed for testing the potential influence of reading ability on math tests that include reading as a requirement to complete most problems. It was hypothesized that the fluency-based math measures would have a significant loading on the Processing Speed factor as fluency-based tests are administered under brief time constraints. Regarding the Reading factor, it was hypothesized that reading would have no or few effects on the applications tests because the tests administered, while including words on many items, focused more on mathematics skills than reading skills.

Method

Participants

Participants in this study were students in Grade 2 and Grade 3 at a university-based public elementary school. Students in 4 second grade classes and 3 third grade classes participated in data collection in May of 2010. Data from 69 Grade 2 and 54 Grade 3 students was used for the analysis (N = 123). Participants included 56 girls and 66 boys. Children ranged in age from 6 years, 7 months to 10 years, 2 months (*M* = 8 years, 7 months, *SD* = 7.88 months). No additional demographic information was obtained about participants as the passive consent procedure did not provide a means for obtaining this information from parents.

Measures

A total of 14 brief measures of mathematics, processing speed, and reading ability were administered. Eight of the mathematics measures were fluency-based measures, which were completed within a brief period of time. The remaining two mathematics measures were accuracy-based measures and were administered under generous time restrictions. Two processing speed measures were administered as measures of automaticity in completing cognitive tasks when measured under pressure to maintain attention. Two reading maze passages were administered to produce measures of reading ability. When grade-specific forms of measures were used, students in Grade 2 and Grade 3 completed Grade 3 measures as it was postulated that Grade 2 students would be able to respond to Grade 3 measures as data collection occurred at the end of the school year and some Grade 3 skills would have been developed at this time by Grade 2 students.

Math computation measures. AIMSWeb (Pearson, 2005) provides M-CBM probes based on expected grade-level skills for Grades 1 through 6 according to National Research Council (NRC; 2001) standards. A total of 40 alternate forms of these probes are available for use as benchmarking, strategic monitoring, progress monitoring, special education decision making, and program evaluation tools for assessing computation skills. M-CBM probes include two pages (front and back) of rows of math computation problems. Scoring of M-CBM items is based on number of digits correct. Skills assessed by Grade 3 probes include addition and subtraction of single- and multiple-digit numbers. Grade 3 M-CBM probes include 72 computation problems for a total of 169 points; students are allowed 2 minutes to answer as many problems as possible.

A study including students in Grade 5 general education classes reported 1-week test-retest reliability of .93 for measures similar to M-CBM (Tindal, Germann, & Deno, 1983). Tindal and colleagues (1983) also reported alternate-form reliability of .91 among across given to students in Grade 4 general education classes. Thurber et al. (2002) reported that M-CBM probes of mixed-operations items were typically correlated more highly with other measures of computation skills, including simple addition, subtraction, multiplication, and division problems as well as accuracy-based tests of computation skills included as part of standardized assessments, such as the Stanford Diagnostic Mathematics Test (SDMT; Beatty at al., 1985) and California Achievement Test (CAT; CTB/McGraw-Hill, 1992). Fluency-based M-CBM probes and fluency-based probes of basic math facts were highly correlated with a median correlation of .82. Mathematics CBM measures were moderately correlated with computation measures from the SDMT (median correlation of .58) and the CAT (median correlation of .62).

MBSP Computation (Fuchs et al., 1999) includes 30 probes, which measure grade-level skills, and these tests can be used for benchmarking and progress monitoring. MBSP Computation tests include 25 computation problems presented on one page. Scoring for this study was based on the number of correct items. Skills included with Grade 3 tests include addition and subtraction with regrouping, basic multiplication and division facts, and multiplication with regrouping. For Grade 3, a total of 50 points (based on number of digits correct) or 25 points (based on number of problems correct) is possible; students are allowed 3 minutes to complete as many problems as possible.

Reliability evidence for the MBSP Computation probes was obtained using two techniques. Two stability studies using one-week, alternate forms correlations were conducted. One study included students with high-incidence disabilities (e.g., learning disabilities and behavior disorders), and the other involved students with no identified disabilities. One-week, alternate-form reliability for Computation probes was .81 for Grade 3 students without disabilities (Fuchs et al., 1999). Another technique for evaluating the stability of the measurement involved correlating the average of students' first and third scores with the average of their second and fourth scores as suggested by Epstein (1979) and used in previous studies concerning measurement systems that rely primarily on aggregated estimates of student performance (Fuchs, Deno, & Marston, 1983). Correlations for aggregated odd and even scores were .81 for Grade 3 students without disabilities. Criterion validity of Computation test scores was assessed by correlating Computation scores of students with mild to moderate disabilities with scores from the Math Computation Test (MCT; Fuchs, Fuchs, Hamlett, & Stecker, 1991) and two accuracy-based subtests of the Stanford Achievement Test (SAT). Criterion validity coefficients between Computation and the MCT for Grade 3 were .81 when considering number of correct digits and .87 when considering number of correct problems. Criterion validity coefficients between Computation and the SAT Math Computation subtest were .55 for Grade 3.

The WIAT-III Math Fluency (Pearson, 2009) test includes three subtests, Addition, Subtraction, and Multiplication. Each subtest consists of 48 single-digit, basic fact problems; students are allowed 1 minute to complete as many problems as possible from each subtest. Problems are presented on two adjacent pages and are completed by answering problems across rows and down the first page before moving to problems on the second page (Breaux, 2009). The WIAT-III Math Fluency subtests are scored based on number of items correct and standard scores are obtained from age-based norms. At intervals ranging from 2 to 32 days, Addition has been shown to have test–retest reliability of .87 for grades 2 through 5, Subtraction has been shown to have test–retest reliability of .91 for grades 2 through 5, and Multiplication has been shown to have test–retest reliability of .90 for grades 3 through 5. The WIAT-III includes two additional, untimed, subtests measuring mathematics achievement, Math Problem Solving and Numerical Operations, and all Math Fluency subtests have been shown to correlate substantially with them. The Addition subtest was shown to have moderate correlations with both Math Problem Solving (r = .56) and Numerical Operations (r = .60). The Subtraction subtest was shown to have moderate correlations with both Math Problem Solving (r = .60) and Numerical Operations (r = .65) as well. The Multiplication subtest was also shown to have moderate correlations with both Math Problem Solving (r = .53) and Numerical Operations (r = .63; Breaux, 2009).

The WJ III ACH Calculation test (Woodcock, et al., 2001) measures the ability to perform math computation problems. Calculation includes 45 problems; it is presented on two consecutive pages in the WJ III ACH Response Booklet. Calculation is scored based on items corrects, and standard scores are obtained from age-based norms. Initial problems are simple problems with single number responses. The remaining problems are a mix of addition, subtraction, multiplication, and division problems; problems involving combinations of basic mathematical operations; and problems involving geometry, trigonometry, logarithmic operations, and calculus. Calculation problems include problems with negative numbers, percents, decimals, fractions, and whole numbers (Mather & Woodcock, 2001). Based on its standardization, Calculation is untimed; however, to reduce total administration time for this study, participants were given a time limit of 10 minutes to complete Calculation problems. It was assumed that 10 minutes will be sufficient for students to complete all problems involving skills they have learned. One year test–retest correlations for Calculation for all ages were high (r = .76 to .87). For ages 6 to 8, the Calculation test correlates moderately with the WJ III ACH Math Fluency test (r = .68) and with the WJ III ACH Applied Problems test (r = .49; McGrew, Schrank, & Woodcock, 2007).

The WJ III ACH Math Fluency test measures the ability to solve simple addition, subtraction, and multiplication facts quickly (Mather & Woodcock, 2001). This test consists of 160 problems presented on two consecutive pages. The initial 60 problems are simple, single-digit addition and subtraction problems, and the remaining 100 problems are a mixture of simple, single-digit addition, subtraction, and multiplication problems. Students are allowed 3 minutes to complete as many problems as possible.

Math Fluency is scored based on problems correct and standard scores are obtained from age-based norms. One year test–retest correlations for Math Fluency for ages 7 to 11were high (*r* = .94). For ages 6 to 8, the Math Fluency test correlates moderately with the accuracy-based WJ III ACH Applied Problems test (*r* = .50; McGrew et al., 2007).

Math applications measures. AIMSweb MCAP (Pearson, 2009) probes are short-duration assessments of mathematics problem-solving ability based on expected grade-level skills according to National Council for Teachers of Mathematics (NCTM, 2006) standards. For each grade level, alternate-form probes have been developed for use as benchmarking, strategic monitoring, frequent progress monitoring, special education decision making, and program evaluation tools for assessing mathematics problem-solving skills. To receive credit for a response on MCAP probes, the entire response must be correct; no credit is given for partial responses or individual correct digits. Skills included in MCAP probes include number sense, operations, patterns and relationships, measurement, geometry, and data and probability for Grade 3. MCAP probes are presented on three stapled pages for Grade 3, with the first two pages including problems on both sides of the page. Grade 3 MCAP probes include 29 problems; students are allowed 8 minutes to complete as many problems as possible. MCAP probes have been shown to have Cronbach's alpha of .80 for Grade 3 (Pearson, 2009).

MBSP Concepts and Applications (Fuchs et al., 1999) probes include 30 alternate-form probes that can be used for benchmarking and progress monitoring. Problems are presented on two attached pages containing problems on both sides of the first page. Points are awarded based on correct responses (Fuchs et al., 1999). Skills included with Grade 3 probes include counting, number concepts, names of numbers, measurement, money, charts and graphs, fractions, decimals, applied computation, and word problems. Grade 3 probes include 24 problems; students are allowed 6 minutes to complete as many problems as possible. Test-retest reliability for Concepts and Applications probes was observed to be .97 for Grade 3 (Fuchs et al., 1999). Concurrent validity coefficients were calculated between normal curve equivalent scores on the mean of students' final three Concepts and Applications scores and scores from the subtests of the CTB/McGraw Hill TerraNova (1997) annual assessment. Concurrent validity coefficients between Concepts and Applications probes scores and CTB/McGraw Hill Concepts and Applications scores were .64 for Grade 3.

Math probes from easyCBM (University of Oregon, 2009) are designed to align with NCTM (2006) standards and assess grade-level skills for kindergarten through Grade 8 (Alonzo & Tindal, 2009). For each grade level, 10 alternate-form probes containing 16 items have been developed. One Math Geometry probe (Grade 3) was used in this study as measures of accuracy in completing mathematics problems. Skills assessed with these probes include number sense, shapes, and measurement. Grade 3 probes include 16 problems. There is no specified time limit for easyCBM measures; however, in an effort to reduce the amount of time for administration while allowing students time to attempt all problems, students were given up to 10 minutes to complete easyCBM Math Geometry probes. In previous studies, researchers have given participants 15 minutes to complete easyCBM probes (J. Alonzo, personal communication, April 26, 2010). Because no standardized instructions are provided for easyCBM probes, instructions from the AIMSweb MCAP probes were used and altered slightly to align with the structure of easyCBM probes. No reliability or validity evidence has been reported for the Math Geometry test and Math Geometry Measurement and Algebra test.

Processing Speed measures. The WJ III COG Decision Speed test measures the ability to quickly select two pictures in a row that are most alike conceptually (Mather & Woodcock, 2001). Decision Speed has a time limit of 3 minutes. This test is presented in the WJ III COG Response Booklet and spans 4 pages, each with 10 rows, for a possible 40 points total. Prior to administration of the test, a practice exercise is given to ensure understanding of the directions for the test. One day test–retest reliability for Decision Speed was .80 for ages 7 to 11. For ages 6 to 8, a negligible correlation was reported for Decision Speed and WJ III COG Pair Cancellation tests (r = .14) and a moderate correlation was reported for Decision Speed and WJ III COG Visual Matching tests (r = .49; McGrew et al., 2007).

