TECHNICAL REPORT #36:

Monitoring Mathematics Progress in Middle School

Anne Foegen and Linda Impecoven-Lind

November 2009

RIPM Year 4: 2006 – 2007

Dates of Study: January to May 2007

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 extends work on middle school mathematics measures by providing an exploratory examination of their use for progress monitoring. Six teachers (two general education and four special education) administered six different middle school mathematics progress monitoring measures to their students in grades five to eight over a period of four months. Teachers' initial plans to collect data weekly for all six measures could not be sustained, producing a data set with variable numbers of administrations of the measures across students, grade levels, and class types (special education, at risk, general education). Three research questions were investigated, examining the reliability, criterion validity, and sensitivity to growth of the measures when administered on a more frequent basis. Results indicated that teacher administration of the measures results in more variable levels of alternate form reliability than have been obtained in previous studies in which the researchers administered the measures. Criterion validity results supported the conclusion that students' performance on the measures is related to other indicators of their mathematics proficiency, including teacher ratings and standardized test scores. The pattern of scores obtained across the six different measures differed somewhat from previous research in that a broad concepts/application measures did not produce markedly stronger criterion validity coefficients than the other measures. With respect to sensitivity to growth over time, the results reflected positive mean weekly slope values. Given the small sample sizes, erratic data collection schedules, and variable levels of alternate-form reliability, future research should be conducted to replicate the study.

Monitoring Mathematics Progress in Middle School

Background Information

The research on progress monitoring tools in mathematics provides only limited guidance for teachers at the middle school level (Foegen, Jiban, & Deno, 2007). Foegen (2008) described the existing measures that have been used or are currently under investigation for middle school mathematics. Foegen and her colleagues (Foegen, 2000, Foegen & Deno, 2001) have conducted initial research on potential measures involving estimation and basic fact proficiency. In more recent research conducted by the Research Institute on Progress Monitoring, additional measures (Missing Number, Complex Quantity Discrimination) involving numeracy concepts have been explored (Foegen, 2008; Foegen, Klein, Lind, & Jiban, 2008, Foegen & Lind, 2009).

Evidence of technical adequacy varies by type of measure. For some measures, such as basic facts and estimation, evidence exists for the measures' technical adequacy as a static measure (necessary for screening) and as an indicator of weekly growth (necessary for progress monitoring). Foegen et al.'s (2008) study was a first effort to establish the reliability and criterion validity of the measures as static indicators and to explore rough estimates of growth across a school year. Foegen et al. administered six measures (Monitoring Basic Skills Progress-Computation, Monitoring Basic Skills Progress-Concepts and Applications, Basic Facts, Estimation, Missing Number, Complex Quantity Discrimination) at three points in time (fall, winter, spring) to an inclusive general education sample including large numbers of sixth and seventh grade students (N = 242 and 264, respectively), and a smaller number (N = 57) of eighth grade students. They found that these measures generally had acceptable levels of test-retest and alternate form reliability, with the exception of the Missing Number measure (across grades) and the Monitoring Basic Skills Progress Computation measure (Fuchs, Hamlett, & Fuchs, 1998) in

Grade 7. Although the Basic Facts and Monitoring Basic Skills Progress-Concepts and Applications measures consistently produced reliable scores with a single administration, the other four measures (Monitoring Basic Skills Progress-Computation, Estimation, Missing Number, Complex Quantity Discrimination) proved to be more reliable when the mean of two forms was used in the analyses. The strongest concurrent criterion validity coefficients were obtained for the Monitoring Basic Skills Progress-Concepts and Applications (Fuchs, Hamlett, & Fuchs, 1999) in Grades 6 and 7 and the Complex Quantity Discrimination measure for students in Grade 8 (who did not complete either of the Monitoring Basic Skills Progress measures). Consistently weak relations were obtained with the Monitoring Basic Skills Progress Computation measure for students in Grade 7.

Predictive validity was examined with students in Grades 6 and 7; these results paralleled the concurrent validity results. For students in Grade 6, strong relations were obtained for the Monitoring Basic Skills Progress Concepts and Applications and Computation measures. For students in Grade 7, relations were strongest for the Monitoring Basic Skills Progress Concepts and Applications measures (although the coefficients for three other measures were near $r = .60$: Missing Number, Basic Facts, and Complex Quantity Discrimination).

Foegen and Lind (2009) replicated this work with a sample of over 470 students in Grades 7 and 8. The brief middle school measures were administered in the spring; the study examined the alternate-form reliability and the concurrent criterion validity of the measures. For students in Grade 7, relations between the brief mathematics measures and the criterion measures (teacher ratings administered in the spring and a standardized achievement test that had been administered in the winter) were compared with relations obtained between three district-selected screening tools that had been administered in the fall and the standardized achievement test scores. The results of this study indicated that the district screeners (particularly an algebra aptitude measure and a multiple choice measure reflecting the full array of key concepts in the Grade 7 curriculum.) had stronger relations with the high stakes assessment than the majority of the progress monitoring measures. Among the progress monitoring measures, the Grade 6 concepts and applications measure produced criterion validity coefficients and classification accuracy results that were similar to the district screeners.

Purpose of the Study

The purpose of the present study was examine the technical adequacy of six middle school mathematics measures, with specific attention on their use for monitoring student progress. This study extends previous research (Foegen et al., 2008, 2009) in which the measures were administered at 1 to 3 points in time by incorporating weekly monitoring. The progress monitoring measures examined included two measures that have been investigated previously with middle school students (facts, estimation; Foegen, 2000, 2008; Foegen & Deno, 2001), two measures that have documented technical adequacy for use at elementary grade levels (Monitoring Basic Skills Progress-Computation, Monitoring Basic Skills Progress-Concepts and Applications; Fuchs et al., 1998, 1999), and two measures that represent extensions of elementary numeracy measures developed in other RIPM research (Complex Quantity Discrimination, Missing Number; Foegen et al., 2008, 2009). This study replicated previous examinations of the measures' alternate-form reliability and criterion validity. It differed from previous research because the students in the sample represented a wider range of grades (Grades 5 through 8, rather than 6 through 8 in earlier studies) and because the analyses focused on examination of growth over time in students' performance on the measures.

Research Questions

The following research questions guided the study:

1. What levels of alternate-form reliability are produced by the middle school mathematics progress monitoring measures?
2. What levels of criterion validity are produced by the middle school mathematics progress monitoring measures?
3. To what extent do the middle school mathematics progress monitoring measures reflect changes in student performance over time?

Method

Setting and Participants

The study was conducted in a midwestern middle school. During the 2006-07 school year, the middle school enrolled 409 students (86, 101, 113, and 109 students in Grades 5, 6, 7 and 8, respectively), with 50 percent males, 96 percent white, 2 percent Hispanic, 1.5 percent Asian, and less than 1 percent other ethnicities. Twenty-seven percent of the students in the district qualified for free or reduced lunch, and 1.2 percent were identified as Limited English Proficient.

