



TECHNICAL REPORT #22:

Iowa Early Numeracy Indicator Screening Data: Iowa 2006-2007

Linda Impecoven-Lind, Jeannette Olson, and Anne Foegen

RIPM Year 4: 2006 – 2007

Dates of Study: September 2006 – May 2007

May 2009

The College of Education
& Human Development

UNIVERSITY OF MINNESOTA



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

This report presents the findings from an ongoing examination of three early mathematics screening measures (now called Early Numeracy Indicators) that were developed by Lembke and Foegen (2005). These measures were used as benchmarking tools in the fall, winter, and spring in a small Midwestern school district. We found that the scores on the Quantity Discrimination measures continued to have larger standard deviations than the Missing Number measures; however, the Number Identification measures which were used for the first time in Iowa during the current study, had the widest distribution. As in the earlier studies, mean scores on all of the indicators increased over the course of the academic year. Very similar means were observed for the Missing Number measures during each administration period for the two cohorts, while the means for the Quantity Discrimination measures were marginally lower in 2006-2007. The levels of alternate-form reliability for Number Identification and Quantity Discrimination were very similar to those from earlier studies, while higher levels of alternate-form reliability were produced by the Missing Number measures. We found that the scores for the Quantity Discrimination measures and the Number Identification measures were highly correlated, and that there was a fair amount of intercorrelation between the Quantity Discrimination measure and the Missing Number measure. If this pattern continues in future studies, it may be that teachers can save time by using the Quantity Discrimination measure by itself; however, we also must examine which of these measures proves to be the best predictor of students' performance on the third grade assessment that is used for determining annual yearly progress.

Iowa Early Numeracy Indicator Screening Data: 2006-2007

This report presents the findings from an ongoing examination of three early mathematics screening measures (now called Early Numeracy Indicators) that were developed by Lembke and Foegen (2005). Two of these screening measures (Quantity Discrimination and Missing Number) were used in earlier studies in Missouri (Lembke & Foegen, 2005) and Iowa (Lembke & Foegen, 2005; Foegen, Lembke, Klein, Lind, & Jiban, 2006). The Number Identification task has been used in Missouri (Lembke & Foegen, 2005); however, this was the first time that the Number Identification task was used with Iowa students. There were no significant changes to the measures during this round of data collection. The major change in the data collection procedures was the design and development of a screening booklet that was used for recording students' responses and scores throughout the academic year to streamline the recordkeeping process.

Research Questions

The following research questions guided the data analysis:

1. Are the scores earned by kindergarten and first grade students similar to those from the 2005-2006 study for the three screening periods?
2. When compared to earlier studies, were similar levels of alternate-form reliability found during the 2006-2007 academic year?
3. To what extent are the different measures intercorrelated?

Method

Setting and Participants

The study was conducted in a small Midwestern district on the fringe of an urban community. The district included four schools: two elementary schools, a middle school, and a high school. During the 2006-2007 school year, the district enrolled 1,424 students, with 52.9 percent being male, 93.3 percent white, 3.7 percent Hispanic, and 3 percent other ethnicities. Nearly 44 percent of the students qualified for free or reduced price lunch, and approximately 17 percent of students were receiving special education services.

All four of the kindergarten and all four of the first grade classes participated in this study for a total of 181 kindergarten and first grade students. Of the total student population, 103 (57%) were males and 78 (43%) were females. Ninety percent of all participants were white, 9 percent were Hispanic, and 1 percent were identified as belonging to another ethnic group. Nine percent of the kindergarten students and seven percent of first grade students were receiving special education services.

Gathering the early numeracy data was a part of the school's typical practices and ongoing commitment to making data driven decisions; therefore, individual consent was not needed for students' participation in the data collection efforts.

Measures

Early Numeracy Indicators. Three early numeracy indicators (Number Identification, Quantity Discrimination, and Missing Number) were used as benchmarking tools for all kindergarten and first grade students three times during the school year (fall, winter, and spring; see Appendix A for samples of one page of each type of measure). As noted earlier, this was the first year that the Number Identification measure was used with Iowa students.