The WJ III COG Pair Cancellation test measures the ability to locate and mark a repeated pattern (Mather & Woodcock, 2001). Pair Cancellation has a time limit of 3 minutes and a total of 69 points are possible. Prior to administration of the test, a practice exercise is given to ensure understanding of the directions for the test. Typically presented as the final page of the WJ III COG Response Booklet, for this study, Pair Cancellation was administered as an individual page that was been removed from the Response Booklet. One day test–retest reliability for Pair Cancellation was .83 for ages 7 to 11. For ages 6 to 8, the Pair Cancellation test had a negligible correlation with the WJ III COG Visual Matching test (r = .19; McGrew et al., 2007).

Reading measures. AIMSweb Maze (Pearson, 2002) is a multiple-choice cloze test used to assess reading skills, and it is completed while students read silently. Within Maze passages, the first sentence of each 150 to 400 word passage is complete. Following the first sentence, every seventh word is replaced by a set of three words in parentheses from which the student chooses the one word that best completes the sentence. Of the two distracter words, one is a word that is the same type of word as the correct word (e.g., a noun, verb, or adjective). The other distracter is the same type of word as the correct word but does not make sense in the context of the sentence. Maze tasks are scored based on number of correct answers; students are allowed 3 minutes to complete as much of the Maze task as possible (Shinn & Shinn, 2002). Prior to completion of the first Maze measure, participants will complete the practice exercise included in the *Maze Administration and Scoring Manual* (Shinn & Shinn, 2002). The practice exercise consists of instruction regarding the design of the measure and how to complete each passage. Students complete three sentences under the lead researcher's guidance before being given instructions for completing the passage. Fuchs and Fuchs (1992) reported test-retest reliability estimates of .90 for maze probes similar to the one administered for this study. Reported validity estimates of maze probes ranged from .77 to .85 for students in Grades 3.

Procedure

Recruitment. Approval was granted by the University of Memphis Institutional Review Board and recruitment and parental consent were obtained using a passive consent procedure. Letters inviting students to participate in the study were distributed to parents twice to reduce the chance of the letter going unnoticed by parents. The letters requested that parents sign and return the form in the event that they do not want their child to participate. Prior to administration of the measures, assent from the students was obtained by providing the students with written and oral descriptions of the nature and procedures of the study.

Administration of measures. Measures were administered to students in their regular classrooms during two sessions lasting 60 to 90 minutes, spanning consecutive days. Students were introduced to the lead researcher and one to two research assistants. Research assistants were students in master's, educational specialist, or doctoral degree programs in psychology and undergraduate students in an honors program in psychology. All graduate research assistants completed at least one course in assessment and were familiar with assessment administration procedures.

Students in each class for whom permission to participate was not granted were provided with independent activities from the teacher. Students with permission to participate were given an overview of the study, assent was obtained from each student, and a folder containing the measures was given to each student. Students were introduced to a digital stopwatch, obtained from an online source (http://www.online-stopwatch.com) and displayed on a computer or projection screen. The purpose of the stopwatch was for students to be able to record the time at which they complete the timed and speeded measures. Students were informed that some of the activities they would complete were timed and that they would need to look at the display and record the time at which they completed the activity.

A total of 14 brief tests of mathematics, reading, and processing speed ability were administered. Within each grade level, the order of administration was counterbalanced across classrooms and days to reduce order effects. Each classroom within grade levels completed the tests in a different random order with five math tests, one processing speed test, and one reading test presented each day. Instructions were given for completing each test prior to beginning and steps were taken to ensure students were not given the opportunity to work on items before the timer was started. Instructions for group administration of AIMSweb and MBSP tests are included in the administration manuals for each AIMSweb test (Pearson, 2009; Shinn, 2005; Shinn & Shinn, 2002) and the *MBSP Basic Math Manual* (Fuchs, at al., 1999). Administration instructions for group administration of easyCBM tests were not found; therefore, instructions for AIMSweb M-CAP tests were modified slightly to coincide with easyCBM test design. Instructions for individual administration of tests from the WJ III ACH (Woodcock et al., 2001, 2007) and WJ III COG (Woodcock et al., 2001, 2007) are provided in the Test Book for each test battery, and they were modified slightly for group administration purposes. Instructions for the WIAT-III Math Fluency tests from the Record Form (Pearson, 2009) were modified slightly for group administration. Several tests included a practice exercise (WJ III COG Decision Speed and Pair Cancellation and AIMSweb Maze); for these tests, practice exercises were administered according to standardized instructions provided and instructions were modified slightly to fit a group administration, when necessary. Instructions for each test as well as general instructions can be found within the Example Script provided in Appendix C.

Following administration of instructions, the students were given the opportunity to ask questions. As students completed each test, the lead researcher and research assistant(s) walked around the room to ensure independent work, answer questions, and provide additional pencils when necessary. At the end of the allotted time for each test, students placed completed tests in the left pocket of their folder and retrieved the next test from the right pocket. Students wrote their name on a removable note page in their folder, which was removed and discarded by the students at the conclusion of the study (i.e., at the end of the second day) to maintain anonymity. After all tests were administered, students were thanked for their participation and completed a demographics form indicating their sex and date of birth.

Administration integrity and scoring accuracy. For each test, a research assistant trained in assessment administration completed an administration integrity checklist (see Appendix D) while observing the lead researcher to ensure standardized administration for all tests. Administration integrity checklists for each test were developed using the AIMSweb Accuracy of Test Administration Rating Scales (Shinn, 2005; Shinn & Shinn, 2002) as a model. Each administration integrity checklist included items for rating the accuracy of administration of standardized instructions, time limits, and basic administration procedures. Results indicated that administration of tests was completed with 98 percent accuracy.

Raw scores for all tests were calculated by the lead researcher (*N* = 118, 96%) and the faculty mentor (*N* = 6, 4%). Inter-scorer reliability and inter-scorer agreement was conducted by comparing scores calculated by the two primary scorers and a third scorer who was an advanced doctoral student in school psychology. Inter-scorer reliability was measured by the correlation between raw scores for each scorer; it represents relative agreement across scores. In contrast, inter-scorer agreement was measured by percentage agreement in producing the exact same raw score across scorers; it represents absolute agreement across scores. A total of 38 participants (30% of the total sample) were selected for this analysis using the random sample generator available through the Microsoft Excel program.

Results indicated that inter-scorer reliability values were above .98 and statistically significant ($p < .05$) for all tests except for one (i.e., easyCBM), and inter-scorer agreement values were all greater that 80% for all tests except for four tests (i.e., AIMSWeb M-CAP, WJ III COG Pair Cancellation, AIMSWeb Maze A, and AIMSWeb Maze B). Specific values are as follows: AIMSWeb M-CBM, r = .99, agreement = 89%; MBSP Computation, r = 1.00, agreement = 95%; WIAT-III Addition, r = .99, agreement = 92%; WIAT-III Subtraction, r = .98, agreement = 86%; WIAT-III Multiplication, r = .98, agreement = 81%; WJ III ACH Calculation, r = 1.00, agreement = 86%; WJ III ACH Math Fluency, r = 1.00, agreement = 80%; AIMSWeb M-CAP, r = .99, agreement = 78%; MBSP Applications r = 1.00, agreement = 89%; easyCBM, r = .87, agreement = 92%; WJ III COG Decision Speed, r = 1.00, agreement = 92%; WJ III COG Pair Cancellation, r = .98, agreement = 77%; AIMSWeb Maze A, r = 1.00, agreement = 76%; AIMSWeb Maze B, r = 1.00, agreement = 78%. Scoring inconsistencies were primarily due to incorrect summing of item scores. The low inter-scorer reliability value for easyCBM is attributed to one instance of a very large difference between scores recorded for one participant. All scoring errors made by the primary scorers identified during this analysis were corrected.

Analysis

Grade differences. Descriptive statistics for all test variables were calculated independently for each grade. A comparison of the means for each test was conducted to examine the differences in scores across grades and inform the decision to calculate z-scores for variables within grade levels for further analyses.

***A priori* model testing.** Data were submitted to first a correlational analysis and then a confirmatory factor analysis. First, the fit provided by *a priori* specified models was examined using the maximum-likelihood estimation of the Amos (Analysis of Moment Structures; Arbuckle, 1999), which accommodates data missing at random. The models to be tested reflected a variety of previously published models of mathematics components as well as theoretically viable alternatives. The fit of each model was examined on the following indices: (a) Chi-square (χ^2) goodness-of-fit test; (b) ratio of χ^2 to degrees of freedom; (c) the Tucker-Lewis index, (d) the comparative fit index, (e) the goodness-of-fit index, (f) the parsimony comparative fit index, (g) the root mean square error of approximation (RMSEA), and (h) Akaike's information criterion (AIC). The AIC was used to compare non-nested models. In addition, a test of the χ^2 difference between nested models was conducted to determine the appropriateness of releasing or imposing any restriction on the parameters.

Four prespecified models considering mathematics, processing speed, and reading were tested. Several variations on the fourth model were tested to investigate construct-irrelevant variance. The first model included a General Performance factor on which all indicators were loaded. It was assumed for this model that there were no distinct constructs aside from the general performance factor. Therefore, all tests served as indicators for the General Performance factor regardless of the skill they were assumed to assess (i.e., computation, applications, processing speed, reading).

In a second model, two first-order factors, an Academic Achievement factor and a Processing Speed factor, were included. Indicators that load on the Academic Achievement factor were those tests that assessed math and reading skills and included AIMSweb M-CBM, MBSP Computation, WIAT-III Math Fluency, WJ III ACH Math Fluency, WJ III ACH Calculation, AIMSweb MCAP, MBSP Applications, easyCBM Math Geometry, and AIMSweb Maze. Tests of processing speed were specified as indicators that load on the Processing Speed factor and included WJ III COG Decision Speed and WJ III COG Pair Cancellation.

A third model of this type retained the Processing Speed factor; however, tests of academic achievement were divided into tests of Math and Reading. Within this model, the first-order Math factor included tests of mathematics ability; indicators for this factor were AIMSweb M-CBM, MBSP Computation, WIAT-III Math Fluency, WJ III ACH Math Fluency, WJ III ACH Calculation, AIMSweb MCAP, MBSP Applications, and easyCBM Math Geometry. Indicators for the first-order Reading factor included tests of reading ability; indicators for this factor were AIMSweb Maze tests.

A fourth model retained the Processing Speed and Reading factors and included two first-order Math factors, Computation and Concepts and Applications. Tests that loaded on the Computation factor included item content that requires completion of basic mathematical operations. Indicators for the Computation factor were AIMSweb M-CBM, MBSP Computation, WIAT-III Math Fluency, WJ III ACH Math Fluency, and WJ III ACH Calculation. Tests including item content related to the application of mathematical concepts and knowledge were loaded on the Concepts and Applications factor. Indicators for Concepts and Applications included AIMSweb MCAP, MBSP Applications, and easyCBM Math Geometry.