Teacher participants in the study included two general education teachers (one each in Grades 7 and 8) and four special education teachers, working with students in Grades 5 to 8. The Grade 7 general education teacher taught two class periods of math that were included in the study; of these, one period was a grade level math class and the other was a class designed for at risk students. The Grade 8 general education teacher taught 1 grade level math class period that was included in the study. The special education teachers taught groups of students with disabilities that ranged in size from 6 to 9.

We obtained parental consent and student assent for 85 of the 90 students in participating class periods involved in the study. Although teachers were provided with data for all students, the data for this technical report include only those students for whom consent and assent were obtained. Table 1 displays the demographic characteristics of the student participants. The Total Percentage column in Table 1 was calculated using the 85 students in the total sample. Readers will note that students with disabilities are over-represented in the present study. This is due to the fact that the district's primary interest in participating in the study was to investigate the use of alternative measures for monitoring student progress, particularly in Grades 7 and 8.

Table 1

Demographic Characteristics of Student Participants

	Grade 5 (*N* = 6)	Grade 6 (*N* = 8)	Grade 7 (*N* = 42)	Grade 8 (*N* =29)	Percentage of Total Sample
Gender					
Male	1	3	24	14	49%
Female	5	5	18	15	51%
Ethnicity					
White	5	7	36	28	89%
African American	1	0	3	0	5%
Asian	0	0	2	0	2%
Hispanic	0	1	1	1	4%
English Language Learners	0	1	1	0	2%
Special Education Services	5	8	13	9	42%
Free/Reduced Lunch	4	6	14	9	39%
Class Type					
General Education	0	0	23	20	51%
At Risk	0	0	10	3	15%
Special Education	6	8	9	6	34%

Note. Categories may not total 100% due to rounding.

Measures

Middle School Mathematics Progress Monitoring Measures. Six mathematics progress monitoring measures were investigated in this study: Monitoring Basic Skills Progress Computation (MBSP-Comp; Fuchs et al., 1998), Monitoring Basic Skills Progress- Concepts and Applications (MBSP-ConApp; Fuchs et al., 1999), Basic Facts, Estimation, Complex Quantity Discrimination, and Missing Number. The MBSP-Comp task reflected 6th grade computation skills and required students to compute addition, subtraction, multiplication, and division problems involving whole numbers, fractions, and decimals. The task consisted of 25 problems; students had six minutes to answer as many problems as possible. This measure was scored using the procedures prescribed by the multi-district education agency within which the school was located. This scoring method considered the number of correct digits in all partial answers obtained in solving the problem, as well as the final answer. Figure 1 illustrates the digits correct scoring procedure used for the MBSP-Comp probes.

Figure 1. Illustration of Digits Correct Scoring

Digits Correct – Full Problem

```
  56
x 35
 280    (3 digits)
1680    (4 digits)
1960    (4 digits)
```

The digits in each partial product and the final answer are scored, resulting in a total of 11 points possible for the problem. Students who write the correct answer were awarded all 11 points, even if they did not show the partial products.

The MBSP-ConApp task reflected Grade 6 concepts and applications curricula and required students to answer problems related to numeration, applied computation, measurement, geometry, percentages, charts and graphs, word problems, ratios and probability, proportions, and variables. Students had seven minutes to complete 24 items. We scored this measure by counting the number of correct responses.

Both the MBSP-Comp and the MBSP-ConApp were currently being used by the district (using Grade 5 versions for students in that grade and Grade 6 measures for students in Grades 6 through 8) for purposes of monitoring student progress toward IEP goals. For the present study, all participating students completed the Grade 6 forms of the measures.

The Basic Facts task required students to solve single digit computation combinations (0-9) in each of the four operations. The task consisted of 80 items; students had one minute to respond. We scored the Basic Facts measure by counting the number of correct responses.

The Estimation measure required students to select the most appropriate estimate from three alternatives when presented with a problem involving either computation (i.e., 1.8 x 4 is about) or contextual (story) problems (i.e., Jared had $15.50 from mowing lawns. He made another $9.25 from working at the retirement center. About how much does Jared have now?). The problems included all four operations and whole numbers, fractions, decimals, and percents. The answer alternatives differed by a factor of 10 (i.e., 4, 40, 400) to encourage students to use estimation and number sense or mental math, rather than exact computation, to select the most appropriate estimate. The task consisted of 40 problems and students had 3 minutes to circle the correct alternative for each problem. We scored this measure by counting the number of correct responses and then subtracting from this value one-half the number of incorrect responses. This scoring method provides a correction for the potential that students will guess in responding to the multiple-choice item format and has been found in previous research to produce scores with higher levels of technical adequacy (Foegen, 2000).

The Complex Quantity Discrimination task consisted of 44 pairs of numbers or equations. The task required students analyze the pairs of quantities (i.e., 40 + 7 and 70 + 4) and write the appropriate symbol (greater than [>], less than [<], or equal to [=]) in a box between the quantities. Students had one minute to complete the task. We scored the task by counting the number of correct responses.

To complete the Missing Number task, students were presented with a series of three numbers and one blank indicating a missing fourth element in the sequence (the position of the blank varied). The number patterns included counting sequences, multiples (i.e., 3, 6, 9, ___), skip counting starting on an off digit (i.e., 11, 14, ___, 20) and halving/doubling patterns. Students responded by writing the missing number in the blank. The task consisted of 44 items; students had two minutes to respond. The duration represents a change from previous studies. We had obtained restricted distributions in earlier studies (Foegen et al., 2008) and wanted to investigate the effects of a longer duration on the technical adequacy of the measure. We scored the measure by counting the number of correct responses. Samples of the researcher-developed mathematics measures, including administration directions, are provided in Appendix A.

Criterion measures. The criterion measures used in the study included teachers' ratings of their students' overall math proficiency, end of course mathematics grades, and standardized test scores. Teachers were asked to rate each student's general proficiency in math relative to other students in that grade using a Likert scale. The scale ranged from one to seven, with one being least proficient and seven being most proficient. General education teachers were asked to try to spread student ratings across the full range of the scale, not clustering students only in the middle or toward one end. All teachers rated students at the end of the spring semester. A sample of the teacher rating form is provided in Appendix B.

Data on student grades were collected by converting each student's end of year letter grade in mathematics to a number on a 4.0 scale. For example, a grade of A was recorded as 4.0, A- as 3.67, B+ as 3.33, B as 3.0 and so forth. Grades of A+ were awarded to a small number of students and were converted to scores of 4.33.

