TECHNICAL REPORT #33:

Exploring the Use of Early Numeracy Indicators for Progress Monitoring: 2008-2009

Jeannette Olson and Anne Foegen

RIPM Year 6: 2008 – 2009

Dates of Study: September 2008 – May 2009

September 2009

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

Two of the Early Numeracy Indicators developed by Lembke and Foegen (Lembke & Foegen, 2005; Olson, Foegen, & Singamaneni, 2009; Nagarkar, Hamptom, Lembke, & Whitaker, 2009) were used as progress monitoring measures with small groups of first grade students. This report describes the alternate-form reliability and sensitivity to growth of the Quantity Discrimination and Mixed Numeracy measures when used to monitor progress. This was the third year Quantity Discrimination was used for progress monitoring. Most of the alternate-form correlation coefficients were at or above the .90 level, which is higher than in previous studies. Mixed Numeracy was used as a benchmarking and progress monitoring measure for the second time during this academic year. We obtained much higher alternate-form correlation coefficients than the previous academic year, with all of the statistically significant correlations at or above the .80 level. Both measures appear to be sensitive to growth, with students showing significant improvement on both measures.

Exploring the Use of Early Numeracy Indicators for Monitoring Progress: 2008-2009

This is the third in a series of studies examining the use of Early Numeracy Indicators (ENIs) that were developed by Lembke and Foegen (Lembke & Foegen, 2005; Olson, Foegen, & Singamaneni, 2009; Nagarkar, Hamptom, Lembke, & Whitaker, 2009) for progress monitoring. In this study, two of the ENIs were used as progress monitoring measures with small groups of first grade students who were struggling with beginning math concepts. This report examines the reliability and sensitivity to growth of these assessments.

Research Questions

The following research questions guided the data analysis:

1. What levels of alternate-form reliability are demonstrated when the ENIs are used as progress monitoring measures?
2. To what extent do the two progress monitoring measures reflect changes in student performance?

Method

Setting and Participants

The study was conducted in an elementary school (grades PreK-3) in a small Midwestern school district on the fringe of an urban community. The school district is composed of four schools. There is one Pre-K through third grade elementary school, one fourth and fifth grade elementary school, one middle school with grades six through eight, and one high school. During the 2008-2009 school year, the district enrolled 1,338 students, with 46 percent being female, 90.5 percent white, 5.4 percent Hispanic, 2.5 percent African American, 1.3 percent Asian, and

less than 1 percent Native American. Nearly 46 percent of the students qualified for free or reduced lunch, and 2.4 percent were identified as English Language Learners.

Although the ENIs were used as benchmarking assessments for all kindergarten and first grade students, the progress monitoring measures were only used with selected groups of students from the four first grade classes at the elementary school. The four classes included a total of 100 students, of whom 51 percent were males, 89 percent were white, 7 percent were Hispanic, 3 percent were African American, and 1 percent were Asian. During the 2008-2009 academic year, 6 percent of the first grade students received special education services and 6 percent were classified as English Language Learners.

Gathering the early numeracy data was a part of the school's typical practices and ongoing commitment to making data driven decisions; therefore, the study was deemed exempt from typical Human Subjects procedures. Individual consent was not needed for students' participation in the data collection efforts.

After all of the fall benchmarking scores had been entered in a database, the Project Coordinator computed each student's percentile rank and prepared class lists of test results. Student scores that fell into the 90^{th} percentile and above, 80^{th} to 89^{th} percentile, 11^{th} through 20^{th} percentile, and below the 10^{th} percentile were color coded on these lists. The Project Coordinator and the Principal Investigator met with the principal and each of the teachers to review the data.

The first grade teachers assigned ten students to the intervention groups during the fall meetings. In all but one case, the teachers chose students who scored below the 10^{th} percentile on at least one of the four benchmarking measures and were not already receiving special education services in mathematics. Table 1 presents the scores and percentile ranges for the students who were assigned to the intervention groups.

Table 1

Fall Benchmarking Scores and Percentiles

		Number Identification		Quantity Discrimination		Missing Number		Mixed Numeracy	
Student	Teacher	Score	Percentile Range	Score	Percentile Range	Score	Percentile Range	Score	Percentile Range
1	2A	12.5	11-20	5	< 10	8.0	11-20	9.5	< 10
2	3A	9	< 10	8.5	< 10	5.5	< 10	13.5	< 10
3	3A	19	21-79	12.5	11-20	4	< 10	13.5	< 10
4	1A	7	< 10	4	< 10	4	< 10	8.5	< 10
5	2A	12.5	11-20	11.5	< 10	7	< 10	11	< 10
6	2A	11.5	< 10	6.5	< 10	4.5	< 10	8.5	< 10
7	3B	16.5	21-79	16	21-79	7.5	11-20	12.5	< 10
8	4B	10	< 10	10	< 10	9.5	11-20	14.5	11-20
9	1B	20	21-79	13	11-20	6	< 10	14	11-20
10	1B	15	11-20	15.5	11-20	6	< 10	12.5	< 10

Note: There are two codes for the teachers in this study because students had one of four classroom teachers (1-4) and one of two intervention teachers (A or B) during the 2008-2009 academic year.