Models testing effects of potential confounds were specified using the model that provided the best fit to the data as a foundation. Eleven variations of the best fitting model were tested to assess the potential confounds of Processing Speed and Reading. Seven models included the addition of one of the timed tests as an indicator for the Processing Speed factor in addition to the original factor for which the test served as an indicator. Additional indicators for the Processing Speed factor included AIMSweb M-CBM, MBSP Computation, WIAT-III Addition, WIAT-III Subtraction, WIAT-III Multiplication, WJ III ACH Math Fluency, AIMSWeb M-CAP, and MBSP Applications. Three models included the addition of one of the concepts and applications tests as an indicator for the Reading factor in addition to the Concepts and Applications factor. Indicators for the Concepts and Applications factor and Reading factor included AIMSWeb M-CAP, MBSP Applications, and easyCBM Math Geometry.

Results

Descriptive Statistics and Grade-Level Differences

Means, standard deviations, skew, and kurtosis values for all tests across both Grades 2 and 3 are reported in Table 1. Means for all norm-referenced standard scores for Grade 2 were higher than the expected mean of 100, and their standard deviations were larger than the expected standard deviation of 15 for all tests with the exception of WIAT-III Multiplication (*SD* = 12.41) and WJ III COG Pair Cancellation (*SD* = 7.71). For Grade 3 tests, means for WIAT-III Multiplication, WJ III ACH Calculation, WJ III COG Decision Speed, and WJ III COG Pair Cancellation were higher than the expected mean, but means for WIAT-III Addition, WIAT-III Subtraction, and WJ III ACH Math Fluency, were lower than the expected mean. Standard deviations for WJ III ACH Math Fluency and WJ III COG Decision Speed were larger than expected; standard deviations for WIAT-III Subtraction, WIAT-III Multiplication, and WJ III COG Pair Cancellation were lower than expected. Standard deviations for other tests were approximately 15. Skew and kurtosis were acceptable (< 2.0) for most tests across grade levels. Only the kurtosis values for WJ III ACH Calculation (3.30) for Grade 2 and for WJ III COG Decision Speed (3.31) for Grade 3 exceeded 2.0.

For tests that utilized raw scores for analysis, means and standard deviations could not be compared to expected values as these tests have not been normed on a standardized sample. Skew and kurtosis were acceptable (< 2.0) for most tests across grade levels. Only the kurtosis values for MBSP Computation (3.81) and AIMSWeb M-CAP (2.26) for Grade 2 exceeded 2.0. A comparison of the distributions of these raw scores across grade levels was conducted. Because there were large differences observed between the means across grade levels and four statistically significant differences between them ($p < .05$), z-scores were calculated by grade level for all tests yielding only raw scores. See the right three columns of Table 1 for the comparison of means. Table 2 provides descriptive statistics for the combined data set, which was used for all analyses. Prior to analyses, data screening procedures were conducted and results indicated assumptions regarding multivariate normality, linearity, and homogeneity of variance were not violated (Tabachnick & Fidell, 2007). The data were screened for univariate outliers, and two individual test scores were eliminated from the data set due to their extreme distance from the mean ($z \geq 4.0$). The data were screened for multivariate outliers using the Mahalanobis Distance method. Ten cases were identified as multivariate outliers; no patterns of outlying scores identified for individual tests or participants and no additional data was eliminated.

Table 1

Descriptive Statistics by Grade and Comparison of Means by Grade

	Grade 2 (*n* = 69)				Grade 3 (*n* = 54)				Grade comparisons		
	M	*SD*	Skew	Kurtosis	*M*	*SD*	Skew	Kurtosis	*M* diff	*t*	*P*
AIMSWeb M-CBM	20.31	9.33	1.18	1.22	22.35	9.74	0.35	0.83	-2.03	-1.16	.250
MBSP Computation PC	7.72	4.76	1.58	3.81	11.07	6.15	0.36	-0.83	-3.35	-3.41	.001
WIAT-III Addition	**102.03**	**15.22**	**-0.52**	**-0.09**	**93.48**	**15.80**	**0.08**	**0.70**	--	--	--
WIAT-III Subtraction	**106.65**	**16.21**	**-0.35**	**0.38**	**95.41**	**13.23**	**0.18**	**-0.55**	--	--	--
WIAT-III Multiplication	**107.21**	**12.41**	**0.08**	**0.53**	**101.46**	**12.66**	**0.45**	**1.09**	--	--	--
WJ III Calculation	**113.10**	**18.74**	**-0.61**	**3.30**	**102.94**	**15.50**	**-0.14**	**-0.25**	--	--	--
WJ III Math Fluency	**106.56**	**17.40**	**0.26**	**1.27**	**94.12**	**17.77**	**0.01**	**-0.68**	--	--	--
AIMSWeb M-CAP	12.88	6.62	1.34	2.26	14.35	7.38	0.46	-0.33	-1.47	-1.16	.250
MBSP Applications	22.94	9.22	0.30	0.49	23.92	10.84	-0.07	-0.65	-0.98	-0.54	.592
easyCBM	11.86	2.59	-0.24	-0.55	12.93	2.09	-0.81	0.01	-1.07	-2.48	.015
WJ III Decision Speed	**102.12**	**21.09**	**-0.44**	**0.79**	**101.21**	**23.56**	**-1.48**	**3.36**	--	--	--
WJ III Pair Cancellation	**100.39**	**7.71**	**-0.16**	**0.41**	**102.29**	**8.12**	**-0.26**	**0.26**	--	--	--
AIMSWeb Maze A	16.48	6.64	0.54	0.43	19.31	8.36	0.55	-0.25	-2.87	-2.12	.036
AIMSWeb Maze B	14.85	6.47	0.45	-0.37	19.88	9.26	0.38	-0.84	-5.03	-3.49	.001

Note. Scores in bold are norm-referenced standardized scores (*M* = 100, *SD* = 15). All other scores are raw scores.

Table 2
Descriptive Statistics for Combined Data Set (N = 123)

	M	*SD*	Skew	Kurtosis
AIMSWeb M-CBM	0.00	1.00	0.81	0.28
MBSP Computation PC	0.00	1.00	1.04	1.68
WIAT-III Addition	**98.21**	**16.00**	**-0.02**	**0.13**
WIAT-III Subtraction	**101.59**	**15.91**	**0.02**	**-0.22**
WIAT-III Multiplication	**104.64**	**12.79**	**0.22**	**0.48**
WJ III Calculation	**108.71**	**18.05**	**-0.28**	**1.71**
WJ III Math Fluency	**101.14**	**18.56**	**0.09**	**0.38**
AIMSWeb M-CAP	0.00	1.00	0.94	1.03
MBSP Applications	0.00	1.00	0.14	-0.04
easyCBM	12.33	2.43	-0.48	-0.35
WJ III Decision Speed	**101.72**	**22.12**	**-0.98**	**2.15**
WJ III Pair Cancellation	**101.21**	**7.91**	**-0.19**	**0.24**
AIMSWeb Maze A	0.00	1.00	0.54	0.08
AIMSWeb Maze B	0.00	1.00	0.41	-0.60

Note. Scores in bold are norm-referenced standardized scores (*M* = 100, S*D* = 15). All other scores are z scores.

Table 2 provides information regarding the mean, standard deviation, skew, and kurtosis for the combined sample. For norm-referenced standardized tests, mean score values were higher than the expected 100 for all tests except WIAT-III Addition (*M* = 98.21). Standard deviations for most norm-referenced standardized tests were larger than the expected value (*SD* = 16.00 to 22.12). Standard deviation values were below the expected value of 15 for WIAT-III Multiplication (*SD* = 12.79) and WJ III Pair Cancellation (*SD* = 7.91). The standard deviation for WIAT-III Subtraction (15.91) most closely approximated the expected value. Skew and kurtosis were acceptable (< 2.0) for most norm-referenced standardized tests across grade levels, except for kurtosis for WJ III Decision Speed (2.15). Skew and kurtosis values for all non norm-referenced standardized tests utilizing raw scores were within acceptable limits (< 2.0).

Correlational Analysis

Table 3 provides results of the correlational analysis of the tests using the combined data of Grades 2 and 3. The following general labels were used for interpreting the magnitude of these correlations: *negligible*, .00 to .19; *weak*, .20 to 39; *moderate*, .40 to .69; *strong*, .70 to .89; and *very strong*, .90 to 1.0. All test scores were significantly correlated at the .05 level, but the magnitude of the correlations ranged from strong to weak. Moderate to strong correlations were evident between all Computation tests (*rs* = .48 to .83). Among Computation tests, correlations were strong between WIAT-III Addition and WIAT-III Subtraction (*r* = .76) and WJ III ACH Math Fluency (*r* = .82); between WIAT-III Subtraction and WJ III ACH Math Fluency (*r* = .83); between WIAT-III Multiplication and WJ III ACH Calculation (*r* = .70) and WJ III ACH Math Fluency (*r* = .75), and between WJ III ACH Calculation and WJ III ACH Math Fluency (*r* = .75). Correlations between all other computation tests were moderate. Within Applications tests, a strong correlation was evident for AIMSWeb M-CAP and MBSP Applications (*r* = .78). Moderate correlations were evident for easyCBM and AIMSWeb M-CAP (*r* = .46) and MBSP Applications (*r* = .50). A strong correlation was evident for the two AIMSWeb Maze reading tests (*r* = .79). A moderate correlation was evident for the WJ III COG Decision Speed and WJ III COG Pair Cancellation processing speed tests (*r* = .41).

Table 3

Correlational Analysis of Tests

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. AIMSWeb M-CBM	--													
2. MBSP Computation	.66*	--												
3. WIAT-III Addition	.66*	.55*	--											
4. WIAT-III Subtraction	.59*	.55*	.76*	--										
5. WIAT-III Multiplication	.62*	.61*	.69*	.69*	--									
6. WJ III ACH Calculation	.48*	.54*	.61*	.69*	.70*	--								
7. WJ III ACH Math Fluency	.67*	.66*	.82*	.83*	.72*	.75*	--							
8. AIMSWeb M-CAP	.60*	.70*	.51*	.56*	.47*	.52*	.58*	--						
9. MBSP Applications	.59*	.69*	.59*	.60*	.54*	.58*	.66*	.78*	--					
10. easyCBM	.31*	.36*	.26*	.37*	.35*	.53*	.34*	.46*	.50*	--				
11. WJ III COG Decision Speed	.43*	.39*	.39*	.44*	.35*	.31*	.49*	.36*	.44*	.20**	--			
12. WJ III COG Pair Cancellation	.29*	.39*	.36*	.38*	.41*	.32*	.34*	.39*	.43*	.32*	.41*	--		
13. AIMSWeb Maze A	.43*	.47*	.38*	.52*	.39*	.40*	.50*	.57*	.62*	.46*	.37*	.37*	--	
14. AIMSWeb Maze B	.50*	.56*	.43*	.49*	.45*	.47*	.57*	.64*	.71*	.48*	.37*	.33*	.79*	--

Note. * p <.01. ** p <.05.

Moderate correlations were evident across most (a) Computation and (b) Applications tests (*rs* = .47 to .69). A strong correlation was evident for AIMSWeb M-CBM and MBSP Computation (*r* = .78) and weak correlations were evident for easyCBM and AIMSWeb M-CBM, MBSP Computation, WIAT-III Addition, WIAT-III Subtraction, WIAT-III Multiplication, and WJ III ACH Math Fluency (*rs* = .26 to .37). Correlations for Processing Speed tests and all other tests were weak to moderate (*rs* = .20 to .49). Moderate correlations were evident across (a) Computation and Applications tests scores and (b) Reading test scores (*rs* = .43 to .64), but there were weak correlations between AIMSWeb Maze A and WIAT-III Addition (*r* = .38) and WIAT-III Multiplication (*r* = .39). Correlations between Reading tests and two of the Applications tests (AIMSWeb M-CAP and MBSP Applications) tended to be more highly correlated than Reading tests and Computation tests.