The district's high stakes achievement test was the Iowa Tests of Basic Skills (ITBS; Hoover, Dunbar, & Frisbie, 2001), which is administered annually for reporting and accountability purposes. The ITBS is a valid and reliable measure that can be used to describe students' developmental level within a content area (Salvia & Ysseldyke, 2007). The test was designed to measure students' critical thinking skills (i.e., interpretation, inferential thinking, classification, analysis, comparisons). The ITBS is designed for students in kindergarten through eighth grade; approximately five and one half hours are required to administer the entire battery of tests. Internal consistency reliability, based on the performance of the standardization sample in 2000, ranged from $r = .66$ to .93 (Salvia & Ysseldyke, 2007). Salvia and Ysseldyke described three types of evidence the test authors reported in support of the content validity of the ITBS. First, they consulted curriculum guides, textbooks, classroom teachers, and school administrators when writing test items. Second, test items were administered to a national sample of students at each grade level. Third, each test item was reviewed for content fit and item bias by a group of experts (Salvia & Ysseldyke, 2007). The Math Problem Solving and Data Interpretation, Math Concepts and Estimation, and Computation subscales, as well as the Math Total Score were used in the analyses. The Math Concepts and Estimation test includes numeration and number systems, whole numbers, geometry, measurement, fractions, money, and number systems problems. The Math Problem Solving and Data Interpretation test is comprised of two sections. One section requires students to solve short word problems while the other section requires students to interpret information in graph and table format. The Computation subtest, which is optional and does not contribute to the Total Math scale score, requires students to apply all four operations to whole numbers and rational numbers. The Total Math scale score is derived from student performance on the Concepts and Estimation and the Problem Solving and Data Interpretation subtests. We recorded the national percentile rank earned by each student on each scale for use in our analyses.

Procedures

The ITBS was group administered to students by district staff using standardized assessment procedures. Standard administration directions (as specified in the teacher directions) and timing requirements were used; the district administered the ITBS in the early spring (February).

The first author visited with teachers prior to the start of the study to explain the study procedures and data collection schedule. Teachers then explained the study to their students and collected the parent consent and student assent forms. Data collection took place from January through April. The original agreement was that the general education teachers would administer one form of each of the six types of probes monthly. Despite cautions from the first author, special education teachers expressed interest (and willingness) to administer all six types of probes weekly for the duration of the study. General and special education teachers were provided with copies of the measures and administration scripts. A research associate observed each teacher administering the measures on at least one occasion to verify adherence to standardized procedures. General education teachers collected students' completed probes, which were then scored by project staff. Special education teachers scored all probes completed by their students.

Data Analysis

Data analyses involving the brief middle school mathematics measures were conducted using number correct scores for the Basic Facts, MBSP-ConApp, Complex Quantity Discrimination, and Missing Number measures. Digits correct scores were used for the MBSP-Comp measure. For the Estimation measure, the final score was computed by subtracting one-half the number of incorrect responses from the total number of correct responses. Any negative scores (resulting when students had more than twice as many incorrect as correct answers) were converted to 0. All of the brief middle school mathematics measures were hand scored using answer keys.

For the criterion measures, teacher ratings were entered as raw scores. For the ITBS, we conducted analyses using national percentile rank scores. Math grades were represented in numerical form using the traditional 4.0 scale (e.g., A = 4.0, B = 3.0). We computed Pearson product moment correlation coefficients in our analyses of reliability and criterion validity. Regression analyses were used to examine slopes produced when the parallel forms of the measures were administered repeatedly over time.

Results

This section begins with a summary of the data collected over the 4-month period. Adherence to the agreed upon schedule was inconsistent across teachers. Table 2 summarizes the data collected each month by teacher and by type of measure. As evident in Table 2, the study data include a range of number of administrations per month, with more frequent administration occurring among special education teachers for students with disabilities. In addition, the MBSP measures, which had been used previously with special education students in the district, were collected less frequently than the measures that were new to the teachers.

Table 2

Number of Middle School Mathematics Progress Monitoring Measures Administered Each Month by Teacher and Class Type

	T1 – GE 22-Gr7 0-IEP	T1 - AR 10-Gr7 3-Gr8 8-IEP	T2 - GE 1-Gr7 20-Gr8 0-IEP	T3 – SE 6-Gr8 6-IEP	T4 - SE 9-Gr7 8-IEP	T5 - SE 3-Gr5 5-Gr6 8-IEP	T6 - SE 3-Gr5 3-Gr6 5-IEP
MBSP-Comp							
January	1	2		2		2	2
February	1	1		1		2	2
March		3		3		2	1
April		2		2		2	
MBSP-ConApp							
January	1	2		2		1	2
February	1	2		3		3	3
March		3		3		2	1
April		2		2		2	
Basic Facts							
January		1	1	2	2	2	2
February	1	2	1	3	4	3	3
March		3	1	3	2	3	1
April		2		2	5	2	
Estimation							
January		1	1	2	2	2	2
February	2	2	1	3	4	3	3
March		3	1	3	2	3	1
April		2		2	5	3	
CQD							
January	1	2	1	2	1	2	2
February	1	2	1	3	4	3	3
March		3	1	3	2	3	1
April		2		2	5	2	
Missing Num							
January		1	1	2	2	2	2
February	1	2	1	3	4	3	3
March		3	1	3	2	3	1
April		2		2	5	3	

Note. Individual students with IEPs in general education classes may have completed more forms than their general education peers in the same class period. The numbers in Table 3 reflect whole class administrations of the measures. T=Teacher, GE = General education, SE = Special education, Gr = Grade, IEP = Individual Education Plan.

Descriptive statistics for all study measures are reported next. Means and standard deviations, along with kurtosis and skewness statistics, for each of the middle school mathematics progress monitoring measures by grade level are presented in Table 3. The data reveal wide inconsistencies in the number of students assessed each month on each of the measures.

In reviewing the data in Table 3, we considered the nature of the distributions produced on each of the middle school mathematics progress monitoring measures. We were particularly interested in any floor or ceiling effects and the size of the standard deviations. Although scores of zero were observed for the MBSP-ConAPP and Estimation tasks, only one or two students had obtained a score of zero in each instance.

We examined the standard deviations produced for each measure and found that the MBSP-Comp produced a much wider distribution of scores, with standard deviations substantially larger than those for the other measures. Among the remaining measures, the standard deviations were greater for the Basic Facts and Complex Quantity Discrimination measures than for the MBSP-ConApp, Estimation, or Missing Number measures. Skewness and kurtosis estimates were examined for all measures and generally fell within the expected parameters for normality. The exceptions to this pattern were primarily in Grades 5 and 6, where small sample sizes likely influenced the distributions. Direct comparisons of mean scores by grade level were not conducted, as sample sizes and group composition (proportion of at-risk and special education students) in each grade level varied widely.

