# TECHNICAL REPORT #26:

## Examining the Long-Term Predictive Validity of the Early Numeracy Indicators for Predicting Success on a High Stakes Mathematics Test

*Subhalakshmi Singamaneni, Anne Foegen, and Jeannette Olson*

**RIPM Years 2, 3 and 4: 2004 – 2007**

**Dates of Study: September 2008 – May 2009**

**July 2009**

The College of Education & Human Development

University of Minnesota

IDEAs that Work

U.S. Office of Special Education Programs

Produced by the Research Institute on Progress Monitoring (RIPM) (Grant # H324H30003) awarded to the Institute on Community Integration (UCEDD) in collaboration with the Department of Educational Psychology, College of Education and Human Development, at the University of Minnesota, by the Office of Special Education Programs. See progressmonitoring.net.

Abstract

The present study examines the long-term predictive validity of three Early Numeracy Indicators (Number Identification, Quantity Discrimination, and Missing Number). Data were drawn from 192 students across three cohorts in the 2004-05 to 2006-07 academic years. The Early Numeracy Indicator data for this study were collected as part of three studies in the years 2004-2007 (Foegen, Lembke, Klein, Lind, & Jiban, 2006; Lembke & Foegen 2005; Lind, Olson, & Foegen, 2009) to identify reliable and valid curriculum-based measures in early mathematics. Third-grade ITBS data for the present study were provided by the district. Short-term criterion measures within those studies were used to evaluate the within-year predictive validity of the measures. The Iowa Test of Basic Skills was used as long-term criterion measure to evaluate the multi-year (across 2 and 3 years) predictive validity of the early numeracy measures. Correlations between the Early Numeracy Indicators in Kindergarten and Grade 1 and Grade 3 ITBS scores were in the moderate range. We used logistic regression and diagnostic accuracy statistics to examine the extent to which the Early Numeracy Indicators predicted success or failure on the ITBS. All the Early Numeracy Indicators produced results that indicated that they are effective in predicting success or failure on the ITBS.

Examining the Long-Term Predictive Validity of the Early Numeracy Indicators for Predicting Success on a High Stakes Mathematics Test

With federal initiatives like the No Child Left Behind Act requiring higher levels of accountability, educators are required to show they are meeting their Annual Yearly Progress goals. For this purpose, teachers today need tools that allow them to predict their students' performance on high stakes tests used to measure the district's performance on annual report cards. Ideal tools for this purpose are those which least disturb the flow of academic programs in the school, and which are already embedded in school activities.

Progress monitoring is an assessment process that is conducted frequently to estimate rates of improvement in students' academic performance and identify those who are not performing at expected levels (Fuchs & Fuchs, n.d.). Progress monitoring also helps in evaluating the efficacy of an instructional method by comparing the performance of students with respect to the form or method of instruction (Fuchs & Fuchs, n.d.). Progress monitoring tools that are technically valid help in understanding the academic standing of students. One of the popular forms of progress monitoring is Curriculum Based Measurement (Deno, 1985). CBM can be used to increase achievement (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Stecker, Fuchs, & Fuchs, 2005) and to monitor long- or short-range achievement goals based on student performance (Fuchs & Fuchs, 1998).

With increasing numbers of schools relying on progress monitoring tools to evaluate student status and progress, it is important to determine whether the Early Numeracy Indicators can be used to determine whether students are performing at the expected level of academic competency. Such results will enable teachers to determine if students have sufficient competency to anticipate success on state tests in upcoming years or whether students are demonstrating performance levels that will prompt changes required to attain necessary levels of competence. The following salient features, gleaned from the literature, demonstrate why CBM progress monitoring tools are an ideal choice as predictors for high stakes tests:

1. CBMs are brief.
2. CBMs can be administered frequently, thus offering multiple samples of student performance to warn in advance about potential difficulties with future high stakes tests.
3. CBMs are general outcome measures (GOMs) that represent overall understanding of a subject, rather than competence in specific or limited topics.
4. CBMs are sensitive to growth and can be used to monitor the effectiveness of intervention programs for individual students.

