wireless generation®

mCLASS®:Math

Identifying Technically Adequate Early Mathematics Measures

May 4, 2007

Young-Sun Lee, Ph.D. Teachers College, Columbia University

Erica Lembke, Ph.D. University of Missouri

Douglas Moore, Wireless Generation

Herbert Ginsburg, Ph.D. Teachers College, Columbia University

Sandra Pappas, Ph.D. Wireless Generation

Technical White Paper

mCLASS®:Math
Identifying Technically Adequate Early Mathematics Measures

Summary

Research shows that early intervention in academic areas is critical for identifying students who may be at risk and in preventing them from future academic failure. The challenge to date has been the development of an assessment system that provides teachers with easy, efficient, and reliable methods to identify students who may be at risk for failure. Research has also shown that the frequent use of formative assessments can be used to inform instruction thereby improving achievement. Although some work in the area has begun, there are few measures that can be used for screening and progress monitoring young students' mathematical performance, particularly students in kindergarten and Grade 1.

The present paper describes a study that was conducted to examine the technical adequacy of Curriculum-Based Measurement (CBM) measures of early numeracy. Early numeracy reflects three broad and theoretically derived categories of mathematical thinking: Quick Retrieval, Written Computation, and Number Sense. These categories represent essential mathematical skills.

Six one-minute early mathematics tasks were administered to 137 kindergarten and first grade students, along with a standardized test of early mathematics. The CBM measures included Count Out Loud, Quantity Discrimination, Number Identification, Missing Number, Next Number, and Number Facts. Students' performance on each measure was correlated with the criterion variable, which was each student's scaled score on a norm-referenced achievement measure. Participating students completed two rounds of data collection over the course of four days.

Reliability of the measures was examined using Cronbach's alpha and split-half reliability, and validity was examined using concurrent criterion validity and construct validity, which was examined through a confirmatory factor analysis. In addition, an item analysis was conducted for each measure. Item difficulty was important for determining whether students had mastered the concept being tested.

Results indicated that the measures produced moderately strong reliability coefficients, moderate criterion validity coefficients with the standardized early mathematics test, and combined, that the tasks appeared to be measuring the construct of early numeracy proficiency. In brief, the measures appear to capture students' performance from kindergarten through Grade 1, are able to discriminate among students at a variety of levels, and display adequate levels of item difficulty and satisfactory discrimination for each grade. Although more research is needed, the results of this study indicate that these early mathematical measures can serve as good screening tools for students in early elementary school.

Introduction

Early intervention in academic areas can play a key role in identifying students who might be at risk and in preventing them from becoming later victims of academic failure. The literature supports intervention in reading and mathematics in kindergarten, if not before (Campbell & Ramey, 1994; Vellutino et al., 1996). However, teachers must have easy, efficient, and reliable methods to identify students that might be at risk for failure. These methods must involve practical assessment that can be utilized to monitor student progress and provide teachers with insights that can guide teaching. Teachers often do not have the means to keep track of how well their students are performing, and have little knowledge of when to intervene with students that are not demonstrating adequate growth. The National Research Council's *Eager to Learn Report* (Bowman, Donovan, & Burns, 2001) calls for assessment of young children to be built into the instructional process. The NAEYC/NCTM joint position paper advocates that the teacher should "… support children's learning by thoughtfully and continually assessing all children's mathematical knowledge, skills, and strategies" (National Association for the Education of Young Children and National Council of Teachers of Mathematics, 2002 p.4). Other reports also stress the integration of assessment and instruction (Pellegrino, Chudowsky, & Glaser, 2001).

The frequent use of formative assessments can improve achievement, particularly when the results are used to adjust instruction (Stecker, Fuchs, & Fuchs, 2005; Speece, Molloy, & Case, 2003; Bergan, Sladeczek, Schwarz, & Smith, 1991; Black & William, 1998). Providing teachers with specific supports for using the resulting information to adjust instruction further raises achievement. In a meta-analysis of 21 controlled studies on the achievement of mildly disabled students, Fuchs and Fuchs (1986) found an average effect size of .70 standard deviations for classrooms that used assessment data to graph children's progress and to adjust instruction. These results were compared to an effect size of 0.26 standard deviations for classrooms that used formative assessments, but did not systematically organize or act on the resulting data. In subsequent studies, they found that using a technology-based expert system to generate graphs of student progress and instructional recommendations enables students to improve more quickly, and achieve higher outcomes, than their peers in a control group (Fuchs et al., 1994; Fuchs, Fuchs, & Hamlett, 1989a, 1989b).

At present, the field is limited with respect to measures that can be used for monitoring young students' mathematics performance. To be sure, some work in the area has begun and appears promising. Researchers are beginning to develop theoretical analyses useful for creating mathematics assessments (Gersten, Jordan, & Flojo, 2005) and to create brief assessment measures that can serve as indicators of mathematics performance. Clarke and Shinn (2004) and Chard et al. (2005) have examined the use of tasks such as number identification, quantity discrimination, and missing number with kindergarten and first grade students. van der Heyden, Witt, Naquin, and Noell (2001) have also examined early math measures that could be used as screening tools in the identification of kindergarten students in need of academic intervention.

With the recognition that mathematics difficulties in early elementary school may lead to later mathematics deficits, it is important to conduct early screening and identification of students who appear to be struggling with mathematics. The measures utilized should entail basic mathematical concepts and skills and also need to be reliable and valid for the purposes for which they are used (in this case, finding the students who might need intervention). One form of screening for students at risk in academic areas is a Curriculum-Based Measurement (CBM). CBM is a simple method of repeated measurement toward long-range instructional goals (Deno, 1985). To use the system, teachers administer and score short assessments on a frequent basis, graph the scores, evaluate student progress toward goals, and use the data to make decisions about the effectiveness of their instruction and to identify students who appear to be struggling and are at risk. Assessments of this type can guide teachers in grouping their students into three instructional tiers: those students who are on track to maintain proficiency in mathematics, those students who are in need of additional intervention, and those students who need the most intensive intervention. Teachers can also use the measures for progress monitoring, administering alternate forms of the CBM measures weekly or bi-weekly, and evaluating whether students remain on track to meet their long-range goals.

One of the primary tasks used when screening and progress monitoring in mathematics has been mastery of "basic facts" (Marston, 1989). However, until recently, there has not been an option for screening and progress monitoring in mathematics for students in grades K-1 who cannot yet master basic facts. As mentioned previously, in studies by Clarke and Shinn (2004) and Chard et al. (2005), kindergarten (n = 436) and first grade (n = 483) students were administered tasks such as counting, number identification, quantity discrimination, and missing number. At the kindergarten level, number identification, quantity discrimination, and missing number tasks had the strongest alternate form reliability and criterion validity, with correlations ranging from .50 to .69 with the criterion, the Number Knowledge Test (Okamoto & Case, 1996). The strongest correlation at the kindergarten level was found between the missing number task and the criterion. Oral counting, quantity discrimination, number identification, and missing number tasks all demonstrated adequate reliability and validity with the criteria, Woodcock Johnson Applied Problems (Woodcock & Johnson, 1989), Number Knowledge Test, and CBM basic facts probes, for first grade students with correlation coefficients ranging from .60 to .79.

van der Heyden et al. (2001) and van der Heyden et al. (2004) examined early mathematics measures as part of studies to identify screening tools that could be used to identify kindergarten students in need of academic intervention and to screen preschool children. Measures such as circle number, write number, and draw circles were administered to kindergarten students, with correlations ranging from .44 to .61 with the math composite on the Comprehensive Inventory of Basic Skills, revised edition (Brigance, 1999). For four-year olds (n = 53), tasks such as choosing number, number naming, counting objects, free counting, discrimination, and choosing shapes were correlated with the Brigance Screens (Brigance, 1985) and the Test of Early Mathematics Achievement-2 (TEMA-2; Ginsburg & Baroody, 1990). The choosing number and discrimination tasks demonstrated the strongest correlations with these criterion variables, ranging from .52 to .57.