Table 3

Descriptive Statistics for the Middle School Mathematics Progress Monitoring Measures

Grade 5							
Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
MBSP-Comp							
January	6	10.5	42.5	24.9	12.3	.57	-1.21
February	6	14.8	54.0	32.7	13.9	.45	-.14
March	6	12.0	38.0	24.5	9.6	.15	-1.20
April	3	22.0	34.5	28.8	6.3	-.80	--
MBSP-ConApp							
January	6	2.0	9.0	5.8	2.5	-.31	0.0
February	6	3.3	7.3	5.5	1.4	-.37	.37
March	6	5.5	12.0	7.8	2.1	1.80	4.13
April	3	5.0	7.0	6.2	1.0	-1.29	--
Basic Facts							
January	6	8.0	20.0	15.2	4.8	-.76	-1.29
February	6	14.0	22.5	18.3	3.5	-.29	-1.79
March	6	13.0	27.7	20.6	6.4	-.16	-2.51
April	2	21.5	22.0	21.8	.35	--	--
Estimation							
January	6	.5	5.5	2.5	2.2	.72	-1.87
February	6	1.5	6.2	3.1	1.7	1.44	2.47
March	6	.7	29.0	9.4	10.8	1.53	2.03
April	3	.5	4.0	2.8	2.0	-1.72	--
Complex Quantity Discrimination							
January	6	7.0	23.0	13.8	6.5	.36	-1.8
February	6	6.0	26.7	16.2	8.7	.08	-2.2
March	6	6.7	26.0	14.8	7.3	.55	-.56
April	3	9.0	10.0	9.3	.6	1.73	--
Missing Number							
January	6	1.0	6.0	3.2	2.1	.61	-1.84
February	6	3.0	7.3	5.2	1.5	-.17	-.17
March	6	1.0	18.5	8.0	6.8	.70	-.93
April	3	3.5	7.7	5.5	2.1	.36	--

Table 3 (continued)

Descriptive Statistics for the Middle School Mathematics Progress Monitoring Measures

Grade 6							
Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
MBSP-Comp							
January	8	21.5	72.5	41.1	18.2	.77	-.62
February	8	10.0	65.5	36.9	19.7	.09	-1.33
March	7	12.5	74.0	41.1	19.1	.49	1.44
April	4	10.0	74.0	34.8	30.0	.87	-1.24
MBSP-ConApp							
January	8	1.0	8.0	3.9	2.3	.53	-.11
February	7	2.5	6.0	3.7	1.3	.79	-.33
March	7	3.0	9.0	4.8	2.2	.79	2.06
April	4	1.5	8.0	5.0	2.9	1.01	-1.65
Basic Facts							
January	8	11.0	32.5	20.4	7.8	.28	-1.09
February	7	13.7	38.3	29.4	9.6	-.72	-1.00
March	7	13.0	45.3	31.7	10.9	-.74	1.59
April	4	16.0	42.5	30.3	12.0	-.32	2.62
Estimation							
January	8	.5	9.8	5.4	3.2	.11	-.89
February	7	1.8	10.2	5.6	3.0	.14	-1.08
March	7	3.0	11.5	7.0	3.5	.11	-1.83
April	3	0	6.0	3.4	3.1	-1.13	--
Complex Quantity Discrimination							
January	8	5.0	25.0	16.4	8.8	-.30	-2.2
February	7	7.5	30.7	20.8	8.7	-.33	-1.0
March	7	8.0	33.0	22.4	8.5	-.49	-.16
April	4	12.0	12.0	20.9	7.0	-.58	-.91
Missing Number							
January	8	6.0	15.0	9.1	3.1	.92	.28
February	7	5.0	11.5	9.2	2.3	-1.22	.43
March	7	5.0	14.7	8.3	3.9	.96	-.79
April	4	4.5	15.0	11.2	4.7	-1.51	2.30

Table 3 (continued)

Descriptive Statistics for the Middle School Mathematics Progress Monitoring Measures

Grade 7							
Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
MBSP-Comp							
January	32	24.0	202.0	91.5	38.5	.45	.61
February	32	28.5	172.0	86.2	37.3	.49	-.66
March	2	28.0	53.7	40.8	18.1	--	--
April	11	31.0	213.0	93.6	55.0	.87	.71
MBSP-ConApp							
January	32	0.0	15.0	8.0	3.7	-.25	-.32
February	31	2.0	15.0	9.8	3.4	-.68	-.36
March	2	4.0	6.7	5.3	1.9	--	--
April	11	7.0	21.08	10.6	4.5	.66	1.49
Basic Facts							
January	12	10.0	30.0	18.8	6.0	.36	-.67
February	41	6.5	60.0	26.0	11.7	1.12	1.40
March	12	5.7	38.0	22.2	10.1	.04	-.95
April	20	9.0	58.0	28.0	13.6	.83	-.10
Estimation							
January	12	1.8	24.0	9.9	7.5	1.06	-.17
February	42	1.3	23.5	10.6	5.7	.18	-.77
March	11	0.0	29.0	7.8	8.7	1.72	2.65
April	20	1.5	22.5	10.4	7.0	.56	-1.16
Complex Quantity Discrimination							
January	40	10.0	44.0	29.3	9.9	-.35	-1.2
February	41	5.0	44.0	29.9	11.0	-.64	-.93
March	12	9.0	43.0	22.5	11.6	.75	-.74
April	20	15.5	44.0	31.4	9.1	-.29	-1.32
Missing Number							
January	12	3.0	28.0	11.1	7.4	1.33	1.29
February	41	3.3	34.0	15.7	7.8	.52	-.33
March	12	3.3	24.0	12.7	5.8	.74	.55
April	20	5.5	37.0	15.2	9.4	1.02	-.10

Table 3 (continued)

Descriptive Statistics for the Middle School Mathematics Progress Monitoring Measures

Grade 8							
Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
MBSP-Comp							
January	8	25.0	97.0	64.1	26.2	-.36	-1.39
February	7	36.5	77.5	53.5	13.0	.87	1.49
March	8	28.0	94.0	57.6	22.2	.54	-.51
April	8	32.0	78.0	61.3	14.5	-1.18	1.72
MBSP-ConApp							
January	8	3.0	10.0	6.8	2.4	-.37	-1.09
February	8	2.0	14.0	8.5	4.4	-.40	-1.56
March	8	4.0	12.0	8.1	2.4	-.23	.51
April	8	3.5	12.5	8.6	3.5	-.45	-1.52
Basic Facts							
January	28	7.0	48.0	30.8	11.4	-.47	-.69
February	25	11.0	52.0	31.5	12.6	.14	-1.03
March	27	10.7	56.0	32.2	12.4	.14	-.81
April	8	12.5	35.0	20.3	7.0	1.46	2.45
Estimation							
January	27	0.0	28.5	15.2	7.8	-.33	-.77
February	27	2.6	32.5	15.4	8.0	.25	-.49
March	27	1.7	30.0	16.2	8.0	-.16	-1.0
April	8	0.0	11.8	7.0	3.9	-.92	.03
Complex Quantity Discrimination							
January	28	11.0	44.0	35.1	10.8	.44	.24
February	27	12.3	44.0	35.9	9.3	.45	.50
March	28	9.0	44.0	36.8	10.1	.44	1.77
April	8	8.0	40.5	29.1	12.0	.75	-.46
Missing Number							
January	28	4.0	35.0	18.2	8.5	.01	-.85
February	27	3.5	33.0	19.8	8.2	-.55	-.63
March	27	4.3	40.0	19.1	9.3	.57	.08
April	8	5.0	17.5	10.8	4.9	.43	-1.88

Descriptive statistics for the criterion measures are presented in Table 4. ITBS scores are reported in percentile ranks and teacher ratings as raw scores, ranging from 1 to 7. The district had administered Computation and Concepts and Applications probes to all students in April; those scores are included in Table 4.

