PROJECT AAIMS: ALGEBRA ASSESSMENT AND INSTRUCTION-MEETING STANDARDS ALGEBRA ASSESSMENT AND INSTRUCTION MEETING STANDARDS


Effects of Teachers' Access to Student Data on Algebra Progress

Technical Report \#15

Anne Foegen, Ph.D.<br>Jeannette Olson, M.S.<br>Iowa State University

December 2007

Project AAIMS is funded by the U.S. Department of Education, Office of Special Education Programs, Grant \# H324C030060


#### Abstract

This studies extended previous research in Project AAIMS by exploring the effects of teachers' access to student data on students' rates of growth in algebra. In addition, the study also replicated previous research examining the technical characteristics (reliability and criterion validity of the measures). The present study was conducted during the fall semester of the 2006-07 academic year and included 168 algebra students in Districts A and B. Students completed four algebra progress monitoring measures each month; all of these probes were administered by students' teachers during class time and scored by Project AAIMS staff. Class periods were matched by class type and assigned to one of two conditions. In the "No Data" condition, teachers (and students) had no access to any of the progress monitoring data until the end of the semester when individual graphs of student performance were shared with both teachers and students. In the "Data" classes, project staff met monthly with teachers to share graphs of individual student data and performance summaries for the class that reported the accuracy of students' responses to the different question types represented in each probe. Analyses of students' slopes in each of these conditions revealed no differences between the rates of growth for students in the two conditions. We hypothesize that several factors contributed to this result. First, the use of project staff for scoring created delays in processing the data, so that the data shared during data conferences had often been gathered 3 to 4 weeks prior. In addition, teachers commented that the performance summaries provided useful information about the difficulties students were experiencing. As a result, they offered review and remedial work to all of their classes (including the "No Data" class) to address these concerns. Future studies will employ strategies to increase teachers' interaction with student data, including scoring student probes and entering student data.


## Introduction

Previous work in Project AAIMS has established the reliability and criterion validity of two measures (Basic Skills and Content-Analysis-Multiple Choice) for monitoring student progress in algebra. In Technical Reports 10, 12, 13, and 14 we reported the technical features of the measures when used for static measurement of student performance. We found that these two measures possess acceptable levels of alternate form reliabilities and moderate levels of criterion validity. Technical Reports 12, 13 , and 14 also reported examinations of students' growth over time on the measures, with the Content Analysis-Multiple Choice probes more sensitive to change than the Basic Skills probes. In the current study, we conducted an additional replication of the technical features of the measures. We also examined the extent to which teachers' access to student performance data was associated with differential rates of growth on the measures.

## Method

The study described in this report was conducted from September 2006 to January 2007 in Districts A and B. District A serves four small towns as well as the rural
agricultural areas between the towns. Approximately 7,000 residents reside in the school district. During the 2006-07 academic year, the junior/senior high school had an enrollment of approximately 670 students; approximately 12 percent of these students received special education services. Eighteen percent of the district's students were eligible for free and reduced lunch; three percent were of diverse backgrounds in terms of race, culture and ethnicity. No students in District A were reported as English Language Learners in 2006-07. District B is located in a community of 26,000 people; the high school (grades 9-12) enrolled nearly 1,300 students during the 2006-07 school year. Forty-seven percent of the district's students were eligible for free and reduced lunch; $18 \%$ were of diverse backgrounds with regard to race, culture, and ethnicity. Approximately $15 \%$ of District B's students received special education services. In both districts, all data collection activities involving students were completed during regular class time. Teachers administered all algebra probes.

## Participants

Student Participants. One hundred sixty-eight students participated in the study, 93 from District A and 75 from District B. Written parental/guardian consent and written student assent were obtained for all of these students using procedures approved by Iowa State University's Human Subjects Review Committee. Students were only included in the data analysis if they had a minimum of four data points for each type of probe administered by their teachers. Descriptions of the participating students are provided in Table 1. Readers should note that demographic data were not available for five students in District B who had dropped the course. In addition, no data on free and reduced lunch status was available for participants in District B. In previous years, this percentage has been approximately $37 \%$ of the district population.