With limited but promising research completed on technically adequate early mathematics CBM measures, the purpose of our study was to examine the reliability and validity of a set of CBM measures reflecting three broad and theoretically derived categories of mathematical thinking: Quick Retrieval, Written Computation, and Number Sense. These three classes of measures cover essential aspects of student mathematics: the abilities to retrieve information quickly, to compute, and to have a good feel for the meaning of the material. In the present study, we examined potential indicators of these broad categories of mathematical ability, including providing oral answers to number facts as a measure of Quick Retrieval, Counting Out Loud, Quantity Discrimination, Next Number, and Missing Number tasks as measures of Number Sense, and a precursor to Written Computation, and Number Identification.

There is widespread agreement (Ansari & Karmiloff-Smith, 2002; Gersten et al., 2005) that quick retrieval of the basic number facts or combinations is both important for learning mathematics and useful for a CBM task. The research shows that knowledge of number combinations predicts achievement, and that children having difficulty with mathematics often do poorly on number combinations (Gersten et al., 2005; Russell & Ginsburg, 1984). From a practical point of view, quick retrieval means that children get the answer quickly and do not have to expend much mental energy on calculation. Quick retrieval may result from rote memory, from fluency with rapid counting or from application of basic principles such as commutativity of addition. In the proposed CBM assessment system, quick retrieval in kindergarten is not assessed, as students at this level would not be expected to master this kind of material. The measure that is examined in this study to assess quick retrieval is number facts in first grade.

Written computation includes basic reading and calculation problems. The task that we included to assess a precursor to written computation was asking students in kindergarten and Grade 1 to verbally identify randomly presented written numerals from 1 through 20. Previous research has found high alternate form reliability, test-retest reliability, and criterion validity for numeral identification (Clarke & Shinn, 2004; Chard et al., 2005; Lembke & Foegen, 2005).

Number sense is a kind of mathematical street smarts. It involves basic "intuitions" and ideas about numbers, including concepts that make computation easier or eliminate the need for it. Number sense includes the ability to compare the magnitude of numbers, to understand the relative effect of arithmetical operations on numbers, and to have meaningful referents for number and quantity (NCTM, 1989). It also includes the ability to know whether certain numbers are plausible answers to certain problems, to break numbers into convenient parts that make calculation easier, or to apply basic operational rules to arithmetical equations. Although notoriously hard to define, number sense has shown to be important in predicting academic achievement (Gersten et al., 2005). Number sense is assessed in the present study using Counting Out Loud, Quantity Discrimination (Berch, 2005; Clarke & Shinn, 2004), and Missing Number (Chard et al., 2005; Lembke & Foegen, 2005) tasks for kindergarten and first grade students in addition to the Next Number (Okamoto & Case, 1996; Wright, 1991) task for first grade students. Quantity discrimination and missing number tasks demonstrated moderate to high alternate form reliability and criterion validity in previous research (Okamoto & Case, 1996; Lembke & Foegen, 2005).

In this study, we hoped to develop and evaluate more powerful and theoretically-based CBM measures of early mathematics. We attempted to replicate and extend existing research on reliable and valid measures, introduce new measures, modify existing measures, and conduct criterion validity studies.

Method

Participants

The site for the study was an elementary school in a large Southern urban city. The total enrollment for the district was approximately 30,000 students, with the primary ethnicity being African American (96%), followed by 2% Hispanic, 1% Caucasian, and 1% other ethnicities. One hundred and thirty seven students (68 in kindergarten and 69 in first grade) participated in the study, with ethnicities representing that of the district (99% African American and 1% Caucasian).

Measures

In kindergarten, students' ability to count out loud (Count Out Loud), their ability to discriminate between the quantities of two sets of numbers (Quantity Discrimination), their ability to name printed numbers between 1 and 20 (Number Identification), and their ability to identify the missing number from a pattern of three numbers (Missing Number) were examined. In the Count Out Loud task, students were prompted to begin with the number 1 and orally count as high as they could. For the Quantity Discrimination task, students were presented with two written numbers between 0 and 20 and were asked to verbally name which had a greater value. During the Number Identification task, students were asked to orally name numbers from 0 through 100. The Missing Number task assessed students' ability to name a missing number from a set of three numbers while counting in a forward sequence beginning as low as 0 and no higher than 20. The placement of the blank in the number sequence was varied randomly.

In Grade 1, we continued to use Count Out Loud, Quantity Discrimination, Number Identification, and Missing Number tasks that were constructed in the same manner as the tasks for the kindergarten students, and added tasks that included Next Number and Number Facts. In the Next Number task, students were verbally prompted to say the number that comes next when presented with a random number between 0 and 100. First grade students were also administered the Number Facts task, where students verbally answered simple addition and subtraction facts with digits less than 10. All tasks for both kindergarten and Grade 1 were administered for 1 minute. Table 1 provides an overview of all early mathematics measures, including the number and range of items, how items were constructed, and a sample item for each measure.

Also, students' performance on each of the early numeracy tasks was correlated with the criterion variable, which was each student's scaled score on a norm-referenced achievement measure, the TEMA-3 (Ginsburg & Baroody, 2003). The TEMA-3 is designed for use with children ages 3 to 8 and consists of 72 items in both informal and formal mathematics. Within the area of informal mathematics, four areas are assessed primarily through verbal, pictorial, or concrete tasks: counting and enumeration skills, number-comparison facility, calculation skills, and understanding of concepts. In the area of formal mathematics, four areas are assessed using written and other tasks and include: numeral literacy, mastery of number facts, calculation skills, and understanding of concepts. As reported in the test manual, reliability coefficients are strong, with internal consistency reliability coefficients of .94 to .96, alternate form reliability coefficients of .93 to .97, and test-retest reliability coefficients of .82 to .93. Evidence of content validity is provided through the detailed explanation of item selection, through a quantitative item analysis, and through an analysis of differential functioning on items due to differences in student demographics.

Criterion validity correlation coefficients relating TEMA-3 performance to scores on other standardized, norm-referenced measures of mathematics achievement ranged from .54 to .55 (Key Math-R/NU, Basic Concepts [Connolly, 1998] and Woodcock Johnson III Tests of Achievement, Applied Problems [Woodcock & Mather, 2001], respectively) to .91 (Young Children's Achievement Test [YCAT; Hresko, Peak, Herron, & Bridges, 2000], Mathematics Quotient). The authors provide evidence of the measure's construct validity by demonstrating changes in mean performance by age at each of six age intervals and by showing how the various items are derived from basic research on children's mathematical thinking.