The Number Identification tasks had 84 boxes with numerals (ranging from 0 to 100) in them. Each student was to say the names of as many of the numerals as he or she could in the time allotted. All of the 63 items in the Quantity Discrimination measures had a pair of numerals (ranging from 0 to 20). Students were to say the name of the greater number in each pair. For the Missing Number measures, each item was a box with a sequence of three numerals and a blank line. The position of the blank line varied across the four possible positions. Students were to state the name of the missing number in the sequence. Most sequences involved counting by ones; however, some involved counting by fives or tens.

Two forms of each measure were individually administered by trained data collectors during each data collection period (fall, winter, and spring), for a total of six different forms for each indicator. Students were given one minute to attempt as many items as they could for each task.

Procedure

Trained data collectors gathered all of the data. Each data collector participated in a small-group training session lasting approximately one hour. The project coordinator delivered this training session. During the training session an overview of the study was provided, then the project coordinator modeled how to administer each of the three measures. Data collectors practiced administering each of the tasks and then administered each task to a peer while the trainer observed and completed an 11-item fidelity checklist. All of the data collectors were required to achieve 100% percent accuracy before data collection with students began.

Students participated in three rounds of data collection spread across the academic year. Fall data were collected at the end of October, winter data during the end of February, and spring data during the beginning of May. During each data collection period, two forms of each task

were individually administered, with each data collection session lasting approximately ten minutes per child. Administration of the tasks took place at desks or tables in the hallways outside of the students' classrooms. Data collectors provided a brief introduction to each measure and had each student try three sample problems to ensure that the student understood the task before administering the two forms of a measure. The data collectors wrote down each student's oral responses in a screening booklet. All of the measures were hand scored by counting the number of correct responses.

Students who were absent during a data collection day were assessed if they could be assessed within the one-week time limit. If this could not be accomplished, that student's data was omitted for that period, but the student was assessed in subsequent rounds of data collection using standard procedures.

Project staff completed all of the scoring and data entry. Twenty percent of the measures were rescored during each round of data collection to assess inter-scorer agreement. We computed an estimate of agreement by counting the number of items considered agreements (i.e., scored correctly) and the number of items for which there was a disagreement in scoring (i.e., scoring errors) and dividing the number of agreements by the sum of agreements and disagreements. We computed scoring accuracy by measure type for each of the selected scoring booklets and then averaged across all of the booklets to obtain an overall estimate of inter-scorer agreement. The scoring accuracy results are presented in Table 1. Scorers were very consistent with mean agreement levels of at least 99% or better.

Table 1

Mean Agreement, Range and Number of Probes Examined for Inter-scorer Agreement

	Number Identification			Quantity Discrimination			Missing Number		
	Mean Agree	Range	<i>n</i>	Mean Agree	Range	<i>n</i>	Mean Agree	Range	<i>n</i>
Fall	99%	93-100%	78	100%	94-100%	78	100%	93-100%	78
Winter	99%	73-100%	64	100%	97-100%	64	100%	94-100%	64
Spring	100%	97-100%	64	100%	100%	64	100%	100%	64

Data Analyses

Data analyses were conducted using number correct scores for each of the three Early Numeracy Indicators. Alternate-form reliability was computed by correlating scores from the two forms of each type of measure for each data collection period.

Results

Descriptive statistics for all of the study measures are reported first. These are followed by alternate-form reliability statistics and the intercorrelation data for the different early numeracy indicators. Means and standard deviations for kindergarten students' scores on the three different measures are presented in Table 2. The same data for first grade students appears in Table 3. Tests of skewness and kurtosis were conducted for all study variables and distributions. During the fall administration, the statistics for skewness and kurtosis for the Number Identification tasks for kindergarten students fell outside of the commonly acceptable range, as did the kurtosis values for one form of Missing Number for kindergarten, and both forms for first grade students. For the winter and spring administrations, the statistics were within the commonly acceptable range.