Measures

Early Numeracy Progress Monitoring Measures. Two ENIs (Quantity Discrimination, and Mixed Numeracy) were used monitor the progress of the students in the intervention groups (see Appendix A for samples of one page of each type of measure). Quantity Discrimination required students to name the greater of two numbers. The number of items on the Quantity Discrimination progress measures increased to 63 for the 2008-2009 academic year because many students had reached the earlier maximum of 42 during the 2007-2008 study (Olson & Foegen, 2009). Students responded verbally by naming the number with the greatest value in each pair. Numerals from 0 to 20 were used to create the items. Numbers were randomly selected by using a random number generator. For the Mixed Numeracy measures, students were presented with items that were similar to items from the three other screening measures (see Olson, Foegen, & Singamaneni, 2009, for more information about these measures). This measure began with a row of four Number Identification items (name the number), followed by a row of four Quantity Discrimination items (name the greater of the two numbers), and then a row of four Missing Number items (name the numeral is missing from a sequence of four numbers). This sequence repeated for a total of 84 items. Fifteen forms of each of the two kinds of tasks were developed. Individual progress monitoring booklets were prepared for each measure that included the directions, fifteen alternate forms, and a blank graph for charting the students' score from each administration.

Materials

Both intervention teachers used the 3-Tier Mathematics Model (3TMM) materials that were developed at the University of Texas (Bryant et al., 2008). These materials used a direct instruction approach with scripted lessons that addressed number sense, place value/relationships

of 10, problem solving, and addition and subtraction combinations. Most lessons included modeling, guided practice, individual practice, and suggestions for corrective feedback. These lessons were designed to be used for 15 to 20 minutes during four days of each week. Two classroom teachers had used the 3TMM materials in the two preceding years to provide supplemental instruction for students struggling in mathematics.

Procedures

For the first time, intervention teachers gathered all of the progress monitoring data instead of classroom teachers. Intervention Teacher A had participated in the progress monitoring process as a classroom teacher during the 2006-2007 and 2007-2008 academic years; therefore, the Project Coordinator reviewed the administration procedures with this teacher. The Project Coordinator met with the Intervention Teacher B, who had not previously used the measures, to provide an overview of the measures and administration details. The second teacher had extensive experience using other progress monitoring measures; therefore, she understood how to use the ENIs very quickly. Both intervention teachers were providing supplemental instruction in reading and mathematics to students experiencing difficulty; the school district funded their positions through Title 1.

The measures were administered once a week (generally on Friday) from January through April. Intervention Teacher A alternated between the Quantity Discrimination and Mixed Numeracy measures, giving one during one week, and the other during the next week, while Intervention Teacher B administered both measures each week. Both tasks were individually administered using printed testing materials. Students were given one minute to verbally respond to as many items as they could during this time period. If a student hesitated for three seconds on a particular item, he/she was prompted to "Try the next one." Teachers recorded student

responses in the appropriate booklet and later scored the measures by counting the number of correct responses. Scores were plotted on a graph that was included in the progress monitoring booklet. At a later date, a member of the research team entered all of the scores into a spreadsheet for data analysis.

Data Analyses

Data analyses were conducted using number correct scores for the two kinds of progress monitoring probes. Alternate-form reliability was computed by correlating scores from adjacent administrations of the same measure. To determine the measures' sensitivity to growth, we used ordinary least squares regression to calculate the slope of each student's scores on the two measures. The obtained slope values were calculated to reflect the amount of weekly progress a student achieved. We used the Wilcoxon Signed Ranks Test to determine if students' growth from the beginning to the end of the progress monitoring period was statistically significant.

Results

Descriptive statistics for all of the study measures are reported first. These are followed by alternate-form reliability statistics and the student growth data for the two ENIs that were used as progress monitoring assessments. Means and standard deviations for the Quantity Discrimination and Mixed Numeracy progress monitoring measures are presented in Table 2, as are data for the intervention students on the Fall, Winter, and Spring screening measures. Table 3 includes the same information for the Mixed Numeracy measures.

As we considered the data in Tables 2 and 3, we looked at the distributions for each of the progress monitoring measures. We were most interested in floor or ceiling effects and the progression of means over time. There were no scores of zero for any of the weeks that the two measures were used. With the revised Quantity Discrimination forms, which were changed to

include 63 instead of 42 items, none of the students earned a maximum score. The highest score for these progress monitoring measures was 45. The Mixed Numeracy progress monitoring measures have 84 items, and the highest score ever achieved by the students in the intervention groups was 34, so there were no ceiling effects with either assessment.