Table 4

Descriptive Statistics for Criterion Variables

Grade 5									
	Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
ITBS	Problem Solving & Data Interpretation	6	4	69	24.00	25.70	1.41	0.91
	Concepts & Estimation	6	2	34	14.50	12.50	0.69	-0.82
	Computation	6	12	58	31.00	15.17	1.08	2.52
	Total Math	6	4	56	18.17	20.55	1.65	2.20
Teacher	Rating	6	1	3	1.83	0.98	0.46	-2.39
Math	Grade	6	2.00	3.67	2.95	0.77	-0.56	-2.0
District	April Comp probe	5	19	41	28.00	9.82	0.68	-2.39
District	April ConApp probe	5	4	11	7.60	2.97	-0.31	-2.27

Grade 6								
	Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
ITBS	Problem Solving & Data Interpretation	8	5	59	30.38	18.86	0.14	-1.11
	Concepts & Estimation	8	3	36	14.38	10.64	1.33	1.63
	Computation	8	3	58	19.38	18.81	1.35	1.75
	Total Math	8	2	49	21.25	14.45	0.77	1.14
Teacher	Rating	8	1	3	1.88	0.99	0.31	-2.36
Math	Grade	7	0.67	3.33	2.19	0.86	-0.73	0.87
District	April Comp probe	7	33	140	80.14	37.71	0.17	-0.44
District	April ConApp probe	7	2	12	5.57	3.60	1.00	0.17

Table 4 (continued)

Descriptive Statistics for Criterion Variables

	Grade 7							
	Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
ITBS	Problem Solving & Data Interpretation	42	6	95	50.05	27.02	-0.14	-1.22
	Concepts & Estimation	42	3	89	43.12	25.58	0.07	-1.04
	Computation	42	2	90	47.69	27.82	-0.01	-1.28
	Total Math	42	4	81	46.76	25.70	-0.20	-1.35
Teacher	Rating	41	1	6	3.44	1.53	-.005	-1.01
Math	Grade	41	0.67	3.67	2.20	0.85	-0.09	-0.64
District	April Comp probe	41	17	195	78.27	34.17	0.82	1.96
District	April ConApp probe	41	3	24	11.88	5.00	0.18	-0.19

	Grade 8							
	Measure	*N*	Min	Max	*M*	*SD*	Skew	Kurt
ITBS	Problem Solving & Data Interpretation	28	12	99	72.36	28.12	-1.17	-0.03
	Concepts & Estimation	28	2	97	69.82	33.60	-1.06	-0.59
	Computation	28	9	98	68.82	23.25	-1.23	1.13
	Total Math	28	5	99	71.75	31.42	-1.18	-0.24
Teacher	Rating	29	1	7	3.79	1.72	0.39	-0.57
Math	Grade	28	0.00	4.00	2.33	1.21	-0.50	-0.93
District	April Comp probe	28	49	195	119.61	37.45	0.14	-0.71
District	April ConApp probe	28	4	31	20.89	7.76	-0.73	-0.63

Note. Skew = skewness; Kurt = kurtosis.

Research Question 1: What levels of alternate-form reliability are produced by the middle school mathematics progress monitoring measures?

We examined alternate-form reliability by computing Pearson product moment correlation coefficients between the scores students obtained on consecutive weekly administrations of each of the six middle school mathematics progress monitoring measures. Because of the variability in administration schedules, the number of cases on which individual correlations were based varies dramatically. We opted to examine all week-to-week correlations for cases for which more than five students' data had been included. In Table 5 we report the results of the alternate-form reliability analyses, including the number and range of the obtained correlations, the median and range for the number of cases, and the number of coefficients that were above $r = .70$, as well as the number that were statistically significant ($p < .05$).

The results reveal substantially lower levels of alternate-form reliability for some of the measures than obtained in previous research, despite the use of cross-grade correlations (due to small sample sizes in Grades 5 and 6). The coefficients for the MBSP-ConApp and Estimation measures were especially troubling, with very few meeting a liberal criterion of $r > .70$. The most positive results were obtained for the MBSP-Comp and the CQD measures, with the Basic Facts and Missing Number measures producing moderate levels of alternate-form reliability.

Table 5

Alternate-Form Reliability of the Middle School Mathematics Progress Monitoring Measures

	Number of Correlations	Mean Correlation and Range	Median and Range for Number of Cases	Proportion of Coefficients Above .70	Proportion of Statistically Significant Coefficients
MBSP Comp	6	.83 .77 – .93	11 9 – 23	6 of 6	6 of 6
MBSP ConApp	8	.47 -.08 – .79	13 9 – 19	1 of 8	4 of 8
Basic Facts	12	.74 .45 – .92	17 7 – 25	9 of 12	10 of 12
Estimation	12	.62 .47 – 88	16 7 – 29	2 of 12	8 of 12
CQD	11	.84 .72 – .94	15 8 – 22	11 of 11	11 of 11
Missing Number	12	.70 .35 – .99	19 7 – 28	6 of 12	9 of 12

Note: CQD = Complex Quantity Discrimination.

Research Question 2: What levels of criterion validity are produced by the middle school mathematics progress monitoring measures?

The initial examination of criterion validity included the full sample; mean scores (across a month) for each type of probe were correlated with the five criterion variables (teacher rating, grade in mathematics, Math Total score on the ITBS, and April scores on district-administered MBSP-Comp and MBSP-ConApp measures). The results of these initial analyses are reported by month in Table 6, using the mean scores for probes administered in each month across the full sample.