Confirmatory Factor Analysis

***A priori* models.** Confirmatory factor analysis was conducted to compare different theoretical models to the patterns of correlations between each pair of test scores. The models included a range of factors, from one factor to four factors including two different math factors (Computation and Applications), a Processing Speed factor, and a Reading factor. Table 4 provides the results of the model fit comparisons between the single-factor, two-factor, three-factor, and four-factor models.

Table 4
Fit Statistics for Models

Model	χ^2	*df*	χ^2 Diff Comparison	$\Delta\chi^2$ / Δdf	CFI	TLI	RMSEA (90%CI)	AIC
1	316.00	77	-	-	.80	.72	.16 (.14/.18)	400.02
2	310.23	76	Model 2 to Model 1	-5.77/-1**	.80	.72	.16 (.14/.18)	396.23
3	245.05	74	Model 3 to Model 2	-65.18/-2*	.86	.79	.14 (.12/.16)	335.05
4	154.90	71	Model 4 to Model 3	-90.15/-3*	.93	.89	.10 (.08/.12)	250.90

Note. * p < .01. **p < 05.

The first model (Model 1) included a single General Performance factor. The single-factor model did not provide a particularly good fit to the data (see Table 4). Standardized path coefficients from the factor to the test scores were all statistically significant ($p < .05$), and they ranged from .48 to.90. Factor coefficients for all Computation tests, all Reading tests, and two of the Applications tests (AIMSWeb M-CAP and MBSP Applications) were above .60. Factor coefficients for Processing Speed tests (WJ III COG Decision Speed and WJ III COG Pair Cancellation) were .50 and .46, respectively, and the factor coefficient for easyCBM was.48.

For the two-factor model (Model 2), the Computation, Applications, and Reading tests were specified as indicators of an Academic factor. Processing Speed tests were specified to indicators of a Processing Speed factor. The Academic and Processing Speed factors were specified to be correlated. The two-factor Model 2 provided a significantly better fit to the data than the single-factor Model 1, $\Delta\chi_2 = 5.77$, $\Delta df = 1$, p <.05. Standardized path coefficients were all statistically significant ($p < .05$), and the two factors were significantly correlated, $r = .75$.

For the three-factor model (Model 3), Computation and Applications tests were specified as indicators of the Math factor, whereas Reading tests were specified as indicators of the Reading factor. Thus, indicators of the Academic factor from Model 2 were respecified as indicators of Math or Reading; the Processing Speed factor indicators remained unchanged from Model 2. All three factors were specified to be correlated. The three-factor Model 3 provided a significantly better fit to the data than the two-factor Model 2, $\Delta\chi_2 = 65.18$, $\Delta df = 2$, $p < .001$. Standardized path coefficients were all statistically significant ($p < .05$), and the three factors were significantly correlated, Math and Reading $r = .68$, Math and Processing Speed $r = .73$, and Reading and Processing Speed $r = .59$.

The four-factor model (Model 4) included further separation of academic-related factors with the separation of the Math factor from Model 3 into two separate mathematics skill areas: Computation and Application. Reading and Processing Speed were also included in Model 4. The Computation factor was formed from tests assessing basic mathematics calculation abilities: AIMSWeb M-CBM, MBSP Computation, WIAT-III Addition, WIAT-III Subtraction, WIAT-III Multiplication, WJ III ACH Calculation, and WJ III ACH Math Fluency. The Applications factor was formed from tests assessing abilities regarding applying basic mathematics concepts to more complex problems: AIMSWeb M-CAP, MBSP Applications, and easyCBM. Reading and Processing Speed indicators remained the same as in Model 3. All factors were specified to be correlated with one another. Model 4 proved be a significantly better fit to the data than Model 3, $\Delta\chi_2 = 90.15$, $\Delta df = 3$, $p < .001$, and provided the best fit of all the models (see Table 4). Standardized path coefficients were all statistically significant ($p < .001$)

and the four factors were significantly correlated. These factor correlations ranged from .59 to .81, as evident in Figure 1. Results of the CFA indicated that the addition of Computation and Applications factors accounting for separate, yet related, categories of math skills significantly improved the fit of the model as compared to previous models.

Construct-irrelevant variance models. In order to investigate construct-irrelevant variance contributed by Processing Speed and Reading to scores on mathematics tests, 11 variations of the best fitting *a priori* model, Model 4, were specified with each variation including one additional path from either the Processing Speed factor or the Reading factor to a mathematics test. Model 4a through Model 4h tested the effects of Processing Speed on performance of each of the eight speeded Computation or Applications tests. These models included a path to a speeded math test from the Processing Speed factor in addition to the preexisting path from either the Computation factor or Applications factor; thus, the influence of the Computation or Applications factor was controlled for in each model.

Model 4i through Model 4k tested the effects of Reading on performance of each of the three Applications tests. These models included a path to each of the Applications tests from the Reading factor in addition to the preexisting path from the Applications factor; thus, the influence of the Applications factor was controlled for in each model. As evident in Table 5, results indicated that none of these three models provided a significantly better fit to the data than Model 4. In addition, the standardized regression paths from the Reading factor to each of the tests were nonsignificant and either negative or low in magnitude (range = -.12 to .08).

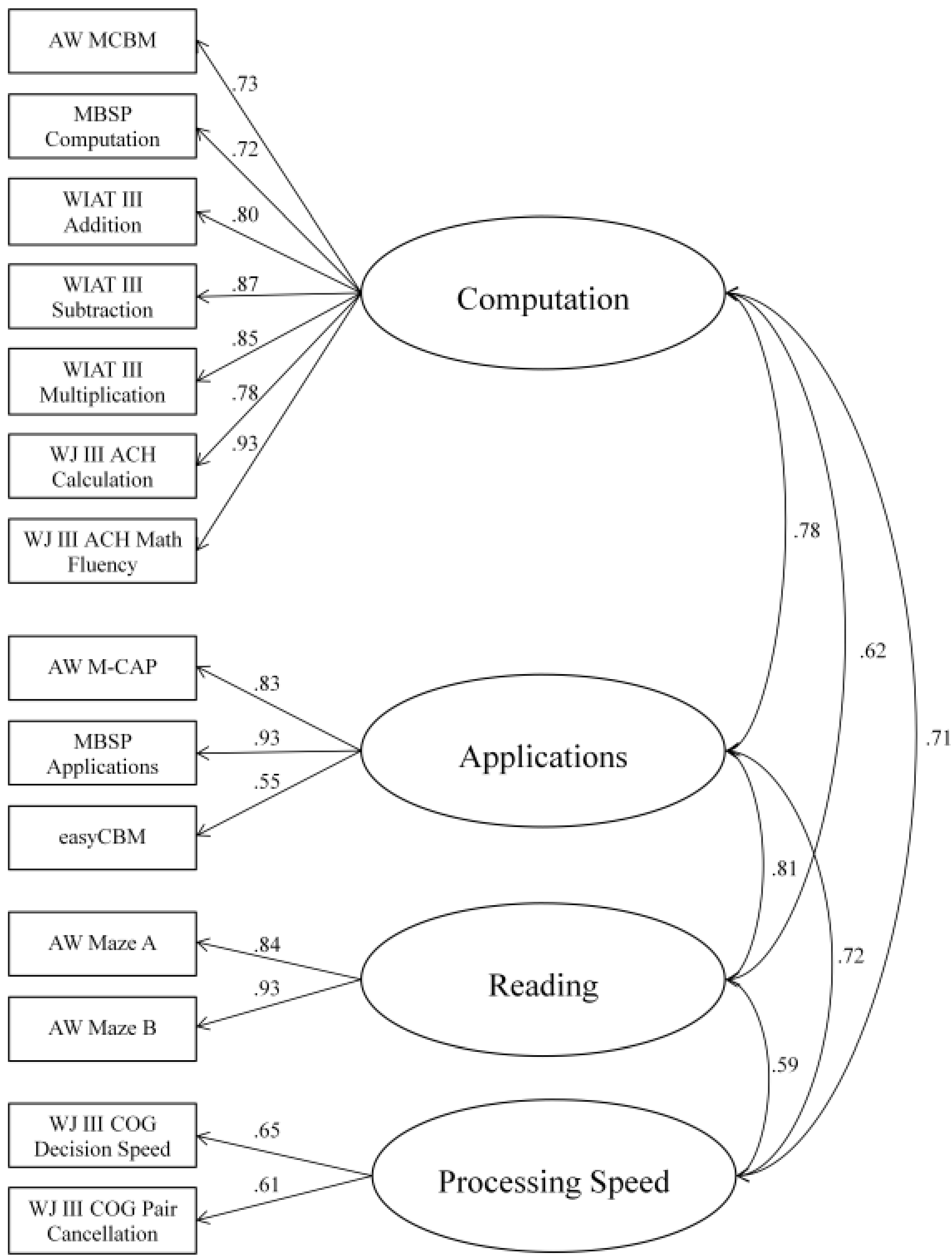

Figure 1. Four-Factor Model (Model 4)

Table 5

Fit Statistics for Model 4 Variations to Test Construct-Irrelevant Variance

	χ^2	*df*	$\Delta\chi^2$ / Δdf	CFI	TLI	RMSEA (90% CI)	AIC
Model 4	154.90	71		.93	.89	.10 (.08-.12)	250.90
4a. AIMSWeb M-CBM loading on Processing Speed	152.23	70	-1.67/1	.93	.89	.10 (.08-.12)	251.23
4b. MBSP Computation loading on Processing Speed	136.10	70	-18.80*/1	.94	.92	.09 (.07-.11)	234.05
4c. WIAT-III Addition loading on Processing Speed	153.30	70	-1.60/1	.93	.89	.10 (.08-.12)	251.30
4d. WIAT-III Subtraction loading on Processing Speed	154.81	70	-0.09/1	.93	.89	.10 (.08-.12)	252.81
4e. WIAT-III Multiplication loading on Processing Speed	154.86	70	-0.04/1	.93	.89	.10 (.08-.12)	252.86
4f. WJ III ACH Math Fluency loading on Processing Speed	153.93	70	-0.97/1	.93	.89	.10 (.08-.12)	251.93
4g. AIMSWeb M-CAP loading on Processing Speed	154.79	70	-0.11/1	.93	.89	.10 (.08-.12)	252.79
4h. MBSP Applications loading on Processing Speed	154.74	70	-0.16/1	.93	.89	.10 (.08-.12)	252.74
4i. AIMSWeb MCAP loading on Reading	154.40	70	-0.50/1	.93	.89	.10 (.08-.12)	252.40
4j. MBSP Applications loading on Reading	152.79	70	-0.16/1	.93	.89	.10 (.08-.12)	252.79
4k. easyCBM loading on Reading	152.07	70	-2.83/1	.93	.90	.10 (.08-.12)	250.07

Note. All comparisons are to Model 4. * p < .01.

Results indicated that most of these models (excluding Model 4b) did not provide a significantly better fit to the data than Model 4. Standardized regression paths from the Processing Speed factor to each of the tests included in these models were nonsignificant and either negative or low in magnitude (range = -.20 to .19). However, Model 4b proved to be a significantly better fit to the data than Model 4, $\Delta\chi_2 = 18.80$, $\Delta df = 1$. However, within Model 4b, the standardized regression path for MBSP Computation as an indicator of Computation was not statistically significant ($p = .77$). The standardized regression path for MBSP Computation as an indicator for Processing Speed was statistically significant (.89, $p = .01$).