Procedures

After parental consent was obtained, participating students completed two rounds of data collection over the course of four days. In one round of data collection, the early mathematics CBM measures were administered to the student. One form of each task was individually administered, with each data collection session lasting approximately five minutes per child. In a separate round of data collection, the TEMA-3 was individually administered to students, with sessions lasting from roughly 20 to 35 minutes per child. Administration of all the tasks took place in the library at the school. Trained data collectors completed all data collection. All of the data collectors were veteran educators and trainers with experience in mathematics, early childhood education, and assessment administration.

Results

Means and standard deviations of each measure for kindergarten and Grade 1 are presented in Table 2. Kindergarten students answered approximately 3 to 33 correct in one minute on the various mathematics tasks, scoring the least correct on the Missing Number task (mean = 3.10) and the most correct on the Counting Out Loud task (mean = 32.52). The greatest dispersion of scores for kindergarten students was on the Counting Out Loud task (s.d. = 10.21), while the measure where the dispersion of student scores was least was on the Missing Number task (s.d = 3.30). Students in Grade 1 scored the least correct on the Number Facts task (mean = 4.47) and the most correct on the Counting Out Loud task (mean = 68.83). The greatest dispersion of scores for students in Grade 1 was on the Counting Out Loud task (s.d. = 15.84), with the smallest dispersion of scores on the Number Facts task (s.d. = 2.78). In all cases, students' scores on the tasks more than doubled when administered to students in Grade 1 as compared to kindergarten students. Tests of skewness and kurtosis were conducted and distributions met the assumptions for use of Pearson product-moment correlations.

Reliability

Measures of internal consistency from a single administration of the CBM, including Cronbach's *alpha* and split-half reliabilities, were calculated for all measures for each grade. For each measure, a high internal consistency reliability coefficient would indicate that the items are homogenous and correspond to the same content domain. *Alpha* was computed as

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_x^2}\right)$$

where k is the number of items on the task, $\sum \sigma_i^2$ is the variance of item i , and σ_x^2 is the total task variance. It is important to note that the alpha coefficient can be considered the lower bound to a theoretical reliability coefficient (Crocker & Algina, 1986).

For the split-half reliability, we divided each CBM measure into two half-tests after matching item difficulty levels, which is the proportion of students who answered the item correctly. We then scored the two half-tests separately for each student. Correlation coefficients between these two sets of scores for each measure of the CBM was calculated, and was corrected for the reliability coefficient of the full-length test by the Spearman-Brown Prophecy,

$$\rho_{XX'} = \frac{2\,\rho_{AB}}{1+\rho_{AB}}$$

where $\rho_{XX'}$ is the reliability of the full-length test and ρ_{AB} is the correlation between the half-tests (Crocker & Algina, 1986).

As shown in Table 3, each measure for both grades had high reliability coefficients (ranging from .87 to .97 for kindergarten and from .80 to .98 for Grade 1, respectively), indicating that they were above the acceptable threshold of reliability coefficient .70. In accordance with the reliabilities, we calculated the measurement precision of the CBM measures using the standard error of measurement (SE), calculated as the amount of error due to the difference between an individual's test score and his/her true score. SE was given as

$$SE = \hat{\sigma}_x \sqrt{1-\hat{\rho}_{XX'}}\ ,$$

where $\hat{\sigma}_x$ is the standard deviation for the scores for the group, and $\hat{\rho}_{XX'}$ is the estimated reliability of the each CBM measure. SE ranged from .37 to 1.42 for kindergarten and from 1.23 to 2.20 for Grade 1 (see Table 3).

Construct Validity

A confirmatory factor analysis (CFA) was performed to evaluate the construct validity of the CBM measures. The CFA determined that Early Numeracy Proficiency was measured by the Count Out Loud, Quantity Discrimination, Missing Number, and Number Identification tasks for kindergarten and by the two additional measures of Next Number and Number Facts tasks for Grade 1. Table 4 presents the results of the CFA models for both kindergarten and Grade 1.

For kindergarten, the χ^2 value was 4.828 with 2 degrees of freedom and its p-value was .0894, indicating that the hypothesized model (i.e., Early Numeracy Proficiency is measured by four CBM measures) fit the data. Other model-fit indices also agreed that the CFA model for kindergarten is a reasonably good fit: 1) standardized root mean square residual (SRMR) = .039, meaning that the model explains the correlation to within an average error of .039 (see Hu & Bentler, 1995); 2) comparative fit index (CFI) = .973, in which a cutoff value close to .95 suggests an adequate fit to the data (Hu & Bentler, 1999), and 3) the Tucker-Lewis index (TLI; Tucker & Lewis, 1973) = .920, with values close to .95 being indicative of good fit. However, the fit measured by the root mean square error of approximation (RMSEA) was not favorable for kindergarten (RMSEA = .141), where values greater than .10 indicate poor fit. The upper bound of a 90% confidence interval exceeds .10, which seems mainly related to the small sample size (n = 68), as suggested in Hu and Bentler (1999). Therefore, although the CFA results pointed out that the overall model fits the data well, the wide confidence interval of the RMSEA suggests that a larger sample may be required in order to obtain more precise results for further study.

The test of model fit for Grade 1 yielded a χ^2 value of 12.367, with 9 degrees of freedom and a probability of .1934, thereby suggesting that the fit of the data to the hypothesized model (i.e., Early Numeracy Proficiency is measured by six CBM measures) is adequate. Other model fit indices for Grade 1 were consistent in suggesting that the model represented an adequate fit to the data: SRMR = .049, CFI = .979, and TLI = .964. A value of RMSEA .076 appeared to be of mediocre fit (see MacCallum, Bowne & Sugawara, 1996). Given the confidence interval around the RMSEA value, ranging from .000 to .171 (*n* = 69), it is suggested that the model needs to be examined with a larger sample due to the fact that confidence interval can be seriously influenced by sample size (ibid).

Figure 1 displays the standardized factor loadings for kindergarten and summarizes the influence of the latent factor (i.e., Early Numeracy Proficiency) on observed variables (i.e., each CBM measure) and the error variances associated with each observed variable. Given that all measurement errors appear to be uncorrelated, all of the CBM measures in kindergarten showed relatively high standardized factor loadings and were statistically significant at a .01 level: Count Out Loud = .594, Quantity Discrimination = .686, Missing Number = .848, and Number Identification = .820. The proportions of explained variances (i.e., one minus the error variance of each indicator) for Count Out Loud, Quantity Discrimination, Missing Number, and Number Identification were 35.3%, 47.1%, 72.0%, and 67.2%, respectively. Figure 2 represents the standardized factor loadings and the error variances of the indicators for Grade 1. All of the indicators showed relatively high standardized factor loadings and were statistically significant at a .01 level: Count Out Loud = .623, Quantity Discrimination = .611, Missing Number = .880, Number Identification = .856, Next Number = .625, and Number Facts = .581. The proportions of explained variance for Count Out Loud, Quantity Discrimination, Missing Number, Number Identification, Next Number, and Number Facts tasks were 38.9%, 43.7%, 77.4%, 73.3%, 33.7%, and 39.1%, respectively.