Predicting success or failure on high stakes tests from performance on CBMs could help educators make instructional decisions well in advance to increase rates of success on state tests and meet Annual Yearly Progress (AYP) goals. There are very few studies that have specifically investigated the long-term predictive power of CBMs to predict performance on state mandated tests and fewer on math CBMs. Given the current emphasis on accountability, it is imperative that teachers identify potential at- risk students early on to provide adequate educational intervention. Generally, state-mandated tests are administered to students starting at Grade 3. It would be of immense importance to educators if they could identify students who are potential candidates for failure in state tests as early as in Kindergarten or Grade 1 so they have sufficient time to implement instructional changes or educational interventions.

The Early Numeracy Indicators (ENIs) namely, Number Identification, Missing Number and Quantity Discrimination, are curriculum-based measures in early math used with Kindergarten and Grade 1 students. Previous studies (Foegen et al., 2006; Lembke & Foegen 2005; Lembke, et al., 2006) have investigated the technical adequacy of these Early Numeracy Indicators (ENIs). These studies indicated solid technical adequacy for the three ENIs, including alternate-form reliability, concurrent and within-year predictive validity, and sensitivity to growth. It is yet to be established how well the three ENIs predict performance on high stakes tests administered two or three years later. This study investigated the predictive validity of the three ENIs administered in Kindergarten and Grade 1, in relation to performance on the Iowa Test of Basic Skills (ITBS) administered in Grade 3. The data for this study were collected from a school district in Iowa where ITBS is being used as the state accountability measure. The following research questions were used to guide this study:

1. To what extent do the ENIs reflect student progress in early numeracy within a grade and across successive grades?
2. To what extent do the ENIs demonstrate short- and long-term predictive validity?
3. How well do the ENIs predict success/failure on the ITBS?

## Method

The primary purpose of this study is to investigate the long-term predictive validity of three Early Numeracy Indicators (Number Identification, Quantity Discrimination, and Missing Number). More specifically, we examined whether students' Kindergarten and Grade 1 scores on the Early Numeracy Indicators effectively predicted their Grade 3 performance on the Iowa Test of Basic Skills (ITBS). The Early Numeracy Indicator data for these analyses were collected in the years 2004-2007 by Foegen and her colleagues (Foegen et al., 2006; Lembke & Foegen 2005; Lind, et al., 2009) for three studies to identify reliable and valid curriculum-based measures in early mathematics. Grade 3 ITBS data for the present study were provided by the district. We examined the degree to which the ENIs administered in Kindergarten and Grade 1 predicted proficiency on the Grade 3 achievement test (ITBS). We also examined the degree to which the ENIs predicted achievement outcomes within a school year on measures of mathematics proficiency (individual standardized tests of mathematics achievement and a first grade computation curriculum-based measure).

*Participants*

The participants in this study were students from a suburban community in Iowa. The district had four schools: one preschool to third grade elementary school, one fourth and fifth grade elementary school, one middle school with grades six through eight, and one high school. For the years during which this study was conducted, the total enrollment for the district was approximately 1,400 students, with about 53% being male and about 93% white, 3.5 % Hispanic, and 3% other ethnicities. Approximately 42% of the students qualified for free and reduced lunch, and about 17% of students were receiving special education services. All four Kindergarten and four Grade 1 teachers in the building were invited and consented to participate in the study during the years in which early numeracy data were collected. For the present study, we examined long-term predictive validity using the data of 192 students from the previous research studies (Foegen et al., 2006; Lembke & Foegen, 2005; Lind et al., 2009) who had both ENI scores in Kindergarten and/or Grade1 and Grade 3 ITBS results. The demographic characteristics of the sample, described by cohort, are presented in Table 1.