Table 2

Descriptive Statistics for Early Numeracy Indicators for Kindergarten Students

<u>Kindergarten</u>							
<u>Measure</u>	<u>Date</u>	<u>Measure</u>	<u>n</u>	<u>Min</u>	<u>Max</u>	<u>M</u>	<u>SD</u>
Number Identification	Fall	Form 1	83	0	51	15.48	10.37
		Form 2	83	0	47	11.34	9.90
		Mean	83	0	49	13.41	9.80
	Winter	Form 1	85	4	60	26.81	11.75
		Form 2	85	0	47	20.51	10.95
		Mean	85	2	54	23.66	11.02
	Spring	Form 1	82	3	54	26.16	13.55
		Form 2	83	1	52	23.29	13.19
		Mean	83	2	52	24.69	13.05
Quantity Discrimination	Fall	Form 1	83	0	33	11.47	8.51
		Form 2	82	0	31	9.44	8.41
		Mean	83	0	30	10.53	8.34
	Winter	Form 1	85	0	35	16.72	9.14
		Form 2	85	0	34	16.20	9.48
		Mean	85	0	34.5	16.46	8.97
	Spring	Form 1	83	0	41	19.55	9.86
		Form 2	83	0	36	18.45	9.29
		Mean	83	0	38	19.00	9.37
Missing Number	Fall	Form 1	83	0	20	6.33	5.04
		Form 2	83	0	16	7.17	4.83
		Mean	83	0	17.5	6.75	4.70
	Winter	Form 1	85	0	21	9.00	5.16
		Form 2	85	0	26	9.18	6.12
		Mean	85	0	23	9.09	5.41
	Spring	Form 1	83	0	23	10.95	4.94
		Form 2	83	0	24	11.96	5.65
		Mean	83	0	22	11.46	5.06

Table 3

Descriptive Statistics for Early Numeracy Indicators for First Grade Students

<u>Grade 1</u>							
<u>Measure</u>	<u>Date</u>	<u>Measure</u>	<u>n</u>	<u>Min</u>	<u>Max</u>	<u>M</u>	<u>SD</u>
Number Identification	Fall	Form 1	87	6	68	34.93	14.61
		Form 2	87	3	62	30.17	15.71
		Mean	87	6	65	32.55	14.80
	Winter	Form 1	86	15	77	51.86	11.37
		Form 2	86	23	74	45.01	12.21
		Mean	86	19	76	48.45	11.41
	Spring	Form 1	83	29	81	54.64	12.04
		Form 2	83	15	83	48.17	13.38
		Mean	83	22	80	51.40	12.36
Quantity Discrimination	Fall	Form 1	87	4	50	29.21	9.61
		Form 2	87	1	49	27.08	10.72
		Mean	87	3	50	28.14	10.04
	Winter	Form 1	86	17	49	35.93	6.55
		Form 2	86	12	49	35.17	7.41
		Mean	86	14.5	48.5	35.55	6.71
	Spring	Form 1	83	11	56	36.71	7.75
		Form 2	83	14	52	35.28	7.30
		Mean	83	14	54	35.99	7.15
Missing Number	Fall	Form 1	87	0	33	15.40	5.37
		Form 2	87	0	28	15.49	4.85
		Mean	87	0	26.5	15.45	4.88
	Winter	Form 1	86	5	30	17.16	4.70
		Form 2	86	4	30	18.37	5.21
		Mean	86	4	30	17.77	4.53
	Spring	Form 1	83	9	34	19.67	4.80
		Form 2	83	5	34	19.63	5.66
		Mean	83	7.5	33	19.65	4.93

Research Question 1. Are the scores earned by kindergarten and first grade students similar to those from the 2005-2006 study for the three screening periods?