Table 6

Full Sample Criterion Validity Coefficients for the Middle School Mathematics Progress Monitoring Measures by Month

	MBSP-Comp	MBSP-ConApp	Basic Facts	Estimation	Complex Quant. Dis.	Missing Number
January Mean Probe Score	*N* = 51 - 54	*N* = 51 - 54	*N* = 51 - 54	*N* = 51 - 53	*N* = 79 - 82	*N* = 52 - 54
Teacher Rating	.71**	.60**	.63**	.57**	.57**	.64**
Math Grade	.18	.38**	.39**	.30*	.21	.25
ITBS-Math Total	.54**	.44**	.73**	.74**	.70**	.74**
April District MBSP-Comp	.73**	.49**	.79**	.69**	.64**	.75**
April District MBSP-ConApp	.52**	.56**	.85**	.77**	.73**	.82**
February Mean Probe Score	*N* = 50 - 53	*N* = 49 - 52	*N* = 76 - 79	*N* = 79 - 82	*N* = 78 - 81	*N* = 78 - 81
Teacher Rating	.72**	.63**	.55**	.54**	.49**	.65**
Math Grade	.19	.07	.35**	.19	.08	.29**
ITBS-Math Total	.44**	.50**	.42**	.78**	.65**	.69**
April District MBSP-Comp	.71**	.49**	.70**	.70**	.63**	.70**
April District MBSP-ConApp	.51**	.70**	.53**	.76**	.67**	.73**
March Mean Probe Score	*N* = 21 - 23	*N* = 21 - 23	*N* = 50 - 52	*N* = 49 - 51	*N* = 51 - 53	*N* = 50 - 52
Teacher Rating	.21	.23	.53**	.52**	.61**	.63**
Math Grade	-.57**	-.21	.33*	.35*	.15	.27
ITBS-Math Total	.25	.24	.52**	.71**	.77**	.68**
April District MBSP-Comp	.58**	-.04	.69**	.73**	.79**	.73**
April District MBSP-ConApp	.80**	.75**	.61**	.82**	.84**	.77**

Table 6 (continued)

Full Sample Criterion Validity Coefficients for the Middle School Mathematics Progress Monitoring Measures by Month

	MBSP-Comp	MBSP-ConApp	Basic Facts	Estimation	Complex Quant. Dis.	Missing Number
April Mean Probe Score	*N* = 25 - 26	*N* = 25 - 26	*N* = 33 - 34	*N* = 32 - 34	*N* = 34 - 35	*N* = 33 - 35
Teacher Rating	.82**	.63**	.66**	.74**	.48**	.65**
Math Grade	.24	.08	.35*	.14	-.28	.15
ITBS-Math Total	.41**	.64**	.20	.55**	.53**	.45**
April District MBSP-Comp	.83**	.41*	.70**	.59**	.56**	.59**
April District MBSP-ConApp	.58**	.67**	.43**	.58**	.64**	.60**

Note. * = $p < .05$; ** = $p < .01$. Complex Quant. Dis. = Complex Quantity Discrimination.

The data in Table 6 indicate that the middle school mathematics progress monitoring measures produced the lowest relations with grades. This result is not surprising, as grades often include factors beyond academic achievement, such as work habits, motivation, or effort. Across the other middle school mathematics progress monitoring measures, many strong relations (correlation coefficients exceeding 0.70) were observed. The MBSP-Comp was most strongly related to teacher ratings and the district-administered MBSP measures given in April. Surprisingly, the MBSP-ConApp, which has demonstrated strong relations to criterion measures in other studies, did not demonstrate the same magnitude of correlation as did the other measures. Among the remaining measures (Basic Facts, Estimation, Complex Quantity Discrimination, and Missing Number), the strongest relations were with the ITBS Total Math score and the April district-administered MBSP measures. This pattern was quite consistent across scores for January, February, and March, however the April coefficients were substantially lower than those obtained in earlier months. Readers should note the varying sample sizes across the four months.

To examine the degree to which the relations between the middle school mathematics progress monitoring measures and the criterion measures might be affected by grade level, the February scores (representing the greatest number of students) were disaggregated by grade level. The results of these analyses are reported in Table 7.

Table 7

February Criterion Validity Coefficients for the Middle School Mathematics Progress Monitoring Measures by Grade

	MBSP-Comp	MBSP-ConApp	Basic Facts	Estimation	Complex Quant. Dis.	Missing Number
Grade 5 (*N* = 5 - 6)						
Teacher Rating	.83*	.67	.57	.79	.66	.49
Math Grade	-.75	-.43	-.09	-.42	-.75	-.57
ITBS-Math Total	.68	.79	.25	.50	.79	.74
April District MBSP-Comp	.04	.59	-.80	-.27	.80	.80
April District MBSP-ConApp	.75	.96*	-.37	.80	.89*	.80
Grade 6 (*N* = 7 - 8)						
Teacher Rating	.36	.35	.24	.40	.77*	-.16
Math Grade	.01	-.06	.14	-.58	-.68	.02
ITBS-Math Total	.62	-.22	.06	.43	.80*	-.50
April District MBSP-Comp	.66	-.21	.51	.65	.84*	-.22
April District MBSP-ConApp	.86*	.65	.36	.45	.71	.10
Grade 7 (*N* = 31 - 41)						
Teacher Rating	.58**	.61**	.58**	.50**	.30	.58**
Math Grade	.45*	.41*	.35*	.07	.00	.22
ITBS-Math Total	.00	.27	.22	.62**	.31	.41**
April District MBSP-Comp	.76**	.60**	.69**	.55**	.40*	.61**
April District MBSP-ConApp	.19	.53**	.33*	.57**	.38*	.53**

Table 7 (continued)

February Criterion Validity Coefficients for the Middle School Mathematics Progress Monitoring Measures by Grade

	MBSP-Comp	MBSP-ConApp	Basic Facts	Estimation	Complex Quant. Dis.	Missing Number
Grade 8 (*N* = 7 - 8 for MBSP; 25 – 27 for others)						
Teacher Rating	-.34	-.17	.59**	.41*	.42*	.59**
Math Grade	-.01	.03	.59**	.53**	.61**	.67**
ITBS-Math Total	.85**	.44	.59**	.79**	.82**	.87**
April District MBSP-Comp	.50	.57	.74**	.74**	.80**	.78**
April District MBSP-ConApp	.94**	.71*	.76**	.79**	.85**	.89**

Note. * = $p < .05$; ** = $p < .01$. Complex Quant. Dis. = Complex Quantity Discrimination.

In considering the data in Table 7, readers are reminded of the small sample sizes for Grades 5 and 6, as well as the high proportion of students with disabilities in these samples. The Grade 7 and 8 samples are larger and include greater numbers of students not receiving special services. Not surprisingly, relations within the grade ranges were smaller in magnitude than those for the full sample. Few of the relations for students in Grades 5 and 6 demonstrated statistical significance; of those that did, the majority involved either the Complex Quantity Discrimination measure or the MBSP-Comp measure. For Grade 7, the MBSP-ConApp, Estimation, and Missing Number measures produced the greatest number of significant relations. In Grade 8, students' scores on the middle school mathematics progress monitoring measures were less correlated with teacher ratings than with the other criterion measures. In general, the Missing Number measure demonstrated the strongest relations with the criterion measures.