Table 1

*Demographic Characteristics of the Three Cohorts of Student Participants*

| | Cohort 1 (entered K in 2003) *N* = 67 | Cohort 2 (entered K in 2004) *N* = 62 | Cohort 3 (entered K in 2005) *N* = 73 |
|---|---|---|---|
| Gender: | | | |
| Male | 34 | 33 | 45 |
| Female | 33 | 29 | 28 |
| Ethnicity: | | | |
| White | 65 | 58 | 63 |
| Hispanic | 1 | 4 | 7 |
| Other Ethnicities | 1 | 0 | 0 |
| Free and reduced lunch | 28 | 27 | 34 |
| Students with IEPs | 6 | 7 | 8 |
| English language learners | 0 | 3 | 5 |

*Measures*

*Early Numeracy Indicators.* Across the three years of data collection, three Early Numeracy Indicators were used: Number Identification, Quantity Discrimination, and Missing Number. Each measure was individually administered for one minute, with students responding orally to the items. For the Number Identification task, students were presented with randomly ordered numbers between 1 and 100, each in a separate box, and asked to name each number. For the Quantity Discrimination task, students were presented with two numbers and asked to name the bigger of two numbers. For the Missing Number task, students were presented with boxes with a series of four numbers, one of which had been replaced with a blank. Readers should note the ENIs were modified slightly between the 2004-2005 and the 2005-2006 school years. Additional details about the construction procedures and the measures are reported in (Foegen et al., 2006; Lembke & Foegen 2005).

*Short-term criterion measures.* At the time of the original studies, data were gathered in the same academic year on several criterion measures to evaluate the concurrent and short-term (within year) predictive validity of the Early Numeracy Indicators. These measures included two individually administered norm-referenced tests of mathematics achievement: the Woodcock-McGrew-Werder Mini Battery of Achievement (MBA; Woodcock, McGrew, & Werder, 1994) and the Test of Early Mathematics Ability-Third Edition (TEMA-3; Ginsburg & Baroody, 2003). We also administered a computation curriculum-based measure, the Monitoring Basic Skills Progress – Computation Grade 1 (MBSP-Comp; Fuchs, Hamlett, & Fuchs, 1998); two forms were group administered to first grade students. The final short-term criterion measure (also used only in Grade 1) was an end-of-course assessment for the district's curriculum program (*Growing with Mathematics*; Irons, 2003). The assessment included 59 problems that addressed topics such as numeration, basic computation, money, time, measurement, geometry, and problem solving. Teachers had administered the assessment under untimed conditions as part of their regular classroom activities at the conclusion of the school year. More detailed information about the characteristics and technical adequacy of the concurrent criterion measures for the studies are reported in the earlier technical reports (Foegen et al., 2006; Lembke & Foegen, 2005).

The short-term criterion measures produced several different types of scores that were used in the present study. The Calculation subtest and Reasoning and Concepts subtest of the Mini Battery of Achievement (Woodcock et al., 1994) generated a Broad Math score for each student. The TEMA-3 produced a Math Ability Score. Both of these scores were norm-referenced standard scores (*M* = 100, *SD* = 15). For the MBSP-Comp, we used the mean total number of digits correct across two forms of the measure. The end of course test was scored as a percent correct of the total number of items.

*Long-term criterion measure.* In Iowa, the Iowa Test of Basic Skills (ITBS; Hoover, Dunbar, & Frisbie, 2001) is the primary achievement measure used in the elementary grades for evaluating educational progress, including accountability for Annual Yearly Progress rates for compliance with the No Child Left Behind Act. As a result, it is important to examine whether the Early Numeracy Indicators, administered in Kindergarten and Grade 1 can effectively predict subsequent mathematics performance on the Grade 3 ITBS. ITBS is a group administered achievement test battery. It is a norm-referenced test that provides a comprehensive assessment of student progress in major content areas (Hoover, Dunbar, & Frisbie, 2001). The ITBS mathematics component consists of three subtests: Mathematics Concepts and Estimation, Mathematics Problem Solving and Data Interpretation, and an optional Mathematics Computation subtest. The Mathematics Total score is the average of the required subtests and is reported in standard scores and percentile ranks. We used the Mathematics Total percentile ranks for our analysis.