As we considered the data in Tables 2 and 3, we looked at the distributions for each of the indicators. We were most interested in floor or ceiling effects and the size of the standard

deviations. Throughout the year, there were kindergarten students who received scores of zero on the screening measures. In the fall, there were the fewest zeroes on the Number Identification tasks (1 on Form 1, and 2 on Form 2), the next fewest on the Missing Number tasks (6 on both forms), and the most zeroes on the Quantity Discrimination tasks (10 on both forms). With each screening period, the number of scores that were zeroes decreased, with no students obtaining a score of zero on a Number Identification measure in the spring, three students scoring a zero on a Quantity Discrimination measure, and two students with scores of zero on a Missing Number measure. In contrast, there was one first grade student who earned a score of zero on a Missing Number task during the fall screening. No other scores of zero were earned during any of the other administration periods on any of the three measures.

There were no concerns about ceiling effects with the kindergarten data; however, as we examined the first grade data, we found one student who scored 81/84 and one student who scored 83/84 on the Number Identification indicator during the spring administration period.

Looking at the standard deviations for each measure for each grade level, we found that the Number Identification measures produced the widest distribution of scores, and the Missing Number measures produced the smallest distributions. Mean scores on all three measures increased for each subsequent administration.

Comparing the descriptive statistics for the Quantity Discrimination and Missing Number indicators from 2006-2007 with those from 2005-2006, we found a different pattern for possible floor effects with a small number of kindergarten students (1 or 2) still scoring zeroes on each of these measures during the spring administration periods of the current study. In addition, more students received zeroes on the Quantity Discrimination tasks than on the Missing Number tasks during the current study, while the converse was true for the 2005-2006 study. When we looked

at the distributions of scores, we found that the Quantity Discrimination indicators continued to have larger standard deviations than the Missing Number indicators; however, the Number Identification indicators, which were used for the first time in Iowa during the current study, had the widest distribution. During both the 2005-2006 and 2006-2007 studies, mean scores on all of the indicators increased over the course of the academic year. Very similar means were observed for the Missing Number measures during each administration period for the two cohorts, while the means for the Quantity Discrimination measures were marginally lower in 2006-2007 when compared to the earlier cohort.

Research Question 2. When compared to earlier studies, were similar levels of alternate-form reliability found during the 2006-2007 academic year?

The alternate-form reliability correlations between the two forms of each type of probe are displayed in Table 4. For the most part, the Number Identification and Quantity Discrimination measures had very good levels of alternate-form reliability with all but one correlation at the .85 level or higher for kindergarten and first grade students. The alternate-form reliability for the Missing Number measures was better for kindergarten students than for first grade students, with the correlation falling below the .80 standard in both the winter and the spring.

Table 4

Alternate-form Reliability

	<u>Kindergarten</u>			<u>Grade 1</u>		
	Fall	Winter	Spring	Fall	Winter	Spring
NI Means	.87	.89	.92	.91	.87	.89
QD Means	.93	.85	.92	.95	.85	.81
MN Means	.81	.84	.83	.83	.67	.78

Note. All correlations are significant at $p < .01$ level.

We found that the levels of alternate-form reliability for the Number Identification measures were very similar to those found in earlier studies with students who lived in Missouri. The levels of alternate-form reliability for the Quantity Discrimination measures were very similar to those reported in earlier studies that included students from Iowa and Missouri. When examining the data for the Missing Number measures, we found that, on the whole, the levels of alternate-form reliability were better than in earlier studies, except for first graders during the winter administration period.

Research Question 3. To what extent are the different measures intercorrelated?

Table 5 displays the intercorrelations between the three early numeracy indicators over the course of the academic year in Kindergarten, while Table 6 includes the same information for first grade. As one might expect, there was a high degree of intercorrelation between the same kind of task given at different times during the academic year, especially for the Number Identification tasks and the Quantity Discrimination tasks for kindergarten students. When considering the intercorrelation between different measures, the highest coefficients were obtained between the Number Identification and Quantity Discrimination indicators, followed by

Quantity Discrimination and Missing Number. The least amount of intercorrelation was between the Number Identification and Missing Number tasks for kindergarten students.