Table 1. Demographic Characteristics of Student Participants by Grade Level

|  | District A |  |  |  |  | District B |  |  |  | $\frac{\text { Total }}{\underline{\mathrm{N}}}$ | $\frac{\text { Total }}{\underline{\%}}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\underline{8}$ | $\underline{9}$ | $\underline{10}$ | 11 | $\underline{12}$ | $\underline{9}$ | $\underline{10}$ | $\underline{11}$ | $\underline{12}$ |  |  |
| Gender |  |  |  |  |  |  |  |  |  |  |  |
| Male | 9 | 27 | 4 | 1 | 2 | 12 | 13 | 5 | 5 | 78 | 48 |
| Female | 13 | 27 | 9 | 0 | 1 | 20 | 9 | 5 | 1 | 85 | 52 |
| Ethnicity |  |  |  |  |  |  |  |  |  |  |  |
| White | 22 | 52 | 13 | 1 | 2 | 28 | 19 | 10 | 6 | 153 | 94 |
| Black | 0 | 1 | 0 | 0 | 0 | 3 | 3 | 0 | 0 | 7 | 4 |
| Hispanic | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | $<1$ |
| Asian | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | $<1$ |

Lunch
Free/Red $\quad 3 \quad 10 \quad 4 \quad 0 \quad 0 \quad$ -
$\begin{array}{llllllllllll}\text { Disability } & 0 & 6 & 6 & 0 & 0 & 5 & 2 & 2 & 3 & 24 & 15\end{array}$
Note. Calculation of the total column was based on the 163 participants for whom demographic data were available.

As the data in Table 1 indicate, a large majority of the participants (94\%) were white and more than half (53\%) were in ninth grade, the traditional grade in which students in these districts complete algebra. Fifteen percent of the participating students were identified as having disabilities and receiving special education services.

The students participating in the study were enrolled in one of four types of algebra classes. A total of 67 students were participating in a traditional Algebra 1 course taught using a conventional time frame (one year for schools such as District A with 45 minute periods, and one half year for school such as District B, using block scheduling with 90 minute periods). Of these, 22 were 8th grade students in District A completing a high school algebra course; these students, who comprised a single class, were identified as advanced in mathematics within their district. The remaining 45 students were enrolled in one of four different sections of Algebra 1. All of the traditional Algebra 1 instruction occurred in District A. Sixty-eight students were enrolled in one of six sections of Algebra 1A. This course spans the same amount of time as Algebra 1, but covers only the first half of the traditional Algebra 1 content. The intent is to provide additional time for mastering algebra content and skills for students who may require this alternative. Thirtythree students were enrolled in one of two sections of Algebra 1B, the companion course to Algebra 1A that addresses algebra content typically taught in the second half of a traditional Algebra 1 course. Readers should note that in District A, the courses previously titled Algebra 1A and Algebra 1B were renamed "Pre-Algebra" and "Basic Algebra" for the 2006-07 academic year. For consistency across districts and with previous Project AAIMS technical reports, we have opted to maintain the Algebra 1A and 1 B language in this report.

Across all participants, the average national percentile rank scores on the Iowa Tests of Educational Development were 47.7on the Problems/Data Subtest and 45.8 on the Computation Subtest. For the 22 8th grade students enrolled in Algebra I, the mean scores on the Iowa Tests of Basic Skills were 91.4, 84.8, and 89.0 on the Concepts/Estimation and Problems/Data subtests and the Mathematics Total scores, respectively. Readers should note that the 8th grade students were considered advanced students in mathematics.

Additional Information on Students with Disabilities. Because the applicability of the algebra probes to students with disabilities is an important part of Project AAIMS, additional information about the 24 students with disabilities participating in the project is provided in Table 2.

## Table 2. Descriptive Information on the Programs of Students with Disabilities Characteristic Quantification <br> Disability category <br> 24 Entitled Individual (EI)

$\%$ time in general education

$$
\text { Range }=76-100 \% ; \text { Mean }=93 \%
$$

\# of students with math goals

In algebra, students with disabilities earned mean grades of 1.75 [D+] (range 0.00 [F] to $4.00[\mathrm{~A}]$ ). In Districts A and B, the Iowa Tests of Educational Development are used as a district-wide assessment for students in grades 9-12. On average, students with disabilities obtained national percentile rank scores of 31 and 31 on the Problems/Data and Computation subtests, respectively. The 8th grade students in District A were assessed with the Iowa Tests of Basic Skills, but none of these students had disabilities.