*Procedures*

The ENI and short term criterion data used for this report were collected from students in Kindergarten and Grade 1 from 2004 to 2007. Long-term criterion (ITBS) data were collected as these students progressed to Grade 3 in subsequent years. Table 2 lists the measures that were administered to students in each of the three cohorts. The rows in this table represent the three cohorts of students: those in Grade 1 in 2004-05 (Cohort 1) and those in Kindergarten in the same year (Cohort 2), as well as the students who were in Kindergarten in 2005-2006 (Cohort 3).

The columns represent the grades the students were in and the data gathered in each of those grades for each of the three cohorts. In 2004-2005, when cohort 1 was in Grade 1 and cohort 2 was in Kindergarten, data were collected only in the winter.

Table 2

*Description of Measures Administered by Year for Cohorts 1 and 2*

| | Kindergarten | | Grade 1 | | Grade 3 |
|---|---|---|---|---|---|
| | ENIs | Criterion measures | ENIs | Criterion measures | Criterion measures |
| Cohort 1 (entered K in 2003) | | | QD, MN | MBA | ITBS |
| Cohort 2 (entered K in 2004) | QD, MN | MBA | QD, MN | TEMA-3<br>MBSP-Comp<br>EOC | ITBS |
| Cohort 3 (entered K in 2005) | QD, MN | TEMA-3 | NI, QD, MN | | ITBS |

*Note.* ENIs = Early Numeracy Indicators, EOC=End of course test, ITBS = the Iowa Test of Basic Skills, MBA = Mini Battery of Achievement, MBSP-Comp = Monitoring Basic Skills Progress – Computation, MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination, TEMA-3 = Test of Early Mathematics Ability, 3rd Edition.

*Research Question 1: To what extent do the ENIs reflect student progress in early numeracy within a grade and across successive grades?*

Descriptive statistics for the ENIs that were administered are presented in Table 3. The means of the ENIs show improvement in students' performance within grades as they progressed from fall to winter to spring. A comparison of mean scores across grades (from Kindergarten to Grade 1) also shows the same result. Though all the three cohorts show similar results, Cohort 3, which has all the data for QD and MN for both Kindergarten and Grade 1 is the best example to illustrate these results. The means of both QD and MN for this cohort indicate progressive growth of students' performance from the fall of Kindergarten (mean of QD = 12.58, MN = 6.64) to the spring of Grade 1 (mean of QD = 35.99, MN = 19.67).

This pattern indicates that the ENIs reflect student progress in early numeracy within the same grade and across successive grades. However, we observed that the rate of increase in the scores from fall to winter is higher than that from winter to spring.

*Research Question 2: To what extent do the ENIs demonstrate long- and short-term predictive validity?*

To examine the long-term predictive validity of the Early Numeracy Indicators, we first computed correlations between the ENIs and Grade 3 ITBS Math scores (see Table 4). Readers should note that during the 2004-2005 school year, ENI data were gathered only in the winter. As a result, the coefficients for Cohort 1 in Grade 1 and Cohort 2 in Kindergarten are based on winter ENI data. The rest of the scores on the ENIs were taken from the fall administrations. Both QD ($r$ = .49 to .69) and MN ($r$ = .48 to .62) demonstrated moderate relations with the Grade 3 ITBS Math scores across all three cohorts. It is interesting to note that the relationships were comparatively stronger across a period of three years (performance on ENIs in Kindergarten correlated with ITBS scores in Grade 3) than across a period of two years (performance on ENIs in Grade 1 with ITBS in Grade 3). The Number Identification (NI) probe was administered for the first time with Cohort 3 in the fall of Grade 1. NI also showed a strong relation with Grade 3 ITBS Math scores. In all, these analyses suggest that ENIs have adequate long-term predictive validity for predicting performance on the ITBS.