Table 5

Intercorrelations Between Early Numeracy Indicators for Kindergarten Students

	FNI Mean	FQD Mean	FMN Mean	WNI Mean	WQD Mean	WMN Mean	SNI Mean	SQD Mean	SMN Mean
FNI Mean	--								
FQD Mean	.83	--							
FMN Mean	.66	.71	--						
WNI Mean	.81	.77	.56	--					
WQD Mean	.70	.83	.72	.83	--				
WMN Mean	.56	.69	.66	.67	.69	--			
SNI Mean	.70	.80	.53	.84	.78	.57	--		
SQD Mean	.64	.75	.61	.72	.78	.70	.84	--	
SMN Mean	.51	.53	.65	.58	.67	.73	.60	.71	--

Note. All correlations are significant at $p < .01$ level.

Table 6

Intercorrelations Between Early Numeracy Indicators for First Grade Students

	FNI Mean	FQD Mean	FMN Mean	WNI Mean	WQD Mean	WMN Mean	SNI Mean	SQD Mean
FNI Mean	--							
FQD Mean	.86	--						
FMN Mean	.73	.78	--					
WNI Mean	.75	.69	.59	--				
WQD Mean	.64	.64	.63	.77	--			
WMN Mean	.65	.55	.71	.65	.73	--		
SNI Mean	.62	.58	.54	.85	.70	.62	--	
SQD Mean	.59	.58	.63	.71	.82	.67	.73	--
SMN Mean	.51	.46	.65	.51	.62	.77	.57	.73

Note. All correlations are significant at $p < .01$ level.

When we examined the intercorrelations between the same tasks during different administration periods for first grade students, we found a different pattern with the highest intercorrelations between the Missing Number measures, and the least overlap with the Quantity Discrimination measures. Nevertheless, the findings with regard to intercorrelation between different types of tasks were the same as those for kindergarten with the greatest between Number Identification and Quantity Discrimination and the least between Number Identification and Missing Number.

One of the dilemmas we are wrestling with as we move forward with the exploration and future dissemination of these Early Numeracy Indicators is the amount of time it takes to individually administer multiple forms of several measures. Considering the intercorrelation data

gathered in this study, we found scores for the Quantity Discrimination measures and the Number Identification measures were highly correlated, with all of the coefficients above .80 for kindergarten students and above .70 for first grade students. In addition, there was a fair amount of intercorrelation between the Quantity Discrimination measures and the Missing Number measures with all but one of the correlation coefficients at .70 or above. If this pattern continues in future studies, it may be that teachers can save time by only using the Quantity Discrimination measures; however, we also must examine which of these measures proves to be the best predictor of students' performance on the third grade assessment that is used for determining annual yearly progress.

References

- Foegen, A., Lembke, E., Klein, K., Lind, L., & Jiban, C. (2006). *Technical adequacy of early numeracy measures: 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.
- 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.

Appendix A

Early Numeracy Indicators

Number Identification

Sample Number Identification Measure Page

Quantity Discrimination

Sample Quantity Discrimination Measure Page

Missing Number

Sample Missing Number Measure Page

Number Identification, page 1—Student copy

6

26

39

9

16

5

18

8

6

8

4

0

18

30

16

2

18

94

17

22

7

64

47

9

1

34

24

97

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
---	---

Missing Number, page 1—Student copy

___ 8 9 10

___ 7 8 9

4 ___ 6 7

30 40 50 ___

1 2 3 ___

4 5 ___ 7

___ 3 4 5

4 5 6 ___

7 8 ___ 10

3 ___ 5 6

7 8 ___ 10

6 7 8 ___

10 15 20 ___

6 7 ___ 9

1 2 ___ 4

3 4 5 ___

4 ___ 6 7

5 6 7 ___

0 1 ___ 3

___ 1 2 3

5 ___ 7 8