Teacher Participants. Two general education teachers and one special education teacher from District A and three general education teachers and two special education teachers from District B were the primary teacher participants in the study. These teachers administered the algebra progress monitoring measures and were responsible for designing and delivering instruction on a daily basis in their classes. All teachers were teaching or co-teaching general education classes in algebra and participated in the data conferences (described later) to review student performance.

The two general education teachers from District A held standard Iowa teacher's licenses with 7-12 Mathematics endorsements. One of these teachers had five years of teaching experience with three years of experience teaching algebra classes. This teacher had earned a Bachelor's degree and had completed some graduate work. The second general education teacher had completed a Master's degree program and had six years of teaching experience with four years of that experience teaching algebra. The special education teacher held an initial Iowa teacher's license with a K-6 Elementary Teacher and a K-6 special education endorsement and had earned a Bachelor's degree and completed some graduate work. She had one year of experience that included teaching algebra.

All of the teachers from District B held standard Iowa teacher's licenses. All of the general education teachers had 7-12 Mathematics endorsements. The special education teachers had $7-12$ special education endorsements. Two of the general education teachers had two years of experience that included teaching algebra. One of these teachers had earned a Master's degree, while the other had earned a Bachelor's degree. The third general education teacher had 18 years of teaching experience with nine of those years teaching algebra. This teacher held a Bachelor's degree. One of the special education teachers had 15 years of experience and had earned a Master's degree and completed additional graduate work. The other special education teacher had two years of experience and had earned a Bachelor's degree and completed some graduate work.

## Measures

Algebra Progress Monitoring Measures. Two algebra measures were examined in this study; sample copies of each are provided in the Appendix. The following paragraphs summarize the characteristics of each of the two types of measures.

## Probe A: Basic Skills Measure

The Basic Skills measure is designed to assess the 'tool skills' that students need to be proficient in algebra. Just as elementary students' proficiency with basic facts is
associated with their ease in solving more complex problems, we hypothesize that there are some basic skills in algebra that serve as indicators of overall proficiency. In our discussions with teachers, they frequently commented that many students had difficulty with integers and with applying the distributive property. The items included in the Basic Skills measure address solving simple equations, applying the distributive property, working with integers, combining like terms and applying proportional reasoning. The Basic Skills probe includes many skills one would assume that students proficient in algebra would be able to complete with reasonable levels of automaticity. Students have five minutes to work on this probe; 12 parallel forms were used in the study. Each Basic Skills probe consists of 60 items; each item is scored as one point if it is answered correctly.

## Probe E: Content Analysis-Multiple Choice Measure

The Content Analysis-Multiple Choice measure consists of 16 items that correspond to the first eight chapters in the textbook that is used in the district. Problems are placed in random order on each probe. Students are directed to circle the correct response from four alternatives and to show their work unless they are confident they can solve the problems mentally. Twelve parallel forms of the measure were used in this study. Students have seven minutes to work on the Content Analysis-Multiple Choice probes. Scoring for the Content Analysis-Multiple Choice probes is done by comparing student responses to a rubric-based key created by the research staff. Each of the 16 problems is worth up to three points. Students earn full credit (three points) by circling the correct answer from among the four alternatives. If students circle an incorrect response and do not show any work, their answer is considered a 'guess;' the total number of guesses is recorded for each probe. In cases where students show work, the scorer compares the student's work to the rubric-based key, and determines whether the student has earned 0,1 , or 2 points of partial credit. The number of points earned across all 16 problems and the number of guesses are recorded and entered in the data files. A final score is computed by subtracting the number of guesses from the total number of points earned on the probe.

Criterion Measures. In order to replicate previous criterion validity analyses involving the algebra progress monitoring measures, we gathered data on a variety of other indicators of students' proficiency in algebra. Some of these measures were based on students' performance. Other measures reflected students' performance on standardized assessment instruments. The classroom-based measures included gradebased measures and teacher ratings. Each student's algebra grade, the grade s/he earned in algebra for the yearlong algebra course, was recorded using a four-point scale (i.e., $\mathrm{A}=$ $4.0, \mathrm{~B}=3.0$ ).

Student performance on standardized, norm-referenced assessments was evaluated using school records and with an algebra instrument administered as part of the project. In District A, students complete either the Iowa Tests of Educational Development (ITED) or the Iowa Tests of Basic Skills (ITBS) each year, depending on their grade level. Students in grades 9 to 12 complete the ITED, while students in grade 8 take the ITBS. District records were used to access students' scores on these instruments;
national percentile ranks were used for the analyses. We recorded the Concepts/Problems subtest score (which was identical to the Math Total score) and the Computation subtest score for the ITED and the Math Total score for the ITBS.