Table 3

*Means, Standard Deviations, and Number of Observations for Early Numeracy Indicators*

| Grades the cohorts were in at the time the measures were administered | | | Fall | | | Winter | | | Spring | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | NI | QD | MN | NI | QD | MN | NI | MN | QD |
| Kindergarten | | | | | | | | | | | |
| | Cohort 1 | | | | | | | | | | |
| | Cohort 2 | *M* | | | | | 18.89 | 8.88 | | | |
| | | *SD* | | | | | (9.42) | (4.10) | | | |
| | | *N* | | | | | 61 | 61 | | | |
| | Cohort 3 | *M* | | 12.58 | 6.64 | | 19.62 | 11.68 | | 24.08 | 14.00 |
| | | *SD* | | (8.79) | (5.45) | | (9.47) | (4.12) | | (9.67) | (4.13) |
| | | *N* | | 44 | 44 | | 45 | 45 | | 45 | 45 |
| Grade 1 | *Cohort 1* | *M* | | | | | 34.03 | 14.37 | | | |
| | | *SD* | | | | | (7.35) | (4.38) | | | |
| | | *N* | | | | | 67 | 67 | | | |
| | *Cohort 2* | *M* | | 31.27 | 16.94 | | 36.60 | 19.01 | | 39.38 | 19.64 |
| | | *SD* | | (10.24) | (4.70) | | (9.60) | (5.30) | | (9.74) | (5.28) |
| | | *N* | | 50 | 50 | | 45 | 45 | | 50 | 50 |
| | Cohort 3 | *M* | 33.32 | 28.73 | 15.82 | 49.56 | 35.94 | 18.10 | 51.59 | 35.99 | 19.67 |
| | | *SD* | (15.71) | (10.20) | (4.87) | (11.10) | (7.14) | (4.55) | (12.73) | (7.14) | (5.07) |
| | | *N* | 71 | 71 | 71 | 72 | 72 | 72 | 71 | 71 | 71 |

*Note*. MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination

Table 4

*Correlations Between Grade 3 ITBS scores and Early Numeracy Indicators*

| ITBS Math | Grade K | | | Grade 1 | | |
|---|---|---|---|---|---|---|
| | NI | QD | MN | NI | QD | MN |
| Cohort 1 | | | | | .49[a] | .55[a] |
| | | | | | *N* = 67 | *N* = 67 |
| Cohort 2 | | .59[a] | .62[a] | | .51[b] | .60[b] |
| | | *N* = 52 | *N* = 52 | | *N* = 44 | *N* = 44 |
| Cohort 3 | | .69[b] | .54[b] | .59[b] | .54[b] | .48[b] |
| | | *N* = 44 | *N* = 44 | *N* = 71 | *N* = 71 | *N* = 71 |

*Note*. All coefficients are statistically significant at $p < .01$level. MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination.

[a] data collected in winter. [b] data collected in fall.

To provide a context against which to evaluate these results, we computed correlations between the ENIs and the within-year short-term criterion measures. These results are presented in Table 5. The MBA was administered to Cohort 1 students when they were in Grade 1 and to Cohort 2 students when they were in Kindergarten. This was the only short-term criterion measure given to Cohort 1.  For both cohorts, moderate correlations were obtained between the MN and QD measures and the MBA. The MBSP was administered only to Cohort 2 in Grade 1 and showed a moderate relationship (*r* greater than .60) with both MN and QD. TEMA-3 was administered to both Cohort 2 and Cohort 3 when the students were in Grade 1. In Cohort 2, though there was not much of a difference, QD ($r = .66$) had a slightly stronger correlation with TEMA-3 than MN ($r = .61$). In Cohort 2, NI had the strongest correlation coefficient with the TEMA followed by QD and MN respectively. The EOC was administered only to Cohort 2 when this cohort was in Grade 1and had moderate correlations with QD ($r = .58$) and MN ($r = .46$). In all, these analyses suggest that the ENIs have comparable validity for predicting performance for short-term criterion measures (MBA, MBSP, TEMA and EOC) within a school year and for predicting a long-term criterion (ITBS).