Discussion

There is evidence to suggest that, despite national educational reform and specific attention given to the subject area, a national deficit exists in the development of mathematics skills across grades, ability levels, and sociodemographic variables. Emphasis on regular skill assessment, intervention, and progress monitoring under the RTI model has created a need for the development of assessment instruments that are psychometrically sound, reliable, universal, and brief. Important factors to consider when developing or selecting assessments for the school environment include what skills are assessed and what additional factors may potentially influence performance on such tests. The current study investigated construct validity of established, widely-used CBM and standardized, norm-referenced tests of mathematics as well as the potential confounding influence of reading and processing speed abilities. Additionally, construct validity of the measures administered was assessed, through an investigation of convergent and discriminant validity, using CFA. Numerous prespecified, theoretical models were tested to replicate previous studies suggesting specific models of mathematics ability (convergent validity) and to identify construct-irrelevant variance (discriminant validity) imposed on tests of computation and applications by processing speed and reading.

Distinct Math Factors

Previous research has resulted in the development of theoretical conceptualizations of the constructs involved in mathematics ability. More recent research suggests two specific constructs within mathematics: Computation and Applications (Thurber et al., 2002). Computation involves the completion of basic mathematical operations and involves no reading requirements. Applications often involves reading instructions and information to successfully complete a problem. In addition to previous research suggesting these to be separate, but related constructs, most mathematics curriculums and assessments focus on basic mathematical operations as well as the application of basic mathematical operations and concepts to more complex problems. The results of previous research indicate that Computation and Applications, while requiring different knowledge and application skills, are related. In a previous study using tests similar to the ones used in the current study (Thurber et al., 2002), the correlation between Computations and Applications factors was strong and statistically significant (r = .83). The current study revealed a correlation between Computation and Applications that was remarkably similar and was only slightly lower (r = .78) than that observed by Thurber et al. (2002).

For the current study, models of mathematics were developed, drawing from previous work including the work of Thurber and colleagues (2002), as a means to investigate the structure of mathematics skills. It was hypothesized that a four-factor model specifying these independent mathematics constructs, Computation and Applications, as well as Reading and Processing Speed as factors would provide the best fit to the data. This hypothesis was confirmed. In fact, comparison of the three-factor model (Model 3) with the four-factor model (Model 4) provided the best test for the concept of two distinct constructs within mathematics ability. The three-factor model included a single general Math factor, Reading factor, and Processing Speed factor; the four-factor model retained Reading and Processing Speed as factors and divided the general Math factor from Model 3 into two separate factors of Computations and

Applications. Results suggested that Model 4 provided a better fit over Model 3, providing additional evidence to support Computation and Applications as two distinct, yet related constructs of mathematics ability.

Convergent and Discriminant Relations

As with many academic skill areas, mathematics skills are not distinct sets of skills with no relation to other skill areas. It was hypothesized that all constructs would be significantly correlated with each other to some degree and related constructs would correlate more highly with each other (i.e., Computation and Applications) than with more related constructs (i.e., Reading and Processing Speed). Results of the CFA supported the hypothesis that all constructs would be significantly correlated with each other indicating moderate to strong correlations (r = .59 to .81). Additionally, results suggested the two mathematics-related constructs (Computation and Applications) were slightly more highly correlated (in most cases) with each other (r = .78) than with unrelated constructs; however, an exception was the correlation between Applications and Reading. Computation was moderately correlated with Reading (r = .62). Regarding Reading, it was expected that Applications tests would be more highly correlated with Reading tests than Computations tests would be, as Applications tests are typically language-loaded while Computation tests have only numerical operations problems. Results indicated that not only were Applications and Reading more highly correlated than Computation and Reading, but Applications was more highly correlated with Reading (r = .81) than with Computation. Computation was strongly correlated with Processing Speed (r = .71), as was Applications (r = .72). Correlations between mathematics factors and the Reading factor exhibited the expected pattern of

Applications and Reading having a higher correlation than Computation and Reading. The higher correlation between Applications and Reading than the correlation between mathematics factors was not expected. Correlations between math factors and the Processing Speed factor was high and almost as high as the correlation between the math factors; this was unexpected as the processing speed tests were chosen specifically because they did not include numbers in the tasks.

Influence of Confounds

While results of previous research and the current study indicate that mathematics is composed of two separate math-related constructs, it is logical that Computation and Applications are not the only factors involved in mathematics. Most CBM tests require the completion of as many problems as possible within a brief period of time, making them fluency-based assessments of skills. Applications problems often require reading, introducing a factor unrelated to mathematics. It is important to consider the potential influence of factors that are not specifically mathematical, how these unrelated factors affect the construct validity of math fluency and accuracy tests, and to what degree do these factors introduce construct-irrelevant variance to scores on mathematics tests. The degree to which Reading and Processing Speed contribute to individual differences in performance on specific tests was investigated through testing several variations of Model 4 and comparing each new model to Model 4 to assess better fit.

Most CBM tests and fluency-based tests from standardized-norm-referenced assessments are administered under brief time constraints and may require greater speed of processing to complete more items correctly than those tests administered under more generous time constraints. For example, students may be given a set of 40 single-digit addition and subtraction problems and are asked to complete as many problems as possible in 3 minutes. It is logical that processing speed ability would contribute individual differences in performance on fluency (or speeded) tests. Fluency measures reflect the speed and accuracy with which an individual can attend to a stimulus, make a decision, and respond (VanDerHeyden & Burns, 2008). Jiban and Deno (2007) suggest that knowledge of basic math facts is used more for “mental computation and estimation” and leads to automaticity with basic computation skills. With rehearsal, skills may reach a level of proficiency in which response is quick, accurate, and requires little to no monitoring (Gersten, Jordan, & Flojo, 2005). Additionally, automaticity may support the ability to succeed at higher-level tasks involving applications. Of the 10 mathematics tests administered as part of this study, only two were administered without strict time constraints. As expected, Processing Speed tests typically showed higher correlations with fluency–based tests ($r = .29$ to 49) than with accuracy tests ($r = .20$ to .32). Regarding the CFA, it is notable that WJ III ACH Calculation, which is not fluency-based, yielded a path coefficient that was similar to fluency-based tests of computation. Therefore, it appears that this factor measures individual differences in math skill development in computation versus only fluency-based skills in computation.

Results of testing the influence of the Processing Speed factor on the eight fluency tests revealed that none of the specified paths were sizeable or statistically significant, with one exception. When a path was added from Processing Speed to MBSP Computation, the model provided a significantly better fit to the data than the original four-factor model ($\Delta\chi_2 = 18.80$, $\Delta df = 1$). Within this model, the standardized regression path from Computation to MBSP Computation was no longer significant, whereas the path from Processing Speed to MBSP Computation was significant. Items on the MBSP Computation test include multiplication and division problems, as well as addition and subtraction problems, and the MBSP Computation test was scored based on the number of total problems correct (as opposed to number of individual digits correct). It is possible that participants were able to respond correctly to more problems at a faster speed due to the nature of the problems on the MBSP Computation test (e.g., fewer digits included in problems, fewer steps involved in the problems when compared to other Computation tests), resulting in scores being influenced more by Processing Speed than Computation skills. Overall, these results indicate that any influence exerted by processing speed ability on these tests was negligible when controlling for other skills (i.e., computation and applications skills). Individual differences in processing speed do not appear to affect performance on the fluency tests included in this study, with the exception of MBSP Computation.

Reading is required on most Applications test problems, as Applications problems require applying knowledge of basic mathematical operations to more complex problems that often include narrative instructions or information. For example, children must read a sentence or set of sentences such as, “Jonathan buys 3 cupcakes for $1.00 each. He gives the cashier $5.00. How much change does Jonathan get back?” and then calculate the answer required by the problem. It is logical therefore, that reading ability would contribute individual differences in performance on Applications tests. As expected, in the best fitting model, the Reading factor was more highly correlated with the Applications factor (r = .81) than it was with Computation or Processing Speed factors (r = .62, .59, respectively). Furthermore, correlations between Reading tests and

Applications tests (with the exception of easyCBM) were moderate to strong and ranged from .57 to .71 for AIMSWeb M-CAP and MBSP Applications, and review of the easyCBM items revealed that they include fewer words than the other Applications test and, therefore, may not be as susceptible to the influence of reading ability.

Results of testing the influence of the Reading factor on the three Applications tests revealed that none of the specified paths were sizeable or statistically significant. These findings indicate that any influence exerted by reading ability on these tests was negligible when controlling for applications skills. Individual differences in reading do not appear to affect performance on any of the application tests included in this study.

Limitations and Future Research

This study has several limitations associated with measurement and sampling including sample characteristics, sample size, and scoring procedures.

Sampling limitations. One limitation to the current study is the sample size of 123. Increasing the number of participants would likely yield results more indicative of the larger population. It is generally agreed that larger sample sizes are preferable when conducting any form of research. In structural equation modeling (SEM), sample size has implications for standard error estimates, chi-square estimates, and various other test statistics (Kline, 2011). Previous research has suggested the ideal minimum sample size can be determined by the ratio of cases to the number of model parameters requiring statistical estimates. An ideal ration of cases to parameters would be 20:1 (Jackson, 2003). Other studies surveyed research utilizing SEM and reported a “typical” sample size of around 200 cases (Breckler, 1990; Shah & Goldstein, 2006). It is noted that 200 may still be too small a sample when analyzing models of greater complexity. Samples of less than 100 cases are almost always too small unless the models tested are very simple (Kline, 2011). Future research should attempt to obtain larger sample sizes by recruiting more classrooms or schools to participate in a study similar to the current study.

Another limitation to the current study associated with sampling is the use of a university-based public school to recruit participants. Typically, schools that are sponsored by institutions of higher learning are seen as holding higher standards for education and development, thereby fostering children's academic skills that exceed that of the general population. Such schools may have access to more cutting-edge curriculums, instructional methods, and technology. As a result, these schools may produce children with academic skills that are more advanced than typical students in more traditional public school settings. Consistent with that expectation, results suggest the participants in this study, in general, displayed higher than average performance on the standardized, norm-referenced tests than would be expected in the general population. All test means for this grade level were above 100 (SS = 100.39 to 113.10). For Grade 3 students, mean standard scores were above the expected 100 for all but three tests (SS = 101.21 to 102.94). Somewhat surprisingly, both grade-based samples showed frequent expansion of range and only occasional restriction of range. For Grade 2 students, standard deviations were greater than the expected value of 15 for all tests with the exception of two. For Grade 3 students, standard deviations were also all above the expected value of 15 for all tests with the exception of three. Generalization of these results to the wider population should be conducted with caution. Future research should attempt to reduce greater than expected scores on standardized tests by including varying student populations.

Measurement limitations. One of the measurement limitations of the current study is the use of group administration for tests that were not developed with the intention of being administered to a group, as opposed to administered to individuals. All standardized, norm-referenced tests included in this study were developed as individually-administered tests, and standardized administration procedures were developed based on this type of administration. In order to expedite data collection, these tests were administered to whole classrooms instead of being administered to individual students. Administration procedures and instructions were modified as little as possible to reflect a group administration. As these tests do not have standardized scores for group administration, some degree of caution must be used when interpreting the scores as standardized scores provided for these tests by the publishers do not reflect group administration-specific variables.