Growth Measure. We also sought to replicate previous findings related to students' growth on the measures over time. The growth measure in this study, which we called probe slope, reflects the growth that students showed on both types of probes over the semester. We used ordinary least square regression to calculate each student's slope on each measure. The obtained slope values were calculated to reflect the amount of weekly progress a student demonstrated on a probe type. Only students having at least four probe scores over the semester are included in this section of the study. If two probes were given within the same week, they were treated as given in the subsequent weeks. For example, if two probes (e.g. E-2, E-3) were given in the same week 7, E-2 probe was considered to given in week 7 and E-3 in week 8 . We opted to use each individual data point, rather than to average the scores and use only the mean because of the frequency with which the data were being collected. In addition, a graphing template used by the teachers was designed to record weekly data points.

## Procedures

Project AAIMS research staff visited each class at the beginning of the school year to present information about the study and gather informed consent. Students completed student assent forms during class and were given parent consent forms to take home. Teachers offered extra credit to students for returning signed consent forms (regardless of whether parents provided or withheld consent). Teachers administered four probes each month. Some teachers (primarily those in District B) opted to administer one measure each week to their students. In District A, the school schedule included approximately two days of early dismissal each month for staff development purposes. Teachers in District A preferred to use the shortened class periods on these days as opportunities to administer two forms of the same type of progress monitoring measure.

Teachers were allowed to choose which measure(s) to administer to their students. In addition to the Basic Skills and Content Analysis-Multiple Choice measures, teachers also had the option to select the Algebra Foundations, though none did. The most frequently selected measure was the Content Analysis-Multiple Choice measure. Table 3 shows the types of courses taught by each of the participating teachers, as well as the measures they chose to administer to their students.

Teachers 2 and 3 in District A planned from the beginning to alternate their use of the Basic Skills and Content Analysis-Multiple Choice measures. They expressed concerns that the students in their co-taught class, many of whom struggled to be successful in mathematics, might find the content of the Content Analysis-Multiple Choice measure to be overwhelming at the beginning of the algebra course. They balanced this concern by alternating between the two measures, administered two forms of each once each month. Teacher 2 in District B began the fall semester planning to use only the Content Analysis-Multiple Choice measure. As the term progressed, her students' frustration with the measure (which they found to be difficult) prompted her to

Table 3. Course Type and Probes Administered, by Teacher

| District | Teacher(s) | Period/Block | Course Type |  | Probe(s) |
| :--- | :--- | :--- | :--- | :--- | :--- |

Note. CA-MC = Content Analysis-Multiple Choice, BS= Basic Skills.
reconsider her decision. At the mid-semester point, she began alternating between the Content Analysis-Multiple Choice and Basic Skills measures each week.

Contrast Conditions. The primary purpose of this research was to conduct a pilot study examining the effects of teachers' access to student progress data on rates of student growth. Some researchers (C. Espin, personal communication) have suggested that merely providing teachers with data on their students' performance may be sufficient to prompt subtle instructional changes that will lead to improved student outcomes. In the present study, we selected two of each teacher's class periods and randomly assigned one to the treatment condition and the other to the comparison condition. Table 4 shows the assignment of class periods by teacher to each condition.

Students in all classes completed four algebra progress monitoring probes each month, following the schedule reported in Table 3. Classroom teachers administered the measures, which were scored by project staff, then entered the data into a graphing template and an error analysis summary spreadsheet. We opted to have scoring completed by project team members to preclude teachers from having any informal access to the data for any classes. In the class designated for the treatment condition (which we labeled the "data" classes), either the lead researcher or the project coordinator met individually with the teachers monthly to review the data for students in the selected period.