Table 5

*Correlations between ENIs and Short-term Criterion Measures*

| Criterion measures | | Grade K | | | Grade 1 | | |
|---|---|---|---|---|---|---|---|
| | | NI | QD | MN | NI | MN | QD |
| MBA | Cohort 1 | | | | | .41<br>*N* = 66 | .56<br>*N* = 66 |
| | Cohort 2 | | .60<br>*N* = 55 | .54<br>*N* = 66 | | | |
| | Cohort 3 | | | | | | |
| MBSP | Cohort 1 | | | | | | |
| | Cohort 2 | | | | | .61<br>*N* = 46 | .61<br>*N* = 46 |
| | Cohort 3 | | | | | | |
| TEMA-3 | Cohort 1 | | | | | | |
| | Cohort 2 | | | | | .66<br>*N* = 48 | .61<br>*N* = 48 |
| | Cohort 3 | | | | .62<br>*N* = 41 | .57<br>*N* = 41 | .45<br>*N* = 41 |
| EOC | Cohort 1 | | | | | | |
| | Cohort 2 | | | | | .58<br>*N* = 46 | .46<br>*N* = 46 |
| | Cohort 3 | | | | | | |

*Note*. All coefficients are statistically significant at $p$ <.01 level. EOC= End of course test, MBA = Mini Battery of Achievement, MBSP-Comp = Monitoring Basic Skills Progress – Computation, MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination, TEMA = Test of Early Mathematics Ability.

*Research Question 3: How well do the ENIs predict success/failure on the ITBS?*

Given the preceding results demonstrating the long-term predictive validity of the ENIs, we wanted to examine whether performance on the ENIs would sufficiently predict later performance on the ITBS. We computed a logistic regression and examined the diagnostic accuracy for predicting outcomes for individual students using the district's metric for proficiency in mathematics (e.g., a percentile rank > 40 constitutes passing). We first recoded each student's ITBS Total Math percentile rank to reflect passing or failing status and then used logistic regression methods to determine how well the ENIs predicted passing or failing the ITBS. We considered the following characteristics: sensitivity, specificity, positive predictive power (PPP) and negative predictive power (NPP).

***Sensitivity*** represents the probability that those who failed the ITBS would have been predicted to fail on the basis of their ENI scores. ***Specificity*** is the probability that those who passed the ITBS would have been predicted to pass on the basis of their ENI scores. ***Positive predictive power*** is the probability that those who scored below the cut scores for the ENIs actually failed the ITBS. ***Negative predictive power*** is the probability that those who scored above the cut scores for the ENIs actually passed the ITBS. In other words, the sensitivity and the specificity refer to the accuracy of the ENIs to identify success or failure on the ITBS whereas the PPP and the NPP refer to the probability that the ENIs will correctly discriminate between who will pass or fail the ITBS. ***Overall correct classification*** gives the percentage of cases correctly predicted (for both success and failure) on the whole.

As reported in Table 6, scores on ENIs in the fall of Grade 1 show adequate diagnostic accuracy indicating that they are potentially useful measures for predicting success or failure on the ITBS. A screening measure is most useful if it can predict failure accurately and can thus help in early identification of students and provision of interventions aimed at preventing failure. We chose to use scores from the fall screening for the diagnostic statistics because this is the point in the school year at which teachers are most likely to seek information about their students with regard to potential risk for poor outcomes. The ENIs are strongest with regard to sensitivity and PPP; all of the ENIs (NI, QD, and MN) have probability rates higher than 82% for predicting failure on the ITBS. However, specificity and NPP rates vary more, suggesting that the measures often miss students who score above the cut points on the ENIs in Kindergarten or Grade 1, but subsequently fail the ITBS in Grade 3. Cohort 2 shows better predictive accuracy for QD and MN than does Cohort 3 (e.g., higher overall correct classification). The results in Table 6 suggest that the ENIs are reasonably successful in predicting long-term outcomes on the ITBS. As is the case with many screening instruments, false negatives (those who pass the screener but fail the later test as represented by the NPP data) represent the most difficult challenge.

Screening data provide teachers with tools to identify students who are at risk for failing later high stakes assessments. Most often, these students are provided with supplemental instruction and teachers monitor the progress of individual students toward a goal representing improved performance in mathematics. In an ideal situation, teachers would set these goals with some indication of the likelihood that attaining this level of performance would improve the student’s probability of being successful on the subsequent high stakes test.