From the CFA results shown in Table 4 and Figures 1 and 2, it can be concluded that the CBM seems to measure one underlying construct, Early Numeracy Proficiency. Since the sample size is relatively small, however, it should be examined with a larger sample to see if the results would be consistent for future study.

Concurrent Criterion Validity

To examine the criterion validity of the CBM early mathematics measures, we computed correlation coefficients between students' scores on the CBM measures at each grade and the criterion variable, students' scaled score on the TEMA-3. Also, we fit structural equation models to evaluate the strength of relationship between the Early Numeracy Proficiency measured by each of the measures of the CBM and the TEMA score, and the interrelationships among the CBM items. Results of the criterion validity analyses are presented in Table 5 and Figure 3. Note that not all error terms and residuals were shown for simplicity.

As shown in Table 5, for students in kindergarten correlations ranged from .53 to .68, with the strongest correlations between the CBM Number Identification measure and the TEMA-3 scaled score. For students in Grade 1, correlations were not as strong, with a range from .40 to .68. Once again, the strongest correlations were between CBM Number Identification and the TEMA-3 scaled score.

In Figure 3, the standardized path coefficients in the path model of kindergarten indicates that a strong relationship exists between the Early Numeracy Proficiency and each of the measures measured in the CBM, Count Out Loud (.621), Quantity Discrimination (.793), Missing Number (.827), and Number Identification (.725). The construct of Early Numeracy Proficiency also has a strong relationship to the TEMA (.815). All path coefficients were significant at .01 level.

For the path model for Grade 1, the standardized path coefficients show that the strongest relationship exists between the Early Numeracy Proficiency and Number Identification (.878).
The paths coefficients to Missing Number (.844), Next Number (.659), Quantity Discrimination (.653), and Number Facts (.599) were significant at .01 level. Also, the relationship between the Early Numeracy Proficiency and the TEMA score turned out to be strong (.748), which was significant at .01 level.

Item Analysis

We examined item difficulty and discrimination for item analysis. Since each item of the CBM was dichotomously scored, the item difficulty (p) was computed as the proportion of students who answered the item correctly. Item difficulty is relevant for determining whether students have mastered the concept being tested. When targeting all levels of math performance, the range for the average item difficulties should be .3 to .7 in order to maximize the information that the items will reveal about the students (Allen & Yen, 1979). Items are classified as "easy" if p is .7 and above and "hard" if p is .3 or below, as a measure of item difficulty. Discrimination coefficient (D) was calculated as a point biserial correlation (r_{pbis}) between scores on an item (e.g., 0 or 1) and total score on all other items on the CBM measure. D classifies items as having "good" discrimination if D is .3 and above; "fair" if D is between .20 and .29; and "marginal to poor" if D is below .20 (Crocker & Algina, 1986). Note that the Count Out Loud task was excluded from the item analysis for both grades because it was intended to measure the ability to count out loud from 1 and continue as high as the student can count in the given time (i.e., 1 minute), thus the item difficulty decreases as the number becomes larger.

Tables 6 and 7 show item analysis results for three CBM measures, Missing Number, Quantity Discrimination, and Number Identification for kindergarten, and two additional measures, Next Number and Number Facts, for Grade 1. The ability of the CBM measures to discriminate between students that have mastered particular skills and those that have not is important for determining particular skills that need to be remediated, and for use in determining which students might be struggling overall in mathematics. For kindergarten, the 10 items of Missing Number showed overall low item difficulties than those for Grade 1 and ranged between .18 and .42. The item, "9 10 11" was the most difficult item (p = .18) while the item, "10 11 12" was easiest (p = .42). The values of pbis r seem reasonably fair to good varying from .24 to .71, indicating that the items of the Missing Number task seem to have a high level of discriminating power. Items on the Missing Number task for Grade 1 all had acceptable and satisfactory item difficulties (ranged between .34 and .83) and discriminations (ranged between .30 to .81).

The 24 items in Number Identification ranged in item difficulty from .21 to .93 and in discrimination from .20 to .80. The items, "2" (p = .93), "1" (p = .80), and "0" (p = .74) showed the largest item difficulties among the items, while items "31", "84", "13", "65", "17", "46", and "71" showed item difficulties less than .3, indicating they were relatively hard items for kindergarten. Discriminating power for these items was satisfactory, except for item "2" which had a marginal discrimination of .20. For Grade 1, all item difficulties were acceptable, except for item "2" (p = 1.00) which showed every student answered the item correctly, thereby resulting in no discriminating power among students. Also, the (r_{pbis}) of "0" and "9" were .18 and .20, respectively, indicating that the items seem to have relatively marginal to poor discriminating power. The rest of the items on the Number Identification task showed acceptable and high (r_{pbis}) values varying from .39 to .87.

For Quantity Discrimination items, the item difficulty ranged from moderate to high item difficulty (ranged from .51 to .92). In addition, the items were generally high in discriminating power. Table 6 indicates that the item "8, 3" showed the smallest item difficulty of p = .51 while the item "10, 0" had the largest item difficulty of p = .92 among the items. Although the item "12, 16" had relatively lower discriminating power (r_{pbis} = .23), most of the items showed high pbis r values. Similarly, items on the Quantity Discrimination task for Grade 1 had moderate to high item difficulty (ranged between .41 and .94). However, item "8, 3" had a poor discrimination (r_{pbis} = .14), indicating it did not function as well as in kindergarten.

The 10 items of the Number Fact task had item difficulties which ranged between .19 to .72, indicating that items "7+8" (p = .19), "7+6" (p = .20), and "2+6" (p = .27) were relatively harder than other items. All items except item "7+8" (r_{pbis} = .18) had satisfactory discriminating power.

For the Next Number task, items showed moderate to high item difficulty (ranging between .48 and .95), suggesting that items "9(10)", "3(4)", and "8(9)" were relatively easy items for Grade 1. Most of the items discriminated well among students, as indicated by the discriminating power ranging from .37 to .67, although three items - "9(10)", "3(4)", and "4(5)" - had poor to marginal discrimination power, r_{pbis} = .03, r_{pbis} = .07, and r_{pbis} = .21, respectively.

Discussion/Implications

The purpose of our study was to replicate and extend existing research on reliable and valid early mathematics measures, introduce new measures, modify existing measures, and replicate results of criterion validity studies with additional students. The proposed CBM mathematics measures were examined with respect to their fit within a model of mathematical reasoning characterized by three components: Quick Retrieval, Number Sense, and Written Computation. For students in kindergarten and Grade 1, measures included Counting Out Loud, Quantity Discrimination, Missing Number, and Number Identification, with the addition of Next Number and Number Facts measures for students in Grade 1.

Curriculum-based measures that are technically adequate embody several characteristics including: not having floor or ceiling effects, so that student performance can be captured throughout the student's academic career; having large standard deviations, so that one can be confident that the measures are discriminating the performances of students at low and high ends of the spectrum; demonstrating reliability in scoring, form, and item type; demonstrating construct validity in that the proposed CBM tasks explain a large proportion of variance in the construct that is being measured; and having criterion validity, which demonstrates how the CBM tasks relate to other mathematic criteria, such as standardized test scores.