Table 4. Assignment to Treatment Conditions

| District | Teacher | Data (Treatment) | No Data (Comparison) |
| :--- | :--- | :--- | :--- |
| A | 1 | 3 | 2 |
|  | 1 | 7 | 6 |
|  | $2 \& 3$ | 5 | 7 |
|  |  |  |  |
| B | 1 | 1 | 2 |
|  | 2 | 4 | 2 |
|  | 3 | 2 | 3 |

The data reviewed in each conference included two elements: student scores with graphs and class-wide summaries of performance on different item types. The student scores/graphs portion of the data conference centered on an Excel spreadsheet showing each student's score on each probe, along with individual graphs (structured as tabs across the bottom of the spreadsheet) for each student. The graphs had a line with the student's data with a trendline imposed on it, along with a second line of comparison data representing the average score on the same probe for all students in the same district enrolled in the same class type (i.e., for students in Algebra 1 this line represented the mean for all students in the same district in Algebra 1). As with the individual student's data, the comparison data line also had a trendline imposed. A sample graph, showing a full year's worth of data, is depicted in Figure 1. The class-wide summaries provided a listing of the different types of items included in each measure, along with the percentage of students obtaining the correct answer, partial credit or incorrect answers, and the percentage who left the item blank. For the Content Analysis-Multiple Choice measure, the student performance data were broken down to show the percentage of students who obtained each of the possible point values (e.g., 3 points for correct responses, 0 points for showing work, but not having any elements of the solution correct). For each item type, the response option (correct, incorrect, blank, partial credit) selected by the largest percentage of students (e.g., the largest number of students) was noted with bold text. A sample of the class-wide summary report for a Content Analysis-Multiple Choice measure is included in the Appendix. As multiple forms of the measures were administered, we listed the four most recent sets of data for teachers' review.

At each data conference, the project staff member began by sharing an electronic version of the graphing template with the teacher and, where applicable, the co-teacher. Teachers were encouraged to begin by clicking through the sequence of graphs for each class, noting each student's slope (increasing, steady, or decreasing) and his/her performance relative to the average. Teachers were asked to comment on any "surprises" in the data. Next, the class-wide summaries were shared and the project staff member highlighted item types on which student performance seemed especially strong or weak. Teachers were asked to note which items had been taught at the point of the data conference and again to identify any aspects of the data that were surprising to them. To close the data conferences, teachers were asked to identify 2 to 3 students as "target students." The target students were to be those about whom the teacher was concerned.

Figure 1. Sample progress monitoring graph


Student 1- Content Analysis (Textbook) Progress


AAIMS Technical Report 15 - Page 9

Although we did not ask teachers to implement a specific intervention with these students, we did suggest that they might pay a bit of extra attention to the target students, by doing things such as noting whether target students were completing their work, or checking in with them during class activities to monitor their level of understanding. During the duration of this study, three data conferences (October, November, and December) were held. January data (representing the end of the course/semester) were shared with teachers, but data conferences were not held because of semester-end demands the teachers were facing.

Students in classes assigned to the comparison treatment (which we referred to as "no data" classes) completed probes on the same schedule as students in the "data" classes. The primary difference between the two conditions was the level of teachers' access to student data. Although teachers were able to review individual student and class-wide data on a monthly basis for their data classes, we withheld any student performance data for students in the no data classes until the end of the semester.

Scorer Training. Scoring of the algebra progress monitoring measures was completed by three pre-service teachers (subsequently referred to as "scorers") who were hired and trained to score the probes and enter the data into the respective summary formats (Excel graphing spreadsheet and class-wide summary templates). The hiring process included a demonstration of correct scoring procedures for each type of probe and guided practice activities in which scorers worked with actual student papers. A final activity was the independent scoring of 10 student papers for each of the probe types. We used these probes to evaluate scoring reliability. For each probe, an answer-by answer comparison was conducted and an interscorer reliability estimate was calculated by dividing the number of agreements by the total number of answers scored. These individual probe agreement percentages were then averaged across all the selected probes of a common type to determine an overall average. After training, the scorers' mean interscorer agreement rates were $99 \%$ for the Basic Skills probes (range $=96 \%$ to $100 \%$ ) and $95 \%$ for the Content Analysis-Multiple Choice probes (range $=81 \%$ to $100 \%$ ). Scorers were informed that we would be checking their scoring accuracy levels throughout the project; they were able to earn bonus pay for maintaining high levels (i.e., $>96 \%$ agreement) of accuracy in their scoring.

Following training, each scorer was assigned approximately five class periods with two forms of a probe per class to score (a total of 10 class sets of probes twice each month). Scorers also completed the data entry for the classes they were scoring. For each scorer, we conducted a scoring reliability on two of the class sets in each scoring period (i.e., twice each month) by re-scoring all of the probes in those sets.