Research Question 3: To what extent do the measures reflect changes in student performance over time?

The final research question examined the degree to which the measures reflected change in student performance over time. To address this question, we computed ordinary least squares regression coefficients for each measure using individual data. We eliminated any students who did not have at least three data points across the four-month period for a particular measure. Data collection dates spanned a 15-week period from mid-January through the end of April. Table 8 shows the means and standard deviations for slopes by measure for the entire sample and for the data disaggregated by class type (general education, at risk, special education). Across the full sample, the Complex Quantity Discrimination measure was the most sensitive to growth, followed by the Basic Facts measure. This pattern held for the at risk student group, but not for the general education and special education student subgroups. Among general education

Table 8

Weekly Slope Data for the Middle School Mathematics Progress Monitoring Measures

	Full Sample Mean (SD) *N*	General Education Mean (SD) *N*	At Risk Mean (SD) *N*	Special Education Mean (SD) *N*
MBSP Comp	-.05 (2.03) 32	-- -- 1	-0.48 (1.63) 18	0.64 (2.45) 13
MBSP ConApp	0.21 (0.20) 32	-- -- 1	0.16 (0.13) 18	0.25 (0.24) 13
Basic Facts	0.33 (0.78) 49	0.17 (1.17) 17	0.45 (0.50) 27	0.20 (.13) 5
Estimation	0.21 (0.43) 51	0.29 (0.43) 19	0.19 (0.43) 27	0.11 (0.50) 5
CQD	0.43 (0.58) 61	0.11 (0.41) 21	0.62 (0.56) 27	0.54 (0.67) 13
Missing Number	0.19 (0.67) 51	0.11 (0.96) 19	0.26 (0.45) 27	0.06 (0.22) 5

Note: CQD = Complex Quantity Discrimination.

students, the most growth was observed on the Estimation measure, however the rates of growth were notably less than those obtained for students in the other two subgroups. Among students receiving special education services, the MBSP-Comp was most sensitive to growth, followed by the Complex Quantity Discrimination measure. Readers should note that in this district, the MBSP-Comp was routinely used for monitoring the progress of students toward IEP goals in mathematics, so the students were likely very familiar with the format and content of this particular measure. The negative growth trend on the MBSP-Comp measure for the at risk students is noteworthy. Anecdotal reports suggested that students found the measure frustrating and tedious, which may have caused them to stop putting forth their best effort as the study went on, producing the downward trends for growth. In general, the results suggest the middle school mathematics progress monitoring measures do reflect changes in student performance over time and could be used to monitor student progress. Practitioners seeking the most sensitive measure would be advised to use the Complex Quantity Discrimination, which produced strong growth rates (above 0.50 points/week) among students in the at risk and the special education groups, which represent the student populations most likely to be considered for progress monitoring.

Discussion and Future Research

The results of this study offer tentative support for the use of the middle school mathematics measures for progress monitoring. The present study adds to the current literature in this area by using a weekly data collection schedule (rather than fall/winter/spring benchmarking data, see Foegen, 2008) and naturalistic data collection conditions. With respect to alternate form reliability, we found that teacher-administered data collection, with minimal supervision by the research team, resulted in lower levels of alternate form reliability than has been obtained in previous studies where researchers were responsible for data collection. In particular, the Estimation and MBSP-ConApp produced particularly low reliability coefficients. The strongest reliability was observed for the Complex Quantity Discrimination and MBSP-Comp measures.

With respect to criterion validity, the results of the present study revealed many correlation coefficients in the moderate to strong range (r = .60 to .80) between the middle school mathematics progress monitoring measures and typical criterion variables (grades, teacher ratings, high stakes achievement test scores). In general, the Basic Facts, Estimation, Missing Number, and Complex Quantity Discrimination measures tended to have stronger relations with the criterion variables across the four-month period than did the MBSP measures. This was notable, as in past research the criterion validity coefficients for the MBSP-ConApp have been stronger than those obtained for the competing measures. When the criterion validity analyses were disaggregated by grade level, small sample sizes in grades 5 and 6 limit conclusions that can be drawn at those grade levels. In grades 7 and 8, the Missing Number measure most consistently resulted in moderate to strong relations with the criterion variables.

The extent to which the middle school mathematics progress monitoring measures were sensitive to student growth over time was the primary focus of the study. The results were encouraging, with the Complex Quantity Discrimination and Basic Facts measures demonstrating the highest mean slope values across the sample. The MBSP-Comp produced mean slope values exceeding 0.6 points/week for students receiving special education services, but it is unclear to what extent this result is influenced by the routine use of this measure (in contrast to the other measures) in the district.

Overall, the results across our examinations of alternate-form reliability, criterion validity, and sensitivity to growth converge to provide tentative support for the Complex Quantity Discrimination measure as an efficient tool for monitoring growth in mathematics among middle school students with or at risk for disabilities. These results must be interpreted with caution given the relatively small sample size, irregular data collection schedule, and disproportionate representation of students with or at risk for disabilities in the full sample. Future research should replicate the study with greater oversight of teacher administration of the measures (and observation of administration fidelity), a larger sample that includes greater numbers of students across the fifth to eighth grade range, and with more diverse criterion measures, including an external measure of growth against which progress monitoring slope data could be compared.

References

Foegen, A. (2000). Technical adequacy of general outcome measures for middle school mathematics. *Diagnostique, 25*, 175-203.

Foegen, A. (2008). Progress monitoring in middle school mathematics: Options and issues. *Remedial and Special Education, 29*, 195-207.

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

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

Foegen, A., Klein, K., Lind, L., & Jiban, C. (2008). *Reliability, criterion validity, and changes in performance across three points in time: Exploring progress monitoring measures for middle school mathematics.* (RIPM Technical Report 15). Minneapolis, MN: Research Institute on Progress Monitoring, University of Minnesota.

Foegen, A., & Lind, L. (2009). *A replication of static use of six brief middle school mathematics measures.* (RIPM Technical Report 21). Minneapolis, MN: Research Institute on Progress Monitoring, University of Minnesota.

Fuchs, L. S., Hamlett, C. L., & Fuchs, D. (1999). *Monitoring basic skills progress: Basic math concepts and applications.* Austin, TX: Pro-Ed.

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

Hoover, H. D., Dunbar, S. B., & Frisbie, D. A. (2001). *Iowa Tests of Basic Skills*. Chicago: Riverside Publishing Company.

Salvia, J., & Ysseldyke, J. E (with Bolt, S.). (2007). *Assessment in special and inclusive education (10th Ed.).* Boston, MA: Houghton Mifflin Company.