It is important to identify measures that do not demonstrate floor or ceiling effects so that students' performance can be measured early in their academic careers, and can continue to be monitored on the same measures as they move through the grades. With respect to floor effects, in kindergarten no students scored 0 on the Count Out Loud measure, 5 scored 0 on the Quantity Discrimination measure, 17 scored 0 on the Missing Number measure, and 3 scored 0 on the Number Identification task. In Grade 1, no students scored 0 on the Count Out Loud, Number Identification, Quantity Discrimination, or Next Number measures, while 5 students scored 0 on Number Facts, and 3 scored 0 on the Missing Number task. In addition, students in both kindergarten and Grade 1 were not approaching a ceiling on any of the tasks. It appears that for kindergarten students, the Count Out Loud, Quantity Discrimination, and Number Identification tasks produced performance that did not approach a floor or ceiling, and for students in Grade 1, all tasks were acceptable, indicating that these measures appear to capture students' performance from kindergarten through Grade 1.

However, measures did vary in difficulty, with the lowest student means on Missing Number at 3.10 and 8.97 for kindergarten and Grade 1 students, respectively. The highest student means were for the Counting Out Loud measure, with kindergarten students receiving a mean of 32.52 and Grade 1 students receiving a mean score of 68.83. Because there is no ceiling on the Counting Out Loud measure, students could continue growing, although one might question the validity of using such a measure long term, particularly when students are learning more complex mathematics skills. The greatest standard deviations were seen in the Quantity Discrimination, Number Identification, and Counting Out Loud tasks for both kindergarten and Grade 1, indicating that these measures have the potential to discriminate among students at a variety of levels.

With respect to reliability, all measures for each grade had high reliability coefficients, above the acceptable threshold of .70. For purposes of educational decision making, however, we expect that the reliabilities will be at least .80, and all measures were above this level, with coefficients for kindergarten measures ranging from .87 (Missing Number) to .97 (Count Out Loud) on *alpha* reliability, with similar results for split-half reliability. For students in Grade 1, *alpha* reliabilities ranged from .80 on Number Facts to .98 on Count Out Loud, with split-half reliabilities mirroring these results. The standard error of measurement ranged from .37 to 1.42 for kindergarten and from 1.23 to 2.20 for Grade 1, indicating that the amount of error due to the difference between an individual's test score and his/her true score was slight. In the future, greater strength of reliability could be assumed if test-retest and alternate-form reliability tests were conducted.

A confirmatory factor analysis (CFA) was performed to evaluate the construct validity of the CBM mathematics measures. The results of the CFA determined that Early Numeracy Proficiency was measured by Count Out Loud, Quantity Discrimination, Missing Number, and Number Identification for kindergarten, and by two additional measures of Next Number and Number Facts for Grade 1. As shown in Figures 1 and 2, we were confident that the measures we were utilizing to measure students' mathematics performance were in fact indicative of the construct of Early Numeracy Proficiency for both kindergarten and Grade 1.

With respect to concurrent criterion validity, correlations were calculated between the students' scaled score on a standardized test of early mathematics, the Test of Early Mathematics Achieve-ment-3 (TEMA-3; Ginsburg & Baroody, 2003), and students' scores on the CBM measures, with all correlations significant at $p \leq .01$. For students in kindergarten, correlations between the TEMA and the CBM measures ranged from .53 to .68, with moderate correlations for Number Identification (.68) and Quantity Discrimination (.64). For students in Grade 1, correlations ranged from .40 (Counting Out Loud) to .68 (Number Identification), with moderate correlations for Number Facts (.50), Missing Number (.56), and Next Number (.59).

An item analysis was conducted to determine whether certain items were disproportionately easy or difficult for students. In a CBM system, it should not be item variance that is causing fluctuations in student performance, but changes in student achievement. Detailed statistics were shown in Tables 6 and 7, and item-by item inspection using item difficulty and discrimination indicated most of items on the CBM measures have adequate levels of item difficulties and satisfactory discrimination for each grade. Although there were a few items (e.g., "2" and "0" in Number Identification and "what comes next?" in Next Number for Grade 1) that were easy for the participants, we are searching for measures that serve as overall indicators of proficiency in early mathematics, and can discriminate among students within and across grades. As you would not exclude items that are important and valuable for the content from a criterion-referenced test because they are too easy (Wieberg, 2004), you would also not exclude items from a CBM measure that are contributing to the overall content unless these items are clearly too difficult for all students (i.e., all students score 0 on this item each time it is administered).

Overall, it appears that for both kindergarten and first grade students, the measures utilized in this study did fit the model for Early Numeracy Proficiency, based on results from the CFA. For kindergarten students, the Counting Out Loud, Quantity Discrimination, and Number Identification tasks appeared to have the greatest technical adequacy, with few floor and no ceiling effects, ability to discriminate among students at different levels, strong reliabilities, and moderate criterion validity with the TEMA-3. For Grade 1 students, all 6 measures (Counting Out Loud, Quantity Discrimination, Number Identification, Missing Number, Next Number, and Number Facts) produced scores that did not approach floor or ceiling limits, while the greatest standard deviations were observed for Quantity Discrimination, Number Identification, and Counting Out Loud. Moderately strong to strong reliabilities were observed for all Grade 1 tasks. With respect to concurrent criterion validity, moderate correlations were noted for Number Identification (.68) and Next Number (.59), indicating that these two measures might be related the most strongly with another form of mathematics assessment that is standardized.

The results of this study give us further confidence that these early mathematics measures can serve as good screening tools for students in early elementary school. However, future research should examine the utility of using these promising measures to demonstrate growth over time. Additionally, further examination with a greater number of students, across several settings and years, is necessary to validate that the measures are effective for the purposes of screening and progress monitoring in early mathematics. This research is promising as we look toward identifying evidence-based systems of screening and progress monitoring in mathematics, intervening early, and improving instruction for all students.

References

Allen, M. J. & Yen, W. M. (1979). Introduction to Measurement Theory. Monterey: Brooks/Cole.

Ansari, D. & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: a neuroconstructivist perspective. Trends in Cognitive Sciences, 6(12), 511-516.

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. Journal of Learning Disabilities, 38(4), 333-339.

Bergan, J. R., Sladeczek, I. E., Schwarz, R. D., & Smith, A. N. (1991). Effects of a measurement and planning system on kindergarteners' cognitive development and educational programming. American Educational Research Journal, 28(3), 683- 714.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. Phi Delta Kappan, 80(2), 139-148.

Bowman, B. T., Donovan, M. S., & Burns, M. S. (Eds.). (2001). Eager to learn: Educating our preschoolers. Washington, DC: National Academy Press.

Brigance, A. (1985). Brigance Preschool Screen. North Billerica, MA: Curriculum Associates. Brigance, A. (1999). Comprehensive inventory of basic skills (rev. ed.) North Billerica, MA: Curriculum Associates, Inc.

Campbell, F. A. & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. Child Development, 65(2), pp. 684-698.

Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. Assessment for Effective Intervention, 30(2), 3-14.

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.