Table 2

Descriptive Statistics for the Quantity Discrimination Progress Monitoring Measures

Measure	Week of School	Measure	*n*	Min	Max	*M*	*SD*
Quantity Discrimation		Fall Screen	10	4	16	10.25	4.20
	19 and 20	1	7	11	31	20.86	6.67
	21	2	7	10	35	22.71	8.77
	22 and 23	3	7	16	36	24.43	6.85
		Winter Screen	10	12	37.5	22.45	7.47
	23 and 25	4	7	19	39	27.57	6.97
	24 and 27	5	9	17	40	28.11	8.01
	25 and 30	6	9	22	40	28.56	6.89
	26 and 32	7	10	19	41	29.80	6.56
	27	8	4	28	42	33.50	5.97
	28	9	4	28	43	32.75	6.90
	29	10	4	28	42	34.00	6.06
	30	11	4	29	43	34.00	6.38
	31	12	4	29	44	34.25	6.85
	32	13	4	29	45	35.00	7.12
	33	14	4	28	45	34.25	7.59

		Spring Screen	10	19.5	40	28.05	8.61

Table 3

Descriptive Statistics for the Mixed Numeracy Progress Monitoring Measures

Measure	Week of School	Measure	*n*	Min	Max	*M*	*SD*
Mixed Numeracy		Fall Screen	10	8.5	14.5	12.05	1.99
	20	1	7	14	24	19.14	3.18
	21 and 22	2	7	17	28	20.71	4.35
	22 and 24	3	7	18	28	21.14	4.49
		Winter Screen	10	14.5	25	19.30	3.83
	23 and 26	4	8	19	31	24.00	4.24
	24 and 29	5	9	23	32	27.11	2.89
	25 and 31	6	8	20	30	25.75	3.80
	26	7	9	19	30	25.79	3.63
	27	8	4	20	34	27.75	6.13
	28	9	4	26	31	28.75	2.63
	29	10	3	26	30	27.33	2.31
	30	11	4	21	33	28.00	5.29
	31	12	4	21	33	28.50	5.45
	32	13	4	22	33	28.75	4.99
	33	14	4	23	32	29.25	4.27
		Spring Screen	10	20	34.5	25.60	4.83

As we examined the progression of mean scores on the two measures, we found that the scores on both measures increased over time. For the Quantity Discrimination measures that were only administered to four students, there was not a lot of change in the mean scores from one week to the next, with only modest gains or losses each week. We found more incremental improvement for the Mixed Numeracy measures with the limited sample sizes of three or four students.

Research Question 1. What levels of alternate-form reliability are demonstrated when the Early Numeracy Indicators are used as progress monitoring measures?

The alternate-form reliability correlations between scores from each type of task for each pair of subsequent administrations are reported in Tables 4 and 5. Eleven of the thirteen correlations for Quantity Discrimination were statistically significant, with these coefficients ranging from .84 to 1.00. Eight of the thirteen alternate-form reliability correlations for the Mixed Numeracy progress monitoring assessments were statistically significant, with these coefficients ranging from .89 to .97.

Table 4

Alternate-form Reliability for Quantity Discrimination

Measure	Forms	*N*	*r*	*p*
Quantity Discrimination	QD 1 and QD 2	7	.96	.00
	QD 2 and QD 3	7	.70	.08
	QD 3 and QD 4	7	.90	.01
	QD 4 and QD 5	7	.96	.00
	QD 5 and QD 6	9	.88	.00
	QD 6 and QD 7	9	.84	.01
	QD 7 and QD 8	4	.98	.02
	QD 8 and QD9	4	.98	.02
	QD 9 and QD 10	4	.93	.07
	QD 10 and QD 11	4	.98	.02
	QD 11 and QD 12	4	1.00	.00
	QD 12 and QD 13	4	.96	.02
	QD 13 and QD 14	4	1.00	.00

Table 5

Alternate-form Reliability for Mixed Numeracy

Measure	Forms	*N*	*r*	*p*
Mixed Numeracy				
	MX 1 and MX2	7	.89	.01
	MX 2 and MX 3	7	.96	.00
	MX 3 and MX 4	7	.89	.01
	MX 4 and MX 5	8	.16	.71
	MX 5 and MX 6	8	.85	.01
	MX 6 and MX 7	8	.52	.19
	MX 7 and MX 8	3	.96	.18
	MX 8 and MX9	4	.95	.05
	MX 9 and MX 10	3	.98	.12
	MX 10 and MX 11	3	.80	.41
	MX 11 and MX 12	4	.97	.03
	MX 12 and MX 13	4	.99	.01
	MX 13 and MX 14	4	.97	.03

Research Question 2: *To what extent do the two progress monitoring measures reflect changes in student performance?*

The weekly growth rate data for the two progress monitoring measures are displayed in Table 6. Growth rates on the Quantity Discrimination assessments ranged from -.38 to 1.32 correct responses per week, with a mean rate of improvement of .64 correct responses per week.