Analyses of interscorer agreement rates revealed that scorers had high reliability on both types of probes. A total of 62 interscorer reliability checks were conducted across the three scorers throughout the 2006-07 school year. The range of agreement for Basic Skills probes was between $98.6 \%$ and $100 \%$ with a mean of $99.7 \%$. For Content Analysis-Multiple Choice probes, the interscorer agreement rates ranged from 94.8\% to $100 \%$, with a mean of $99.2 \%$.

## Results

In the following sections, we report the results of our analyses. Because of differences in the instructional context in each district (e.g., traditional vs. block scheduling), we conducted the analyses separately by district. We first report the descriptive data on student performance on the measures first, followed by analyses related to replications of the measures' reliability and criterion validity. We conclude this section with analyses of differential performance across the two contrast conditions (data and no-data) investigated in this study.

## Descriptive Data on Score Ranges and Distributions

In this section, we report the ranges, means, and standard deviations for each type of measure by the school week during which the measure was administered. Because specific data collection schedules varied, we did not track which specific measure was administered which week and the weekly data likely include multiple versions of the parallel forms of each type of measure. The data are reported separately by district and disagreggated by class type.

Content Analysis-Multiple Choice: Tables 5 and 6 list the ranges, means, and standard deviations for the Content Analysis-Multiple Choice probes by class type within each district. On the Content Analysis-Multiple Choice measure, the score represents the number of points earned on the probe (each of the 16 problems was worth up to 3 points. The total possible score was 48 . On occasion, students who demonstrated high rates of guessing earned computed scores that were negative. We converted all negative scores to zeros.

In District A, students' scores improved from the early portion of the semester to the later portion of the semester, with the exception of the Algebra 1B course, where scores tended to decline over the period of the study (the teachers of this course reported significant concerns with student motivation in this course. The rates of improvement were greatest (and most regular from week to week) in the 8th grade algebra course. The amount of improvement was consistent with expectations for types of course offerings. Students in the 8th grade algebra and Algebra 1 courses demonstrated more improvement than did students in Algebra 1A and 1B. Standard deviations ranged from 2.7 to 10.0, with most in the 4 to 8 point range. The magnitude of these values suggests that the measures provide sufficient dispersion of the scores, so that teachers should be able to discriminate between students performing at different levels.

In District B, the same pattern of improvement from Week 1 to Week 17 was observed, but there was much greater variability from week to week. Students in District $B$ tended to decline in average performance levels during the first weeks of the semester, then show improvements in Weeks 9 and 10, followed (in the 1B class) by another smaller decrease for a period of three to four weeks before additional gains at the end of the semester. Standard deviations in District B were generally in the range of 5 to 8 points.

Table 5. Descriptive Statistics for Content Analysis-Multiple Choice Measures in District A

|  | 8th Grade Algebra |  |  |  | Algebra 1 |  |  |  | Algebra 1B |  |  |  | Algebra 1A |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Wk | N | Range | Mean | SD | N | Range | Mean | SD | N | Range | Mean | SD | N | Range | Mean | SD |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 | 21 | 7-18 | 11.05 | 2.7 | 44 | 0-42 | 13.15 | 8.1 |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 | 21 | 4-18 | 11.90 | 3.9 | 41 | 3-37 | 13.54 | 6.9 |  |  |  |  |  |  |  |  |
| 5 | 21 | 3-18 | 10.95 | 4.4 | 45 | 0-45 | 11.38 | 8.1 | 6 | 11-24 | 19.17 | 5.0 | 17 | 0-18 | 7.82 | 5.5 |
| 6 | 22 | 10-23 | 14.45 | 3.4 | 43 | 0-41 | 13.70 | 7.8 | 6 | 6-29 | 20.50 | 8.3 | 17 | 4-23 | 10.00 | 5.2 |
| 7 | 22 | 5-23 | 14.45 | 5.2 | 43 | 0-48 | 14.33 | 8.6 |  |  |  |  |  |  |  |  |
| 8 | 22 | 8-23 | 15.00 | 4.7 | 44 | 0-45 | 11.64 | 8.8 |  |  |  |  |  |  |  |  |
| 9 | 21 | 8-26 | 17.19 | 4.4 | 43 | 0-48 | 16.74 | 8.1 | 7 | 4-22 | 14.00 | 7.3 | 19 | 0-19 | 6.58 | 5.8 |
| 10 | 20 | 10-29 | 19.05 | 5.1 | 38 | 0-48 | 16.79 | 9.5 | 7 | 0-25 | 14.71 | 8.3 | 19 | 2-22 | 10.00 | 5.8 |
| 11 | 21 | 4-28 | 19.19 | 5.2 | 43 | 0-48 | 16.88 | 8.1 |  |  |  |  |  |  |  |  |
| 12 | 21 | 9-28 | 19.95 | 5.5 | 43 | 0-47 | 17.00 | 8.5 |  |  |  |  |  |  |  |  |
| 13 | 22 | 7-33 | 22.27 | 7.2 | 41 | 0-48 | 18.15 | 9.8 | 6 | 5-20 | 13.83 | 5.3 | 19 | 0-23 | 11.63 | 6.5 |
| 14 | 22 | 18-39 | 28.32 | 5.8 | 41 | 0-48 | 20.44 | 10.1 | 6 | 9-30 | 17.50 | 8.1 | 19 | 2-21 | 10.58 | 5.9 |
| 15 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 16 | 22 | 17-43 | 31.59 | 6.0 | 42 | 4-41 | 22.95 | 9.3 |  |  |  |  |  |  |  |  |
| 17 | 22 | 22-42 | 31.68 | 5.4 | 42 | 0-45 | 24.07 | 10.0 |  |  |  |  |  |  |  |  |