Connolly, A. J. (1998). Key Math-revised, normative update. American Guidance Service, Inc.: Circle Pines, MN.

Crocker, L. & Algina, J. (1986). Introduction to classical modern test theory. Belmont, CA: Wadsworth Group.

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

Fuchs, L. S. et al. (1994). Strengthening the connection between assessment and instruction with expert systems. Exceptional Children, 61(2), 138-146.

Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A metaanalysis. Exceptional Children, 53(3), 199-208.

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

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989b). Effects of alternative goal structures within curriculum-based measurement. Exceptional Children, 55(5), 429-438.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. Experimental Child Psychology, 88, 121 - 151.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for mathematical disability. Journal of Experimental Child Psychology, 74, 213-231.

Gersten, R. & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. The Journal of Special Education, 44, 18- 28.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. Journal of Learning Disabilities, 38(4), 293-304.

Ginsburg, H. P. & Baroody, A. J. (1990). Test of Early Mathematics Ability, 2nd edition. Austin, TX: Pro-Ed, Incorporated.

Ginsburg, H. P. & Baroody, A. J. (2003). Test of Early Mathematics Ability, 3rd edition. Austin, TX: Pro-Ed, Incorporated.

Hresko, W., Peak, P., Herron, S., & Bridges, D. (2000). Young Children's Achievement Test. Austin, TX: Pro-Ed, Incorporated.

Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal, 6, 1-55.

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), Structural Equation Modeling: Concepts, issues, and Applications (pp. 76-99). Thousand Oaks, CA: Sage.

Jordan, N. C., Hanich, L. B., & Uberti, H. Z. (2003). Mathematical thinking and learning difficulties. In A. J. Baroody & A. Dowker (Eds.), The development of arithmetic concepts and skills: Constructing adaptive expertise (pp. 359-383). Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Lembke, E. & Foegen, A. (2005, February). Monitoring student progress in early math. Paper presented at the 14th annual meeting of the Pacific Coast Research Conference, San Diego.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. Psychological Methods, 1, 130-149.

Marston, D. (1989). Curriculum based measurement: What it is and why do it? In M. R. Shinn (Ed.), Curriculum-based measurement: Assessing special children (pp. 18-78). New York: Guilford.

National Association for the Education of Young Children and National Council of Teachers of Mathematics. (2002). Position statement. Early childhood mathematics: Promoting good beginnings., from http://www.naeyc.org/about/positions/psmath.asp

National Council of Teachers of Mathematics (1989). Curriculum and evaluation NCTM Standards for school mathematics. Reston, VA: The National Council of Teachers of Mathematics, Inc.

Okamoto, Y. & Case, R. (1996). Exploring the microstructure of children's central conceptual structures in the domain of number. In R. Case & Y. Okamoto (Eds.), The role of central conceptual structures in the development of children's thought: Monographs of the Society for Research in Child Development (Vol. 1-2, pp. 27-58). Malden, MA: Blackwell Publishers.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment. Washington, DC: National Academy Press.

Russell, R. L. & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematics difficulties. Cognition and Instruction, 1, 217-244.

Speece, D. L., Molloy, D. E., & Case, L. P. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. Learning Disabilities Research & Practice, 18, 145-156.

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

Tucker, L. R. & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38, 1-10.

van der Heyden, A. M., Broussard, C., Fabre, M., Stanley, J., Legendere, J., & Creppell, R. (2004). Development and validation of curriculum-based measures of math performance for preschool children. Journal of Early Intervention, 27(1), 27-41.

van der Heyden, A. M., Witt, J. C., Naquin, G., Noell, G. (2001). The reliability and validity of curriculum-based measurement readiness probes for kindergarten students. School Psychology Review, 30, 363-382.

Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. Journal of Educational Psychology, 88(4), 601-638.

Wiberg, M. (2004). Classical test theory vs. Item response theory: An evaluation of the theory test in the Swedish driving-license test (No. 50). Umeå Universite.

Woodcock, R. M. & Johnson, M. B. (1989). Woodcock-Johnson psycho-educational batteryrevised. Allen, TX. DLM Teaching Resources.

Woodcock, R. W. & Mather, N. (2001). Woodcock-Johnson III tests of achievement, examiner's manual. Itasca, IL: Riverside Publishing.

Wright, B. (1991). What number knowledge is possessed by children beginning the kindergarten year of school? Mathematics Education Research Journal, 3(1), 1-16.

Table 1. Early Mathematics Measure Descriptives and Exemplary Items

	Count Out Loud	Quantity Discrimination	Number Identification	Missing Number	Next Number	Number Facts
Total number of items Range	120 1-120	63 0-20	84 0-100	54 0-20	84 0-100	60 Digits in problems ≤10
Construction	Numbers placed in order from 1 through 120.	Generate number sets between 0 –10 and 0 – 20 randomly (67% between 0 –10 and 33% between 0 – 20). Alternate the order by placing the larger number first or second.	Randomly placed numbers on a page with 50% between 0 – 20, 30% between 21 – 50, and 20% between 51 – 100.	2 numbers and a blank are presented in a forward counting sequence by 1's, with the blank varying.	Randomly presented numbers (50% between 0 – 20, 30% between 21 – 50, and 20% between 51 –100).	Randomly presented addition and subtraction problems with digits less than ten.
Sample Item	1…2…3	3 14	7 50	2, 3, __	Oral task ("What comes next?")	Oral task: "Four plus two?"

Table 2. Descriptive Statistics

	Kindergarten		Grade 1	
	Mean	St. Dev.	Mean	St. Dev.
Count Out Loud	32.52	5.35	68.83	15.84
Quantity Discrimination	11.27.	5.35	22.78	7.69
Missing Number	3.10	3.30	8.97	4.89
Number Identification	8.62	7.30	23.53	10.36
Next Number	-	-	12.02	3.59
Number Facts	-	-	4.47	2.78

Table 3. Alpha and Split-Half Reliabilities, and Standard Error (SE)

	Count Out Loud		Quantity Discrimination		Number Identification		Missing Number		Next Number		Number Facts	
	K	1	K	1	K	1	K	1	K	1	K	1
Alpha	.97	.98	..88	.94	.87	.92	.95	.96	-	.84	-	.80
Split-Half	-	-	.87	.96	.88	.92	.97	.98	-	.85	-	.84
SE	.37	2.20	1.40	1.84	1.07	1.39	1.42	2.01	-	1.42	-	1.23

Table 4. Results of CFA for Grades K and 1

Test of Model Fit		Grade K	Grade 1
Chi-Square Test of Model Fit	Value	4.828	12.367
	Degrees of Freedom	2	9
	P-value	.0894	.1934
CFI/TLI	CFI	.973	.979
	TLI	.920	.964
RMSEA (Root Mean Square Error of Approximation)	Estimate	.141	22.78
	90% C.I.	.000 .307	.000 .171
	Probability RMSEA ≤ .05	.128	.300
SRMR (Standardized Root Mean Square Residual)	Value	.039	.049
R-SQUARE		R-Square	R-Square
	Count Out Loud	.353	.389
	Quantity Discrimination	.471	.437
	Missing Number	.720	.774
	Number Identification	.672	.733
	Next Number	-	.337
	Number Facts	-	.391

Table 5. Concurrent Criterion Validity Coefficients

	Count Out Loud		Quantity Discrimination		Number Identification		Missing Number		Next Number		Number Facts	
	K	1	K	1	K	1	K	1	K	1	K	1
TEMA-3	.53	.40	.64	.48	.59	.56	.68	.68	-	.59	-	.50

Note that all correlations were significant at the .01 level.