A second measurement limitation relates to inter-scorer reliability. To evaluate reliability of scoring a sample of participant tests were selected to be scored by an independent second scorer. Inter-scorer reliability results indicate significant correlations between scores recorded by both scorers for all tests. However, inter-scorer agreement ranged from 77% to 92% when looking at agreement between scores recorded by both scorers. Differences in scores recorded were generally due to raw score summation errors. The results suggest that error may remain in scores recorded for remaining participants. Similar studies should consider scoring all tests twice in an effort to reduce scoring error.

Implications for Practice. Overall, the research provides evidence to support the conceptualization of mathematics ability as two distinct but related constructs: Computation (basic mathematics operations) and Applications (the application of basic math operations and concepts to more complex problems). The results also suggest that, while reading and processing speed do appear to have some role in completing mathematics problems, little to no influence is exerted on tests of mathematics skills by reading and processing speed abilities. However, because no significant influence was observed in the current study does not suggest that these factors cannot influence mathematics performance. For example, if a child is not able to read he will not be able to do well on applications tasks. The current study included a sample of typical children and included both good readers as well as poor readers. Continued research is needed in the area of mathematics to further develop understanding of mathematics-based assessments, particularly those CBM tests used in the brief testing of students to determine skill development and progress in the curriculum. With the national movement towards a Common Core curriculum, it is becoming more important to understand the reliability and validity of various mathematics assessments, identify the extent to which these assessments generalize across varying state curriculums and requirements, and understand what additional factors have the potential to influence performance on brief assessments of mathematics skill.

REFERENCES

Aiken, L. R. (1998). *Tests and examinations: Measuring abilities and performance*. New York, NY: Wiley.

Arbuckle, J. L. (1999). *Amos 4.0* [Computer software]. Chicago, IL: SmallWaters.

Beatty, L. S., Gardener, E. G., Madden, R., & Karlsen, B. (1995). *The Stanford Diagnostic Mathematics Test* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Betts, J., Pickart, M., & Heistad, D. (2009). Construct and predictive validity evidence for curriculum-based measures of early literacy and numeracy skills in kindergarten. *Journal of Psychoeducational Assessment*, *27*, 83-95.

Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *Behavior Analyst*, *19*, 163-197.

Breaux, K. C. (2009). WIAT-III Technical Manual. *Wechsler Individual Achievement Test-Third Edition*. San Antonio, TX: Pearson.

Breckler, S. J. (1990). Applications of covariance structure modeling in psychology: Cause for concern? *Psychological Bulletin*, *107*, 260-273.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the miltitrait-multimethod matrix. *Psychological Bulletin*, *56*, 81-105.

Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review*, *33*, 234-248.

Connell, J. E. (2005). *Constructing a math applications, curriculum-based assessment: An analysis of the relationship between applications problems, computation problems, and criterion-referenced assessments.* Unpublished doctoral dissertation, Louisiana State University Agricultural and Mechanical College.

CTB/McGraw Hill (1992). *California Achievement Tests* (5th Edition). Monterey, CA: CTB Macmillan/McGraw-Hill.

CTB/McGraw Hill (1997). *TerraNova: Technical manual*. Monterey, CA: Author.

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

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, *49*, 36-47.

Epstein, S. (1979). The stability of behavior: On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, *37*, 1097-1126.

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

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

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

Fuchs, L. S., Deno, S. L., & Marston, D. (1983). Improving the reliability of curriculum-based measures of academic skills for psychoeducational decision making. *Diagnostique*, *8*, 134-149.

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

Fuchs, L. S., Fuchs, D., & Courey, S. J. (2005). Curriculum-based measurement of mathematics competence: From computation to concepts and applications of real-life problem solving. *Assessment for Effective Intervention*, *30*, 33-46.

Fuchs, L., Fuchs, D., Hamlett, C. L., & Stecker, P. M. (1991). Effects of curriculum-based measurement and consultation on teacher planning and student achievement in mathematics operations. *American Educational Research Journal*, *28*, 617-641.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Thompson, A., Roberts, P. H., Kubek, P., & Stecker, P. M. (1994). Technical features of a mathematics concepts and applications curriculum-based measurement system. *Diagnostique, 19*, 23-49.

Fuchs, L. S., Fuchs, D., & Zumeta, R. O. (2008a). A curricular-sampling approach to progress monitoring: mathematics concepts and applications. *Assessment for Effective Intervention, 33*, 225-233.

Fuchs, L. S., Fuchs, D., & Zumeta, R. O. (2008b). Response to intervention: A strategy for the prevention and identification of learning disabilities. In E. L. Grigorenko (Ed.), *Educating individuals with disabilities: IDEIA 2004 and beyond*. New York, NY: Springer.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1999). *Monitoring Basic Skills Progress*. Austin, TX: Pro-Ed.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*, 293-304.

Hintze, J. M., Christ, T. J., & Keller, L. A. (2002). The generalizability of CBM survey-level mathematics assessments: Just how many samples do we need? *School Psychology Review*, *31*, 514-528.

Howell, K. W., Fox, S. L., & Morehead, M. K. (1993). *Curriculum-based evaluation: Teaching and decision making*. Pacific Grove, CA: Brooks/Cole.

Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C. § 1400 et seq. (2004). (Reauthorization of the Individuals with Disabilities Education Act of 1990)

Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the *N*:q hypothesis. *Structural Equation Modeling*, *10*, 128-141.

Jiban, C. L., & Deno, S. L. (2007). Using math and reading curriculum-based measurements to predict state mathematics test performance: Are simple one-minute measures technically adequate? *Assessment for Effective Intervention*, *32*, 78-89.

Kelley, B., Hosp, J. L., & Howell, K. W. (2008). Curriculum-based evaluation and math: An overview. *Assessment for Effective Intervention*, *33*, 250-256.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.

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

Leh, J. M., Jitendra, A. K., Caskie, G. I. L., & Griffin, C. C. (2007). An evaluation of curriculum-based measurement of mathematics word problem-solving measures for monitoring third-grade students/ mathematics competence. *Assessment for Effective Intervention, 32*, 90-99.

Marsh, H. W. (1993). Multitrait-multimethod analyses: Inferring each trait-method combination with multiple indicators. *Applied Measurement in Education*, *6*, 49-81.

Mather, N. & Woodcock, R. W. (2001). Administration and scoring manual. *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.

McGrew, K. S., Schrank, F. A., & Woodcock, R. W. (2007). Technical Manual. *Woodcock-Johnson III Normative Update.* Rolling Meadows, IL: Riverside Publishing.

Minneapolis Public Schools. (2004). *Minneapolis Kindergarten Assessment*. Minneapolis, MN: Minneapolis Public School Research, Evaluation & Assessment Division.

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: Author.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academy Press.

Pickart, M., Betts, J., Sheran, C., & Heistad, D. (2005). *Minneapolis Kindergarten Assessment. Beginning and end of kindergarten assessment: Technical Manual*. Minneapolis, MN: Minneapolis Public Schools.

Pearson. (2009). *Wechsler Individual Achievement Test* (3rd ed.). San Antonio, TX: Author.

PsycCorp/Pearson. (2004). *AIMSweb*. San Antonio, TX: Author.

PsycCorp/Pearson. (2009). *AIMSweb*. San Antonio, TX: Author.

Salvia, J., & Ysseldyke, J. E. (1991). *Assessment* (5th ed.). Boston, MA: Houghton Mifflin.

Shah, R., & Goldstein, S. M. (2006). Use of structural equation modeling in operations management research: Looking back and forward. *Journal of Operations Management*, *24*, 148-169.

Shapiro, E. S. (2004). *Academic skills problems* (3rd ed.). New York, NY: Guilford Press.

Shapiro, E. S., Edwards, L., & Zigmond, N. (2005). Progress monitoring of mathematics among students with learning disabilities. *Assessment for Effective Intervention, 30*, 15-32.

Shinn, M. R., & Shinn, M. M. (2002). *Administration and Scoring of Reading Maze for Use in General Outcome Measurement*. EdenPrairie, MN: Edformation, Inc.

Shinn, M. R., Tindal, G., & Stein, S. (1988). Curriculum-based assessment and the identification of mildly handicapped students: A research review. *Professional School Psychology*, *3,* 69-86.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (5th ed). Boston, MA: Allyn and Bacon.

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

Tindal, G., Germann, G., & Deno, S. L. (1983). Descriptive research on the Pine County norms; A compilation of findings (Res. Rep. No. 132). Minneapolis, MN: University of Minnesota Institute for Research on Learning Disabilities.

University of Oregon. (2009). *easyCBM*. Eugene, OR: Author.

VanDerHeyden, A. M., & Burns, M. K. (2008). Examination of the utility of various measures of mathematics proficiency. *Assessment for Effective Intervention, 33*, 215-224.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock Johnson III*. Itasca, IL: Riverside.


Dear Parent,

Ms. Scott, assistant director of Campus School, has agreed to allow Campus School students to participate in a study to better understand mathematics fluency.

We will be asking Campus School students to complete several brief math tests in their classrooms. These tests are designed to evaluate basic math skills, including applying knowledge of basic addition and subtraction facts, as well as math applications skills, such as using graphs. For a contrast, students will also complete tests designed to assess reading comprehension skills and general speed in completing paper-and-pencil tasks.

Every child in second, third, fourth, and fifth grade at Campus School is being asked to participate; no child is being singled out. The tests will be completed in the classroom in 2 group sessions lasting approximately 45 minutes during the regular math instruction period. No research information about any individual child will be made available to any teacher or administrator at the school. Our information will be kept completely confidential. All data will be coded with numbers, and all publications and reports to the school resulting from this research will refer to group results and not results for individual students. Again, no individual child will ever be identified by name.

Campus School students will be at minimal risk of psychological or physical discomfort or harm during the completion of this research. Children will be told that they are not required to participate, and they will be assured that there will be no consequences should they make this decision. Participation or non-participation in this study will not impact the services or education your child receives at school.

We would like for as many students as possible to participate in this project because it is important to understand students abilities at all levels. **You need not respond to this letter unless you do not want your child to participate in this research.** If you have any questions concerning this project please call us at 678-2145 and ask to speak to Jennifer Maynard. For answers to questions regarding research subjects' rights, you may contact the Chair of the Committee for the Protection of Human Research Participants at 678-2533.

Sincerely,

Jennifer L. Maynard, M.S.
Doctoral Student, School Psychology Program
The University of Memphis

Please Keep This Form For Your Records

Please sign and return this portion of the form to school with your child if you DO NOT want your child to participate in this project.

I DO NOT wish for my child, ______________________, to participate in the mathematics study. My child is in the following teacher's class: ______________________

____
Parent Signature

University of Memphis IRB # H10-68
Approval of this form expires 4/4/11

University of Memphis IRB # H10-68
Approval of this form expires 5/4/11

APPENDIX B: Participant Assent

Your mother, father, or guardian has decided that you *can* work with us today, but we need your permission, too. We need to make sure that you know about what we are going to do together today and that you want to work some problems with me. We hope it will be fun. Here are some of the things we will do.

- We want to see how you answer different math problems. Some of the problems will be on worksheets and some will be shown to you on the board. Some of these activities are timed, some are untimed. Some are very short, and others take a little longer. We'll use the papers in the envelopes you were given.

- We'll spend about an hour and a half in your classroom doing these activities together.

- You can decide at any time today that you don't want to do these activities, and it will be OK. We think that we will have a good time today though.

- We want you to do your best on each of these activities, but we won't tell your teacher, your friends, or anyone else at the school or at home how you did. When we look at the results, we'll look at what you did and what other children at your school did all at once. You each have a special number that is written on your packet instead of your name. After you are finished I will not be able to tell which answers are yours just by looking at your packet.