Table 6. Descriptive Statistics for Content Analysis-Multiple Choice Measures in District B

|  | Algebra 1B |  |  |  |  |  | Algebra 1A |  |  |  |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :---: | :---: |
| wk | N | Range | Mean | SD | N | Range | Mean | SD |  |  |
| 1 |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  | 47 | $0-22$ | 7.62 | 5.3 |  |  |
| 3 |  |  |  |  | 15 | $0-22$ | 7.53 | 6.5 |  |  |
| 4 | 22 | $7-35$ | 20.27 | 6.5 | 48 | $0-24$ | 8.27 | 6.3 |  |  |
| 5 | 23 | $4-38$ | 19.22 | 8.7 | 41 | $0-16$ | 5.00 | 5.0 |  |  |
| 6 | 23 | $2-27$ | 14.91 | 7.4 | 14 | $2-16$ | 7.86 | 4.5 |  |  |
| 7 | 21 | $3-32$ | 18.57 | 7.3 | 32 | $0-18$ | 8.13 | 5.7 |  |  |
| 8 | 23 | $4-33$ | 19.48 | 8.5 | 16 | $0-15$ | 6.19 | 5.1 |  |  |
| 9 | 20 | $6-43$ | 20.05 | 8.3 | 41 | $0-24$ | 10.05 | 6.7 |  |  |
| 10 | 23 | $11-36$ | 25.00 | 6.6 | 24 | $0-24$ | 11.05 | 7.1 |  |  |
| 11 | 22 | $12-37$ | 22.41 | 6.9 |  |  |  |  |  |  |
| 12 | 21 | $4-34$ | 21.67 | 6.9 |  |  |  |  |  |  |
| 13 | 18 | $8-35$ | 22.11 | 7.2 | 25 | $0-28$ | 13.16 | 8.1 |  |  |
| 14 | 19 | $4-32$ | 20.37 | 8.1 | 23 | $2-29$ | 14.13 | 7.1 |  |  |
| 15 | 22 | $18-39$ | 27.09 | 5.0 | 21 | $0-31$ | 13.81 | 9.3 |  |  |
| 16 | 21 | $8-37$ | 29.38 | 7.2 |  |  |  |  |  |  |
| 17 | 20 | $11-41$ | 29.25 | 6.9 |  |  |  |  |  |  |

Basic Skills. Tables 7 and 8 report ranges, means and standard deviations for the Basic Skills measure. On this measure, students in all three classes showed improvement from the first to the final administration, but the magnitude of the improvement (number of problems correct) tended to be smaller than what we observed for the Content Analysis-Multiple Choice measure.