Table 6. Item Difficulty and Discrimination for the CBM Measure for Kindergarten

CBM Measure	Item	p	r_{pbis}	CBM Measure	Item	p	r_{pbis}
Missing Number	2 3 4	.39	.24	Number Identification	2	.93	.20
	8 9 10	.41	.29		10	.57	.56
	9 10 11	.18	.59		31	.23	.55
	10 11 12	.42	.64		15	.33	.70
	4 5 6	.25	.70		20	.36	.66
	12 13 14	.36	.67		84	.28	.67
	7 8 9	.35	.60		13	.27	.64
	15 16 17	.27	.71		0	.74	.22
	16 17 18	.26	.66		9	.54	.47
	5 6 7	.37	.70		40	.33	.62
Quantity Discrimination	8, 3	.51	.34		65	.21	.67
	11, 17	.55	.44		17	.26	.67
	9, 5	.70	.57		46	.24	.62
	2, 7	.64	.57		5	.71	.25
	13, 9	.81	.61		32	.36	.56
	6, 8	.70	.57		71	.28	.67
	7, 4	.78	.48		14	.31	.80
	12, 16	.57	.23		96	.31	.75
	6, 4	.73	.46		1	.80	.27
	1, 3	.82	.48		43	.40	.72
	15, 2	.84	.46		57	.35	.68
	0, 8	.71	.54		4	.68	.38
	10, 0	.92	.41		37	.30	.72
	4, 16	.76	.48		100	.58	.51
	8, 5	.79	.42				
	3, 4	.72	.58				
	18, 7	.79	.57				
	6, 9	.85	44				

Table 7. Item Difficulty and Item Discrimination for the CBM Measure for Grade 1

CBM Measure	Item	*p*	r_{pbis}	CBM Measure	Item	*p*	r_{pbis}
Missing Number	4 5 6	.55	.30	Quantity Discrimination	8, 3	.59	.14
	12 13 14	.73	.40		11, 17	.78	.41
	7 8 9	.61	.64		9, 5	.89	.34
	15 16 17	.84	.51		2, 7	.89	.59
	16 17 18	.61	.66		13, 9	.94	.38
	5 6 7	.53	.51		6, 8	.83	.63
	3 4 5	.63	.71		7, 4	.91	.40
	18 19 20	.70	.62		12, 16	.70	.30
	11 12 13	.63	.71		6, 4	.94	.26
	1 2 3	.69	.67		1, 3	.89	.46
	17 18 19	.69	.81		15, 2	.91	.49
	14 15 16	.44	.70		0, 8	.91	.60
	6 7 8	.47	.72		10, 0	.91	.64
	13 14 15	.34	.69		4, 16	.88	.72
Number Identification	2	1.00	.00		8, 5	.80	.32
	10	.94	.39		3, 4	.77	.63
	31	.55	.70		18, 7	.86	.61
	15	.75	.51		6, 9	.78	.73
	20	.81	.60		1, 0	.86	.73
	84	.67	.78		14, 19	.70	.67
	13	.56	.63		5, 2	.73	.72
	0	.91	.18		3, 5	.66	.78
	9	.86	.20		15, 4	.69	.80
	40	.73	.43		4, 8	.61	.77
	65	.70	.59		4, 0	.63	.76
	17	.59	.45		3, 12	.53	.75
	46	.72	.69		7, 1	.53	.70
	5	.84	.46		3, 8	.42	.59
	32	.69	.81		17, 8	.41	.56
	71	.66	.65		4, 9	.44	.62
	14	.58	.58		7, 5	.42	.58
	96	.69	.84				

Number Identification (continued)	1	.84	.59	**Number Facts**	3+1	.72	.35
	43	.69	.76		7+8	19	.18
	57	.67	.82		2+4	.56	.46
	4	.75	.45		5+4	.59	.56
	37	.61	.87		9+4	.38	.64
	100	.80	.49		8+0	.72	.43
	49	.66	.79		7+6	.20	.43
	12	.63	.62		3+7	.34	.66
	34	.64	.83		3+3	.50	.48
					2+6	.27	.58
	6	.70	.71	**Next Number**	9 (10)	.95	.03
	3	.63	.61		21 (22)	.83	.37
	16	.48	.82		54 (55)	.75	.51
	41	.44	.81		3 (4)	.92	.07
	50	.50	.73		73 (74)	.63	.38
	67	.47	.80		18 (19)	.78	.55
	44	.44	.66		4 (5)	.94	.21
	19	.34	.72		67 (68)	.80	.57
					8 (9)	.92	.52
					14 (15)	.81	.55
					30 (31)	.73	.52
					15 (16)	.63	.48
					36 (37)	.72	.67
					0 (1)	.58	.60
					27 (28)	.55	.62
					91 (92)	.48	.45

Figure 1. Standardized Factor Loadings and the Error Variances of the Indicators for Kindergarten

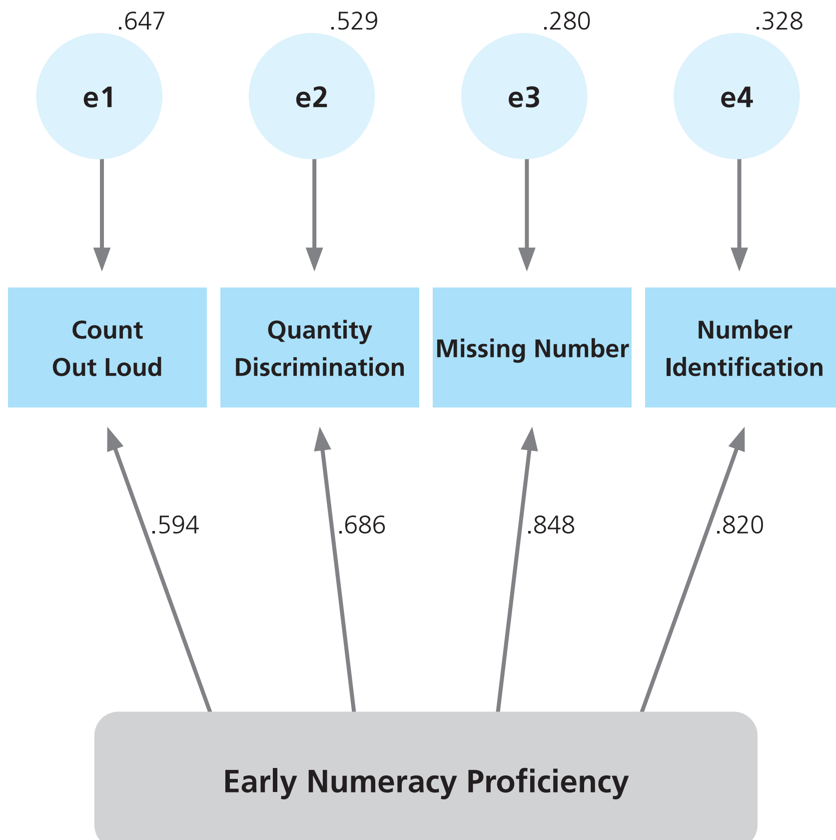


Figure 2. Standardized Factor Loadings and the Error Variances of the Indicators for Grade 1