Table 6

*Diagnostic Accuracy Statistics for Predicting Success on the ITBS*

| Predictors | *N* | Cut Scores | Sensitivity (%) | Specificity (%) | PPP (%) | NPP (%) | Overall Correct Classification (%) |
|---|---|---|---|---|---|---|---|
| Cohort 2 | | | | | | | |
| F QD Gr1 | 50 | 14.5 | 97.37 | 50 | 92.5 | 75 | 90.91 |
| F MN Gr1 | 50 | 10.5 | 100 | 50 | 92.68 | 100 | 93.18 |
| F QD Gr1 | 50 | 14.5 | 97.37 | 50 | 92.5 | 75 | 90.91 |
| Cohort3 | | | | | | | |
| F NI Gr1 | 71 | 18 | 90.57 | 61.11 | 87.27 | 68.75 | 83.10 |
| F QD Gr1 | 71 | 23 | 94.34 | 61.11 | 87.72 | 78.57 | 85.92 |
| F MN Gr1 | 71 | 12.5 | 98.11 | 38.89 | 82.54 | 87.50 | 83.10 |

*Note*. F = Scores taken from fall ENIs, Gr 1 = Grade 1, MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination.

To explore the degree to which the ENIs might be useful to teachers in setting such goals, we examined the relations between students' spring ENI scores in Kindergarten and Grade 1 and their later passing status on the Grade 3 ITBS. We drew from the work of Espin and her colleagues to create a Table of Probable Success (Espin, Wallace, Campbell, Lembke, Long, & Ticha, 2008). We chose to use spring scores to create our 'Table of Probable Success' because teachers can use these scores to set end of the year goals that are associated with performance levels students need to demonstrate to achieve success on the Grade 3 ITBS. Table 7 shows the probability of success on the ITBS for scores on Spring ENIs in Kindergarten and Grade 1.

Table 7

*Table of Probable Success for Passing the ITBS in Grade 3 based on ENI Scores*

| Predictors | Probability of Passing ITBS | Cohort2 Scores | Cohort3 Scores | Combined Cohort 2 and Cohort 3 Scores |
|---|---|---|---|---|
| S QD Grade K | 70 | - | 15.5 | - |
| | 75 | - | 17.5 | - |
| | 80 | - | 19.5 | - |
| | 85 | - | 22 | - |
| | 90 | - | 25.5 | - |
| | 95 | - | 31 | - |
| S MN Grade K | 70 | - | 9 | - |
| | 75 | - | ~ 10.5 | - |
| | 80 | - | 12.5 | - |
| | 85 | - | 14.5 | - |
| | 90 | - | 17 | - |
| | 95 | - | ~ 24 | - |
| S NI Grade 1 | 70 | - | 44.5 | - |
| | 75 | - | 48 | - |
| | 80 | - | 52 | - |
| | 85 | - | 57.5 | - |
| | 90 | - | 63.5 | - |
| | 95 | - | 74 | |
| S MN Grade 1 | 70 | ~16 | 17 | ~16.5 |
| | 75 | ~16.5 | 18.5 | ~17.5 |
| | 80 | ~17 | ~ 20 | ~18.5 |
| | 85 | ~17.5 | 21.5 | 19 |
| | 90 | 18 | 23.5 | 21 |
| | 95 | 19 | 28 | 23.5 |
| S QD Grade 1 | 70 | ~33.5 | 32 | 31.5 |
| | 75 | 34 | 34.5 | 33.5 |
| | 80 | 35 | 37.5 | 35.5 |
| | 85 | ~36.5 | 41.5 | ~39 |
| | 90 | 38 | 47 | 42 |
| | 95 | 40.5 | 54 | 47.5 |

*Note*. Dashes indicate absence of data. ~ = approximated score, MN = Missing Number, NI = Number Identification, QD = Quantity Discrimination, S = Scores taken from spring ENI data.