If you have any questions, please ask them now.

If you want to do these activities with us today, write your name or a letter in your name in this box.

If you <u>don't want to do these activities</u> with us today, write your name or a letter in your name in this box.

______________________________ ______________________

Examiner's Signature Date

APPENDIX C1: Introduction Script

Lead Researcher (LR): Good (morning/afternoon)! We are researchers at the University of Memphis and we would like for you to do some activities with us today. Everyone will need to have a pencil on their desk. Please open your folder, take only the first page out, and place them it on your desk. You can read silently with me as I read it to you.

Your mother, father, or guardian has decided that you *can* work with us today, but we need your permission, too. We need to make sure that you know about what we are going to do together today and that you want to work some problems with me. We hope it will be fun. Here are some of the things we will do.

- We want to see how you answer different math problems. Some of the problems will be on worksheets and some will be in workbooks. Some of these activities are timed, some are untimed. Some are very short, and others take a little longer. We'll use the papers in the folders you were given.

- We'll spend about 45 minutes in your classroom doing these activities together.

- You can decide at any time today that you don't want to do these activities, and it will be OK. We think that we all will have a good time today though.

- We want you to do your best on each of these activities, but we won't tell your teacher, your friends, or anyone else at the school or at home how you did. When we look your answers, we'll look at what you did and what other children at your school did all at once. You will each have a special number we will write on your papers instead of your name. After you are finished I will not be able to tell which answers are yours just by looking at your packet.

Do you have any questions about what we will be doing today?

If you want to do these activities with us, write your name or a letter from your name in the first box.

If you do not want to do these activities with us, write your name or a letter from your name in the second box.

Research assistants will walk around the room and collect the student consent forms. Students who indicate that they do not want to participate will go to the library or to another chosen location for the duration of the testing

***Folders will be distributed to the students in the classroom ***

LR: Now that you all know what we are going to do today, let's begin! All of the activities that we will be doing over the next two days (today) are in a special order. Be sure to leave each activity in the folder until I tell you to take it out. The pages are all facing the right way, so don't turn the pages over or open

the booklets until I tell you to. After we finish each activity, you will place the pages you just worked on in the left pocket of the folder.

There are a few things I want you to remember when we're doing these activities.

- Each activity is different and has different instructions. Listen carefully when I give the instructions for a new activity.
- Don't write your name, the date, or your teacher's name anywhere.
- For some of these activities, you will only have a short amount of time to finish them. On others, you will have longer. I will tell you which tests happen fast and which tests are longer.
- On the math tests, we can't tell you how to work a problem. If you don't know how to do a problem, skip the problem and keep going. You can go back to any problems you skipped when you finish the problems you know how to do.
- You do not have to show your work if you don't want to. It's up to you.
- If an activity has more than one page, when you get to the bottom of a page, turn the page and keep working until you get to the end of the activity or I say, "Stop."
- If you finish an activity before I say, "Stop," put your pencil down and sit quietly while others finish the activity.
- On some activities I will ask you to write down what the time is when you finish. Look at the computer screen and write down the minutes and seconds that are showing when you finish that activity.

Do you have any questions? Always ask questions before we begin because I can't answer questions once we start each activity.

Are you all ready? Great! Let's begin. Remember to listen carefully to the directions for each activity. Open your folder and take out the next activity, close your folder, and put the activity on top of your folder.

APPENDIX C2: AIMSweb M-CBM

AIMSweb: Computation

We're going to take a 2(4)-minute math test. I want you to write your answers to several kinds of math problems. Some are addition and some are subtraction. Look at each problem carefully before you answer it.

When I say 'BEGIN' write your answer to the FIRST problem **(demonstrate by pointing)** and work ACROSS the page. Then go to the next row.

Try to work EACH problem. If you come to one YOU REALLY DON'T KNOW HOW TO DO, put an 'X' through it and go to the next one.

If you finish the first side, turn it over and continue working. Are there any questions? **(Pause)**

If you finish before I say, "Stop," look at the computer screen and write down what time it says.

APPENDIX C3: AIMSweb M-CAP

AIMSweb: Concepts and Applications

We're going to take an 8-minute math test.

Read the problems carefully and work each problem in the order presented. Do not skip around. If you do not know how to work a problem, mark it with an X and move on. Once you have tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.

Write the answers to the problems in the blanks. For multiple choice questions, place the letter (A, B, or C) of the correct answer in the blank. You do not have to show your work, but you may if that is helpful for you in working the problems. Keep working until you have completed all of the problems or I tell you to stop. Do you have any questions? **(Pause)**

If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENDIX C4: MBSP Computations

MBSP: Computation

On the next test I want you to do as many problems as you can. Work carefully and do the best you can. Start at the top left. Work from left to right. Some problems will be easy for you; others will be harder. When you come to a problem you know you can do, do it right away. When you come to a problem that's hard, skip it and come back to it later.

Go through the entire test doing the easy problems. Then go back and try the harder ones. Remember, you might get points for getting part of a problem right. So, after you've done all the easy problems, try the harder problems. Try to do each problem even if you think you can't get the whole problem right.

When I say, "Begin," turn your test over and start to work. Work for the whole test time. You should have enough room to do your work in each block on the page. Write your answers so I can read them! If you finish early, check your answers. At the end of 3 (5) minutes, I will say, "Stop."

Do you have any questions? If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENDIX C5: MBSP Applications

MBSP: Concepts and Applications

On the next test I want you to do as many problems as you can. Work carefully and do the best you can. Start at the first problem, work down the first column, and then down the second column. Some problems will be easy for you; others will be harder. When you come to a problem you know you can do, do it right away. When you come to a problem that's hard for you, skip it and come back to it later.

Go through the entire test doing the easy problems. Then go back and try the harder ones.

When I say, "Begin," start to work. Work for the whole test time. You should have enough room to do your work in each block on the page. Write your answers so I can read them! If you finish early, check your answers. At the end of 6 (7) minutes, I will say, "Stop." Put your pencil down.

Do you have any questions? If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENDIX C6: easyCBM Math Geometry

easyCBM

Now you’re going to do some more math problems. You will have up to 10 minutes to work these problems, so take your time and carefully answer each problem. For this activity, it doesn’t matter how fast you finish. When you finish all the problems, put your pencil down so we will know you are finished.

For this activity, read the problems carefully and work each problem in the order presented. Do not skip around. If you do not know how to work a problem, mark it with an X and move on. Once you have tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.

Circle the letter (A, B, or C) of the correct answer. You do not have to show your work, but you may if that is helpful for you in working the problems. Keep working until you have completed all of the problems or I tell you to stop. Do you have any questions? **(Pause)**

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENDIX C7: WJ III ACH Calculation

WJ III: Calculation

For this activity, you will have up to 10 minutes to work these problems, so take your time and carefully answer each problem. It doesn't matter how fast you finish these problems. You will not know how to do some of these problems- that's ok! Do your best to answer as many problems as you can. When you finish all the problems you know how to do, put your pencil down so we will know you are finished. This activity is only on pages 6 and 7, so don't turn the page. Turn to pages 6 and 7 now.

I want you to do some math. Begin with number 1 and answer as many problems as you can. If you come to one you do not know how to do, just skip it and try the next one.

You will have 10 minutes. Do you have any questions? **(Pause)**

Begin.
Stop. Put your pencils down.

APPENDIX C8: WJ III ACH Math Fluency

WJ III: Math Fluency

When I say, "Begin," I want you to turn the page.

I want you to work some simple arithmetic problems. Start with the first problem. When you finish a row, go to the next one and work each problem until you finish the page. The go to the top of the next page. If you cannot think of an answer, skip that item and move to the next one. Work as fast as you can without making mistakes. Be sure to watch the signs. If you do make a mistake, just cross out the answer you do not want. You will have 3 minutes. Do you have any questions? **(Pause)**

Remember: do not erase; just cross out any answers you don't want.

If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENDIX C9: WIAT-III Math Fluency

WIAT-III: Math Fluency (all versions)

For the next activity, you will use three worksheets that are folded. Leave them closed until I tell you to begin.

I want you to write the answers to as many math problems as you can in one minute. Start with the first problem and work the problems going across the row before you move on to the next row. If you come to a problem you don't know, just skip it. If you finish the first page, turn to the next page. Do you have any questions? **(Pause)**

If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

APPENIX C10: AIMSweb Practice Maze and Maze 1

AIMSweb Maze Practice and First Maze Probe

When I say, “Begin,” I want you to silently read a story. You will have 3 minutes to read the story and complete the task. Listen carefully to the directions. Some of the words in the story are replaced with a group of three words. Your job is to circle the 1 word that makes the most sense in the story. Only 1 word is correct.

Let's practice one together. Look at your first page. Read the first sentence silently while I read it out loud: 'The dog apple, broke, ran after the cat.' The three choices are apple, broke, ran. 'The dog apple after the cat.' That sentence does not make sense. 'The dog broke after the cat.' That sentence does not make sense. 'The dog ran after the cat.' That sentence does make sense, so circle the word ran."

Let's go to the next sentence. Read it silently while I read it out loud. The cat ran fast, green, for up the hill. The three choices are fast, green, for. Which word is the correct word for the sentence?

(Students answer fast)

Yes, ‘The cat ran fast up the hill.’ is correct, so circle the correct word fast.

(Make sure students circle fast)

Silently read the next sentence and raise your hand when you think you know the answer. *What is the answer?*

(Make sure students know the correct word. Read the sentence with the correct answer)

That’s right, ‘The dog barked at the cat.’ is correct. Now what do you do when you choose the correct word?

(Students answer “Circle it.” Make sure the students understand the task)

That's correct, you circle it. I think you're ready to work on a story on your own.

When I say, “Begin,” turn your page over to the story and start reading silently. When you come to a group of three words, circle the 1 word that makes the most sense. Work as quickly as you can without making mistakes. (2nd and 3rd: If you finish the first side of the page, turn the page and) Keep working until I say, “Stop,” or you are all done.

If you finish before I say, “Stop,” look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder.

APPENDIX C11: AIMSweb Maze 2

AIMSweb Maze

When I say, “Begin,” I want you to silently read a story. You will have 3 minutes to read the story and complete the task. Listen carefully to the directions. Some of the words in the story are replaced with a group of three words. Your job is to circle the 1 word that makes the most sense in the story. Only 1 word is correct.

When I say, “Begin,” turn your page over to the story and start reading silently. When you come to a group of three words, circle the 1 word that makes the most sense. Work as quickly as you can without making mistakes. (2^{nd} and 3^{rd}: If you finish the first side of the page, turn the page and) Keep working until I say, “Stop,” or you are all done.

If you finish before I say, “Stop,” look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder.

APPENDIX C12: WJ III COG Decision Speed

WJ III: Decision Speed

I want to find out how fast you can find two things that either go together or are most alike. Look at the first row of pictures. There are two cats in that row. Draw a circle around each cat.

Now look at the second row of pictures. Draw a circle around the two things that go together. You should have drawn a circle around the shoes.

Now try the next three. Find the two things in each row that go together. Do them as fast as you can.

You should have drawn circles around the cat and dog because they are animals, the moon and sun because they are in the sky, and the apple and pear because they are fruits.

When I say begin, turn the page and draw circles around the two things in each row that go together. Work as fast as you can without making mistakes. If you do make a mistake, just cross out the one you do not want. If you have trouble finding two things that go together, skip that row and move on to the next one. After you get to the bottom of a page, go to the top of the next one. Keep working until I tell you to stop. Do you have any questions? **(Pause)**

If you finish before I say, “Stop,” look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder.