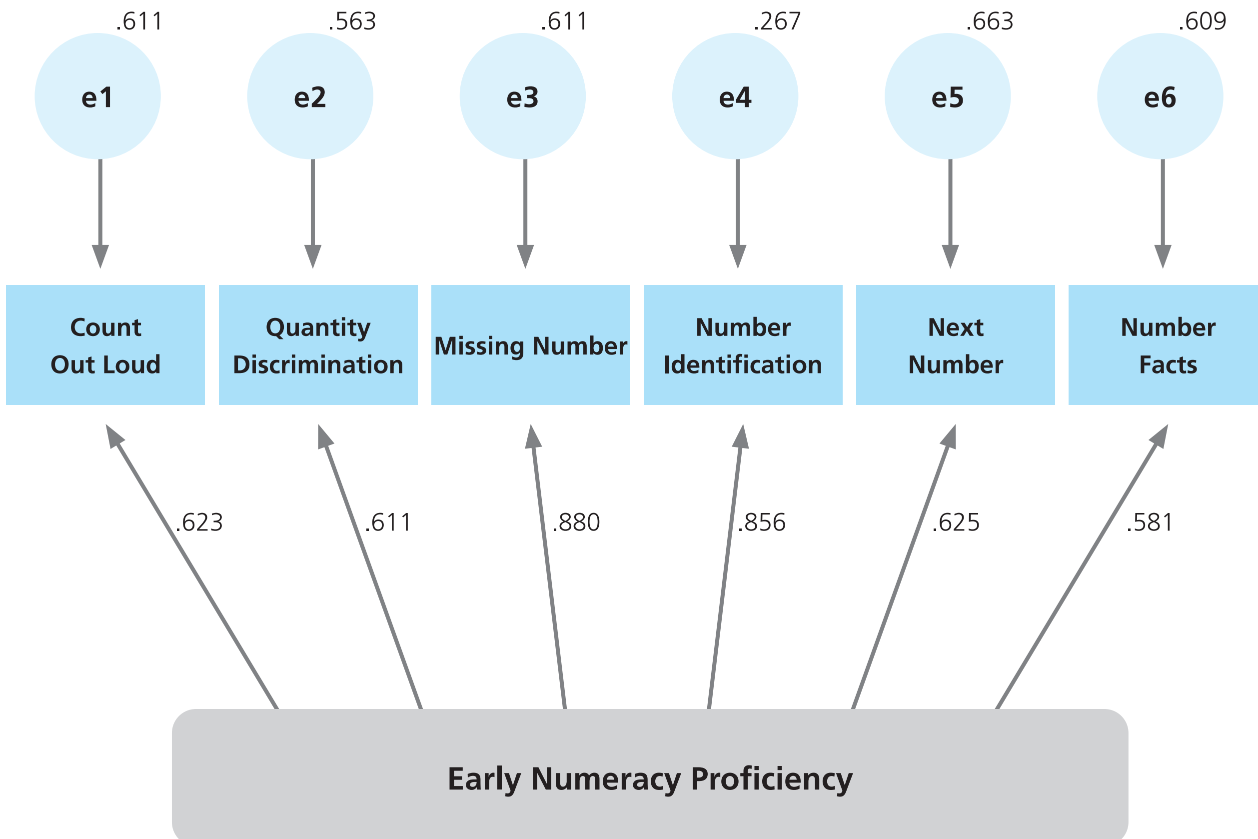
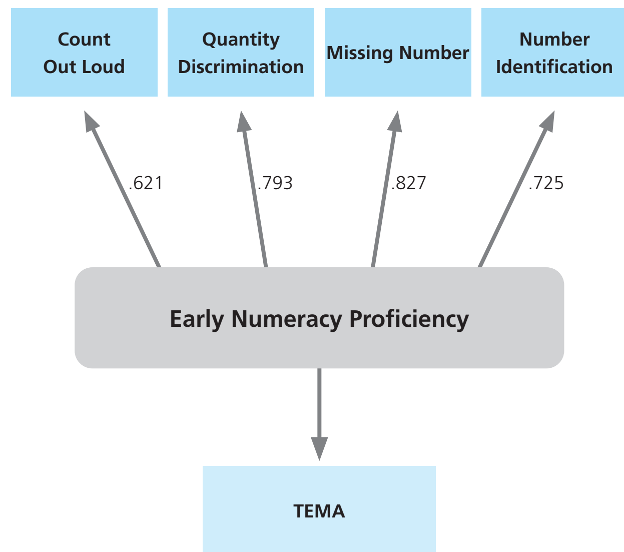
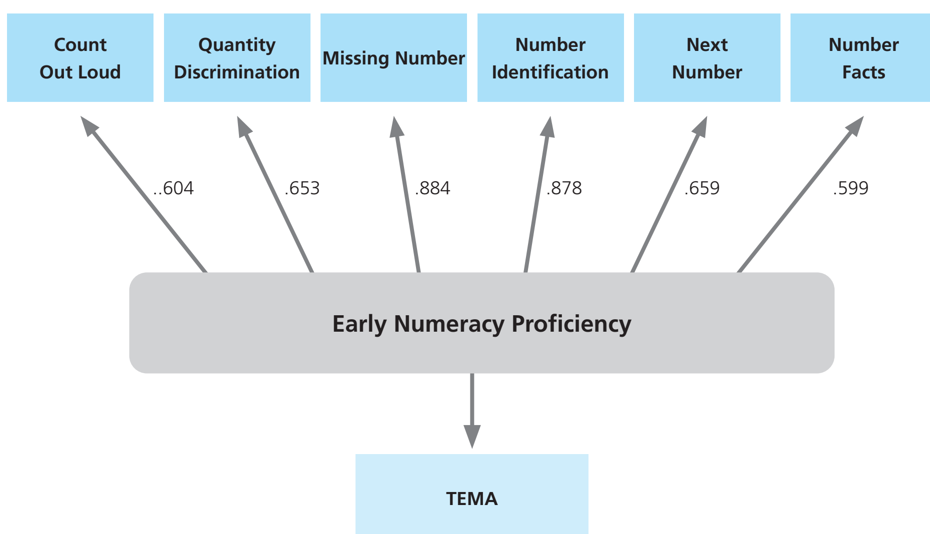


Figure 3. Path Models for Concurrent Validity between the CBM Measures and the TEMA Score



Kindergarten Model



Grade 1 Model

For more information, please visit www.wirelessgeneration.com.

mCLASS®:Math Identifying Technically Adequate Early Mathematics Measures 01/08

Submitted to Journal of Educational Psychology, May 2007

Corporate: 55 Washington Street, Suite 900, Brooklyn, NY 11201-1071 • (212) 796-2200
Inquiries: (866) 212-8688 • www.wirelessgeneration.com