When compared, all the three groups in the table (i.e., Cohort 2, Cohort 3, and the full sample comprising both cohorts) show a similar range of scores for a particular probability of success on the ITBS. For example, to get a 70% probability of success on the ITBS, the scores on spring MN in Grade 1 range from 16 to 17 for all three groups. Similarly, for spring QD in Grade 1, the probability of success on the ITBS is 70% if the scores range from 31.5 to 33.5. The results support the viability of using the ENI scores to set targets for students receiving intervention.

## Conclusion

The results of this study contribute to a growing body of literature supporting the Early Numeracy Indicators as effective and technically sound measures for screening and progress monitoring in the early grades. Progressively increasing mean scores from fall to winter to spring across both Kindergarten and Grade 1 were evident across three cohorts of students, indicating that the ENIs are sensitive to changes in early numeracy proficiency within and across years. The correlations between Kindergarten and Grade 1 ENIs and Grade 3 ITBS scores revealed moderate relations indicating adequate long-term predictive validity for ENIs. These results compared positively to data for short–term (within-year) criterion validity, with the ENIs producing correlations of comparable magnitude for both long-term prediction of ITBS scores and short-term prediction of scores on measures like MBA, MBSP, TEMA-3, and EOC. This result suggests that the strength of prediction does not diminish over longer time periods.

The diagnostic accuracy statistics revealed the probability proportions for predicting failure on the ITBS for the ENIs was above 82% and the percentage for overall correct classification for success/failure in ITBS was above 80% for all ENIs (NI, QD, and MN). An examination of all three cohorts revealed that no particular ENI demonstrated comparatively better capability than the others in predicting success/failure on the ITBS; therefore it appears the measures can be used interchangeably.

Examination of the probability of success on the ITBS for specific scores on two of the ENIs (QD and MN) revealed that across the two cohorts and the combined sample, very similar scores were associated a particular probability of success on the ITBS. This result added more credibility to the predictions. For example, the probability of success on the ITBS was 90% if the scores on MN ranged from 18 to 23.5 or if the scores on QD ranged from 38 to 47. There was not enough data on NI to draw such a conclusion, as this measure had only one set of observations.

Future research should examine the extent to which these findings are replicated across years and with other student populations. In addition, the present results were limited primarily to the QD and MN measures, for which the greatest amount of longitudinal data was available. Additional research should examine whether similar results would be obtained with the NI and the Mixed Numeracy measures that have been used in recent Early Numeracy Indicator studies.

References

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

Espin, C., Wallace, T., Campbell, H., Lembke, E., Long, J. D., & Ticha, R. (2008). Curriculum-based measurement in writing: Predicting the success of high-school students on state standards tests. *Exceptional Children*, *74*, 174-193.

Foegen, A., Lembke, E., Klein, K., Lind, L., & Jiban, C. (2006). *Technical adequacy of early numeracy indicators: Exploring growth at three points in time* (Technical Report 13). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

Fuchs, L. S., & Fuchs, D. (1998). Treatment validity: A unifying concept for reconceptualizing the identification of learning disabilities. *Learning Disabilities Research and Practice, 13,* 204-219

Fuchs, L. S., & Fuchs, D. (n.d.). *Using CBM for progress monitoring*. Washington, DC: National Center on Student Progress Monitoring, American Institutes for Research. Retrieved from http://www.studentprogress.org/library/Training/CBM%20Reading/UsingCBMReading.pdf.

Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review, 22*, 27-48.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1998). *Monitoring basic skills progress: Basic math computation* (2nd ed.). Austin, TX: PRO-ED.

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

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa Test of Basic Skills*. Itasca, IL: Riverside.

Impecoven-Lind, L., Olson, J., & Foegen, A. (2009). *Early Numeracy Indicator Screening Data: 2006-2007* (Technical Report 22). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

Irons, C. (2003). *Growing with mathematics*. Guilford, CT: Wright Group/McGraw Hill.

Lembke, E., & Foegen, A. (2005). *Identifying indicators of early mathematics proficiency in Kindergarten and Grade 1* (Technical Report 6). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

Woodcock, R. W., McGrew, K. S., & Werder, J. K. (1994). *Mini-Battery of Achievement*. Itasca, IL: Riverside Publishing.