APPENDIX C13: WJ III COG Pair Cancellation

WJ III: Pair Cancellation

Look at the two pictures at the top of the page. The first one is a ball, and it is followed by a picture of a dog. There are other pictures in this row. Each picture is either the ball, the dog, or a cup. There is a circle drawn around the ball followed by the dog. Do you see another example of the ball followed by the dog? Draw a circle around the pair of pictures where the ball is followed by the dog.

Here are two more rows of pictures. Draw a circle around each pair of pictures where the ball is followed by the dog. Work as fast as you can without making mistakes. Go ahead.

When I say, "Begin," turn your page over. Start at the top of the next page and draw a circle around every ball followed by a dog in each row. Work as fast as you can without making mistakes. If you do make a mistake, cross out the pictures you do not want. You will have 3 minutes. Do you have any questions? **(Pause)**

If you finish before I say, "Stop," look at the computer screen and write down what time it says.

Begin.

Stop. Put your pencils down and place that sheet in the left pocket of your folder. Take the next activity out of the right pocket of your folder.

Appendix D1: AIMSweb M-CBM

AIMSweb M-CBM
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *We're going to take a 2(4)-minute math test. I want you to write your answers to several kinds of math problems. Some are addition and some are subtraction. Look at each problem carefully before you answer it.* *When I say 'BEGIN' write your answer to the FIRST problem* ***(demonstrate by pointing)*** *and work ACROSS the page. Then go to the next row.* *Try to work EACH problem. If you come to one YOU REALLY DON'T KNOW HOW TO DO, put an 'X' through it and go to the next one.* *If you finish the first side, turn it over and continue working. Are there any questions?*	______________
Provided opportunity for students to ask questions before beginning measure	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Allowed up to 2(4) minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D2: AIMSweb MCAP

AIMSweb MCAP
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *We're going to take an 8-minute math test.* *Read the problems carefully and work each problem in the order presented. Do not skip around. If you do not know how to work a problem, mark it with an X and move on. Once you had tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.* *Write the answers to the problems in the blanks. For multiple choic questions, place the letter (A, B, or C) of the correct answer in the blank. You do not have to show your work, but you may if that is helpful for you in working the problems. Keep working until you ha completed all of the problems or I tell you to stop. Do you have any questions?*	______________
Provided opportunity for students to ask questions before beginning measure	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Allowed 8 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D3: MBSP Computation

MBSP Computation
1 = completed accurately   0 = incorrectly completed

Testing Procedure **Observation**

Students had a pencil and were provided with an extra pencil if needed ______________

Said instructions:

On the next test I want you to do as many problems as you can. Wo carefully and do the best you can. Start at the top left. Work from left to right. Some problems will be easy for you; others will be harder. When you come to a problem yo know you can do, do it right away. When you come to a problem that's hard, skip it and come back to it later.

Go through the entire test doing the easy problems. Then go back a try the harder ones. Remember, you might get points for getting part of a problem right. So, after you've done all the easy problems, try the harder problems. Try to each problem even if you think you can't get the whole problem right.

When I say, "Begin," turn your test over and start to work. Work fo the whole test time. You should have enough room to do your work each block on the page. Write your answers so I can read them! If you finish early, check your answers. At the end of 3 (5) minutes, I will say, "Stop."

____

Provided opportunity for students to ask questions before beginning measure ______________

Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop ______________

Allowed 3(5) minutes for students to complete measure ______________

Instructed students to put away completed measure and retrieve next measure from folder ______________

APPENDIX D4: MBSP Applications

MBSP Applications
1 = completed accurately   0 = incorrectly completed

Testing Procedure **Observation**

Students had a pencil and were provided with an extra pencil if needed ______________

Said instructions:

On the next test I want you to do as many problems as you can. Wo carefully and do the best you can. Start at the first problem, work down the first column, and then down the second column. Some problems wil be easy for you; others will be harder. When you come to a problem you know you can do, it right away. When you come to a problem that's hard for you, skip it and come back to it later.

Go through the entire test doing the easy problems. Then go back a try the harder ones.

When I say, "Begin," turn your test over and start to work. Work fo the whole test time. You should have enough room to do your work in each block on the page. Write your answers so I can read them! If you finish early, check your answers. At the end of 6 (7) minutes, I will say, "Stop." Put your pencil down.

____

Provided opportunity for students to ask questions before beginning measure ______________

Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop ______________

Allowed 6(7) minutes for students to complete measure ______________

Instructed students to put away completed measure and retrieve next measure from folder ______________

APPENDIX D5: easyCBM Math Geometry

easyCBM
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	____________
Said instructions: *Now you're going to do some more math problems. You will have up to 10 minutes to work these problems, so take your time and carefully answer each problem. For this activity, it doesn't matter how fast you finish. When you finish all the problems, put your pencil down so we will know you are finished.* *For this activity, read the problems carefully and work each problem in the order presented. Do not skip around. If you do not know how to work a problem, mark it with an X and move on. Once you had tried all of the problems in order, you may go back to the beginning of the worksheet and try to complete the problems you marked.* *Circle the letter (A, B, or C) of the correct answer. You do not have show your work, but you may if that is helpful for you in working th problems. Keep working until you had completed all of the problem or I tell you to stop. Do you have any questions?*	____________
Provided opportunity for students to ask questions before beginning measure	____________
Allowed up to 10 minutes for students to complete measure	____________
Instructed students to put away completed measure and retrieve next measure from folder	____________

APPENDIX D6: WJ III ACH Calculation

WJ III Calculation
1 = completed accurately 0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *I want you to do some math. Begin with number 1 and answer as many problems as you can. If you come to one you do not know how to do just skip it and try the next one.*	______________
Instructed students to take their time completing the problems because they are not being timed	______________
Told students that they will not know how to do some of the problems and that is OK	______________
Instructed students to do their best	______________
Told students that they will only be working problems on pages 6 and 7 and they should not turn the page if they get to the end of the problems	______________
Provided opportunity for students to ask questions before beginning measure	______________
Allowed up to 10 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D7: WJ III ACH Math Fluency

WJ III Math Fluency
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Instructed students to not turn the page until they are told to begin	______________
Said instructions: *I want you to work some simple arithmetic problems. Start with the first* *problem. When you finish a row, go to the next one and work each problem until you finish the page. The go to the top of the next page If you cannot think of an answer, skip that item and move to the next one. Work as fast as you can* *without making mistakes. Be sure to watch the signs. If you do make mistake,* *just cross out the answer you do not want. You will have 3 minutes. Do you have any questions?*	______________
Provided opportunity for students to ask questions before beginning measure	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Allowed 3 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D8: WIAT-III Math Fluency

WIAT-III Math Fluency
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Told students to not open folded worksheets until they are told to begin	______________
Said instructions: *I want you to write the answers to as many math problems as you c* *in one minute. Start with the first problem and work the problems* *going across the row before you move on to the next row. If you co* *to a problem you don't know,* *just skip it. If you finish the first page, turn to the next page. Do you* *have any questions?*	______________
Provided opportunity for students to ask questions before beginning measure	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Allowed 1 minute for students to complete each measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D9: AIMSweb Maze 1

AIMSweb Maze Practice and First Maze Measure
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________

Said instructions:

When I say, "Begin," I want you to silently read a story. You will had 3 minutes to read the story and complete the task. Listen carefully to the directions. Some of the words in the story are replaced with a group of three words. Your job is to circle the 1 word that makes the most sense in the story. Only 1 word is correct.

Let's practice one together. Look at your first page. Read the first sentence silently while I read it out loud: 'The dog apple, broke, ran after the cat.' The three choices are apple, broke, ran. 'The dog apple after the cat.' That sentence does not make sense. 'The dog broke after the cat.' That sentence does not make sense. 'The dog ran after the cat.' That sentence does make sense, so circle the word ran."

Let's go to the next sentence. Read it silently while I read it out loud. The cat ran fast, green, for up the hill. The three choices are fast, green, for. Which word is the correct word for the sentence?

Yes, 'The cat ran fast up the hill.' is correct, so circle the correct word fast.
Silently read the next sentence and raise your hand when you think you know the answer. What is the answer?

That's right, 'The dog barked at the cat.' is correct. Now what do you do when you choose the correct word?

That's correct, you circle it. I think you're ready to work on a story your own.

____

Allowed students to give correct answers aloud to sample sentences	______________

Said instructions:

When I say, "Begin," turn your page over to the story and start reading silently. When you come to a group of three words, circle tl 1 word that makes the most sense. Work as quickly as you can without making mistakes. ***(2nd and 3rd:*** *If you finish the first side of the page, turn the page an Keep working until I say, "Stop," or you are all done.*

____

Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop ______________

Provided opportunity for students to ask questions before beginning measure ______________

Allowed 3 minutes for students to complete measure ______________

Instructed students to put away completed measure and retrieve next measure from folder ______________

APPENDIX D10: AIMSweb Maze 2

AIMSweb Second Maze Measure
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *When I say, "Begin," turn your page over to the story and start reading silently. When you come to a group of three words, circle tl 1 word that makes the most sense. Work as quickly as you can without making mistakes.* ***(2nd and 3rd:*** *If you finish the first side of the page, turn the page an Keep working until I say, "Stop," or you are all done.*	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Provided opportunity for students to ask questions before beginning measure	______________
Allowed 3 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D11: WJ III COG Decision Speed

WJ III Decision Speed
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *I want to find out how fast you can find two things that either go together or are most alike. Look at the first row of pictures. There are two cats in that ro Draw a circle around each cat.* *Now look at the second row of pictures. Draw a circle around the t things that go together. You should had drawn a circle around the shoes.* *Now try the next three. Find the two things in each row that go together. Do them as fast as you can. You should have drawn circles around the cat and dog because they are animals, the moon and sun because they are i the sky, and the apple and pear because they are fruits.* *When I say begin, turn the page and draw circles around the two things in each row that go together. Work as fast as you can without making mistakes. If you do make a mistake, just cross out the one you do not want. If you had trouble finding two things that go together, skip that row and move on to the next one. After you get to the bottom of a page, go to the top of the next one. Keep working until I tell you to stop. Do you have any questions?*	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Provided opportunity for students to ask questions before beginning measure	______________
Allowed 3 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________

APPENDIX D12: WJ III COG Pair Cancellation

WJ III Pair Cancellation
1 = completed accurately   0 = incorrectly completed

Testing Procedure	**Observation**
Students had a pencil and were provided with an extra pencil if needed	______________
Said instructions: *Look at the two pictures at the top of the page. The first one is a bal and it is followed by a picture of a dog. There are other pictures in this row. Each picture is either the ball, the dog, or a cup. There is a circle drawn around the ball followed by the dog. Do you see another example of the ball followe by the dog? Draw a circle around the pair of pictures where the ball is followed by the dog.* *Here are two more rows of pictures. Draw a circle around each pai of pictures where the ball is followed by the dog. Work as fast as yo can without making mistakes. Go ahead.* *When I say, "Begin," turn your page over. Start at the top of the ne page and draw a circle around every ball followed by a dog in each row. Work as fast as you can without making mistakes. If you do make a mistake, cross o the pictures you do not want. You will have 3 minutes. Do you have any questions?*	______________
Instructed students to record finishing time by looking at the computer screen if they finish the passage before they are told to stop	______________
Provided opportunity for students to ask questions before beginning measure	______________
Allowed 3 minutes for students to complete measure	______________
Instructed students to put away completed measure and retrieve next measure from folder	______________