Appendix A

Researcher-developed math measures

Estimation

Complex Quantity Discrimination

Basic Facts

Missing Number

Estimation Probe 1

A large pizza has 16 slices. 5 kids will share the pizza. About how many slices will each kid get? 0.3 3 30	22 x 59 is about 12 120 1,200	It takes 2 1/2 yards of fabric to make a costume for the play. Mom has 11 yards of fabric. About how many costumes can she make? 0.4 4 40	$8\overline{)555}$ is about 0.7 7 70	The gym shoes cost $82. They are on sale for 25% off. About how much will you save? $20 $200 $2,000
4 x 9.3 is about 0.36 3.6 36	Edward makes $4 per hour doing odd jobs. If he works 11 hours, about how much will he earn? $4 $40 $400	0.45 - 0.14 is about 0.3 3 30	Luis wants to buy 6 new books. If each one costs $12, about how much will Luis pay for the 6? $0.60 $6 $60	$8\overline{)0.19}$ is about 0.02 0.2 2
Christine's car went 300 miles on 11 gallons of gas. About how many miles per gallon did the car go? 27 270 2,700	97.7 - 21.4 is about 0.8 8 80	There are 30 students in the class. Each student paid $3.50 for the bus. About how much money do they have for a bus? $10 $100 $1,000	73 - 18 is about 5 50 500	Joel earns $4.25 per hour stacking shelves at the grocery store. About how much will he earn in 20 hours? $0.80 $8 $80
78 + 17 is about 10 100 1,000	The car's gas tank holds 14 gallons. You just pumped in 11.75 gallons to make the tank full. About how many gallons were already in the tank? 2 20 200	219 + 876 is about 10 100 1,000	For her birthday, Sue received $19 from Grandpa and $32 from Aunt Sue. About how much did she receive from these two people? $5 $50 $500	$82\overline{)71.3}$ about 1 10 100

Complex Quantity Discrimination Form A

50 + 9 ☐ 9 + 50	18 ☐ 2 • 9	67 ☐ 76	.3 ☐ .03
56 ☐ 7 • 8	.5 ☐ .2	50 • 7 ☐ 5 • 70	15 ☐ 15 ÷ 3
32 ÷ 8 ☐ 32 ÷ 5	60 + 2 ☐ 20 + 6	60 • 4 ☐ 6 • 40	10 ☐ 10 − 7
70 + 3 ☐ 72 + 3	20 − 8 ☐ 20 − 9	85 ☐ 80	79 ☐ 70 + 9
53 ☐ 35	.4 ☐ .3	10 − 2 ☐ 13 − 2	$\frac{1}{6}$ ☐ $\frac{1}{2}$
70 − 1 ☐ 71 − 1	50 ☐ 50 + 1	29 ☐ 20	26 ☐ 2 + 6
$\frac{1}{5}$ ☐ $\frac{1}{9}$	90 ☐ 91 − 1	90 ☐ 90 + 2	20 ☐ 20 − 8
9 ☐ 90	48 ÷ 6 ☐ 48 ÷ 9	70 • 80 ☐ 70 • 8	80 + 4 ☐ 40 + 8
3 • 20 ☐ 3 • 2	15 ☐ 15 ÷ 5	74 ☐ 7 + 4	65 ☐ 60 + 5
1 • 90 ☐ 1 • 9	.9 ☐ .09	30 ☐ 34 − 4	80 − 9 ☐ 80 − 4
3 ☐ 30	40 • 50 ☐ 40 • 5	10 + 9 ☐ 9 + 10	60 + 4 ☐ 61 + 4

Project RIPM, Iowa State University

Facts Probe 1

1 - 1 =	8 x 1 =	5 x 5 =	1 x 7 =
4 x 7 =	5 x 7 =	4 + 6 =	9 x 5 =
3 + 0 =	3)12 =	2)14 =	6)6 =
12 - 9 =	7 + 4 =	0 x 7 =	7 - 4 =
5)10 =	8)48 =	11 - 7 =	4)12 =
8 - 2 =	9 + 6 =	6 + 6 =	1 x 2 =
8 + 7 =	0 x 0 =	11 - 2 =	8 - 5 =
6 - 2 =	7 + 0 =	3 + 3 =	17 - 9 =
10 - 4 =	9 x 9 =	4)4 =	1)5 =
1)1 =	2 -2 =	5 + 9 =	7 x 8 =
6)54 =	9 - 3 =	4)32 =	16 - 7 =
4 + 5 =	14 - 9 =	7 + 6 =	2 x 6 =
8 + 8 =	13 - 6 =	2 x 4 =	5)0 =
1 + 0 =	6 x 2 =	2 + 8 =	1 + 8 =
9)63 =	3)27 =	3)15 =	9)36 =
0 + 0 =	8 x 3 =	8 + 5 =	7)42 =
13 - 8 =	6)24 =	2 x 2 =	2 - 0 =
9 + 1 =	6 - 3 =	0 + 7 =	3 x 5 =
8)8 =	4 x 9 =	9 - 7 =	5)40 =
5 + 2 =	7 - 0 =	1 x 6 =	8 + 0 =

Missing Number Form A

37, 40, 43, ___	11, 22, 44, ___	43, ___, 45, 46	12, ___, 24, 30
30, 35, 40, ___	32, 28, ___, 20	21, 16, 11, ___	50, 60, ___, 80
48, 40, 32, ___	25, 20, 15, ___	24, ___, 12, 6	35, ___, 37, 38
18, 16, 14, ___	19, 20, 21, ___	6, ___, 12, 15	8, 16, 32, ___
80, 79, ___, 77	56, 57, ___, 59	88, 44, 22, ___	27, ___, 21, 18
24, 12, 6, ___	36, ___, 54, 63	14, 16, 18, ___	56, 49, ___, 35
70, 60, ___, 40	14, 21, ___, 35	49, 45, 41, ___	28, 25, 22, ___
10, 9, 8, ___	43, 45, 47, ___	47, 57, 67, ___	38, 42, 46, ___
2, 4, 8, ___	28, ___, 26, 25	7, 12, 17, ___	80, 40, 20, ___
40, 48, 56, ___	12, 16, ___, 24	81, ___, 63, 54	48, 38, 28, ___
17, 15, 13, ___	56, 28, 14, ___	13, 26, 52, ___	65, 64, ___, 62

Project RIPM, Iowa State University

Appendix B

Teacher Rating Form

Teacher ______________________

Middle School Math Progress Monitoring Study
XXX Middle School
Spring 2007

Directions: For each student, rate his or her general proficiency in mathematics relative to other students in your class(es). A rating of "1" indicates a very low level of proficiency, "4" indicates average proficiency, and "7" indicates exceptional proficiency. Try to spread student ratings across the full range of the scale, not clustering students only in the middle or toward one end.

Student		Mathematics Proficiency						
Last Name	First Name	Low			Average			High
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7
		1	2	3	4	5	6	7