Mixed Numeracy growth rates ranged from .21 to 1.16 correct responses per week with a mean rate of improvement of .59 correct responses per week.

Table 6

Weekly Growth Rates

		Slopes	
Student	Teacher	Quantity Discrimination	Mixed Numeracy
1	2A	1.29	.64
2	3A	1.25	.40
3	3A	NA	NA
4	1A	1.32	1.16
5	2A	.41	.22
6	2A	-.38	.75
7	3B	.43	.53
8	4B	.75	.70
9	1B	-.19	.21
10	1B	.91	.67

Note: Student 3 was added to the intervention group later in the year; he only took 3 forms of each measure so his slopes were not calculated.

We used the Wilcoxon Signed Ranks Test to determine if student performance improved significantly from the beginning to the end of the progress monitoring period. For each of the measures we used the mean of the first two scores and the mean of the last two scores for these comparisons. Students' scores improved significantly for both sets of measures (Quantity Discrimination [$z = -1.96$, $p < .05$]; Mixed Numeracy [$z = -2.52$, $p < .01$]).

Discussion

All of the statistically significant alternate-form reliability correlations obtained for the Quantity Discrimination progress monitoring measures were higher than our target of .80, with most coefficients greater than .90. These results were similar to the alternate-form reliability coefficients for the Quantity Discrimination benchmarking measures (Olson, Foegen, & Singamaneni, 2009).

This was the second year we used the Mixed Numeracy measures for both benchmarking and progress monitoring. The alternate-form reliability coefficients for the fall/winter/spring benchmarking measures were all at or above the .80 level. As noted in Table 5, the statistically significant correlations were also in this range; however, only eight of the thirteen comparisons were significant. The correlation coefficient between Form 4 and Form 5 was very different from all the other comparisons. Consequently, it will be important to examine these two forms to determine whether the items have similar difficulty levels.

During this study, both of the ENIs used for progress monitoring were administered throughout most of the spring semester. Both of the measures appear to be sensitive to growth even though the data were drawn from a sample of first grade students who were struggling with mathematics concepts and therefore represented a restricted range of ability. These results should be interpreted with caution given that the sample used to examine growth was very small ($n = 3$ and $n = 4$).

References

Bryant, D., Bryant, B., Gersten, R., Scammacca, N., Funk, C., Winter, A., Shi, M., & Pool, C. (2008). *The effects of tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties. Learning Disability Quarterly, 31,* 47-63.

Impecoven-Lind, L., Olson, J., & Foegen, A. (2009). *Exploring the use of early numeracy indicators for monitoring progress in two intervention contexts* (Technical Report 28). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

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.

Ngarkar, S., Hampton, D. D., Lembke, E. S., & Whitaker, T. (2009). *Missouri Early Numeracy Indicator screening data: 2007-2008*. (Technical Report XX). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

Olson, J. & Foegen, A. (2009). *Exploring the use of early numeracy indicators for monitoring progress in two intervention contexts: 2007-2008* (Technical Report 29). Minneapolis, MN: University of Minnesota, College of Education and Human Development, Research Institute on Progress Monitoring.

Olson, J., Foegen, A., & Singamaneni, S. (2009). *Early Numeracy Indicator Screening Data: 2008-2009* (Technical Report 24). Minneapolis, MN: University of Minnesota, College

of Education and Human Development, Research Institute on Progress Monitoring.

Appendix A

Early Numeracy Indicators

Quantity Discrimination

Sample Quantity Discrimination Measure Page

Mixed Numeracy

Sample Mixed Numeracy Measure Page

Quantity discrimination, page 1—student copy

5	2	7	1	8	3
1	18	8	10	7	8
16	8	9	1	10	7
2	6	8	3	9	4
12	5	9	15	10	8
0	14	0	6	8	10
15	14	6	1	5	1

Mixed Numeracy, page 1 - Student Copy

4	7	2	1
12 | 9	6 | 1	3 | 8	10 | 7
2 3 ___ 5	4 5 6 ___	1 ___ 3 4	7 8 9 ___
41	8	21	11
18 | 9	6 | 10	20 | 15	1 | 7
5 10 15 ___	___ 5 6 7	3 4 ___ 6	20 30 40 ___
14	81	21	50

RIPM Research Institute on Progress Monitoring OSEP Award #H324H030003