Table 7. Descriptive Statistics for Basic Skills Measures in District A
Algebra 1B

| wk | N | Range | Mean | SD | N | Range | Mean | SD |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 1 |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |
| 3 | 6 | $9-22$ | 16.00 | 4.6 | 18 | $3-25$ | 12.39 | 5.3 |
| 4 |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |
| 7 | 7 | $0-26$ | 9.71 | 8.3 | 18 | $2-21$ | 10.39 | 5.9 |
| 8 | 7 | $3-21$ | 12.57 | 5.8 | 18 | $2-26$ | 13.94 | 6.9 |
| 9 |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |
| 11 | 7 | $6-25$ | 13.71 | 6.6 | 15 | $4-20$ | 11.47 | 5.1 |
| 12 | 7 | $3-23$ | 12.71 | 6.2 | 15 | $5-26$ | 14.33 | 6.2 |
| 13 |  |  |  |  |  |  |  |  |
| 14 |  |  |  |  |  |  |  |  |
| 15 | 6 | $5-22$ | 14.17 | 5.7 | 19 | $5-23$ | 13.37 | 5.0 |
| 16 | 6 | $8-22$ | 17.50 | 5.0 | 19 | $2-32$ | 16.42 | 7.8 |
| 17 |  |  |  |  |  |  |  |  |

## Table 8. Descriptive Statistics for Basic Skills Measures in District B

## Algebra 1B <br> Algebra 1A

| wk | N | Range | Mean | SD | N | Range | Mean | SD |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |
| 11 |  |  |  |  | 12 | 1-23 | 11.17 | 6.6 |
| 12 |  |  |  |  | 11 | 4-19 | 10.73 | 5.8 |
| 13 |  |  |  |  | 15 | 1-26 | 13.27 | 8.0 |
| 14 |  |  |  |  |  |  |  |  |
| 15 |  |  |  |  | 13 | 1-32 | 14.38 | 9.6 |
| 16 |  |  |  |  | 15 | 1-34 | 11.87 | 10.0 |
| 17 |  |  |  |  | 14 | 2-37 | 14.21 | 10.3 |

## Data on Technical Adequacy

Alternate Form Reliability of Algebra Progress Monitoring Measures. We examined the alternate form reliability of individual probes by computing correlations between scores obtained on consecutive weeks. To account for the variations in teachers' administration schedules, we conducted the reliability analyses separately within each district. The results of these analyses are reported in Tables 9 and 10 for the Content Analysis-Multiple Choice and the Basic Skills measures, respectively. Alternate form reliability estimates for Content Analysis-Multiple Choice ranged from .63 to .83 in District A and from .51 to .85 in District B with some values not being statistically significant. Readers should note the variability in sample size across weeks in both districts that resulted from teachers' choices about which measures to administer. For the Basic Skills measure, alternate form reliability coefficients ranged from .65 to .81 in District A and from . 85 to .93 in District B.

The alternate form correlation coefficients for both measures were somewhat more variable than in previous research studies. In Technical Reports 13 and 14 (Perkmen, Foegen, \& Olson, 2006b, c), coefficients ranged from . 48 to .91 for the Content Analysis-Multiple Choice measure, with the lowest coefficients occurring during the initial administrations, and steady increases with subsequent administrations (most were in the .7 to .8 range). In the same studies, alternate form reliability coefficients for the Basic Skills measure ranged from .81 to .89 . The data for the present study may have been influenced by the variability in sample size. We also suspect that student motivation may play a critical role the consistency of performance over time. Our results in this study suggest that teachers should note variability in students' performance
from one administration to another and explore strategies to enhance students' motivation to give their best effort on the measures.

Table 9. Alternate Form Reliability Coefficients for Content Analysis-Multiple Choice Measures.

|  | District A |  | District B |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Week | $N$ | $r$ | $N$ | $r$ |  |
| $2 / 3$ |  |  | 15 | ns |  |
| $3 / 4$ | 61 | $.76^{* *}$ | 15 | $.63^{*}$ |  |
| $4 / 5$ | 87 | $.78^{* *}$ | 62 | $.71^{* *}$ |  |
| $5 / 6$ | 63 | $.78^{* *}$ | 35 | $.66^{* *}$ |  |
| $6 / 7$ | 65 | $.83^{* *}$ | 20 | ns |  |
| $7 / 8$ | 63 | $.73^{* *}$ | 20 | $.57^{* *}$ |  |
| $8 / 9$ | 84 | $.77^{* *}$ | 33 | $.72^{* *}$ |  |
| $9 / 10$ | 57 | $.63^{* *}$ | 39 | $.79^{* *}$ |  |
| $10 / 11$ | 64 | $.81^{* *}$ | 20 | $.61^{* *}$ |  |
| $11 / 12$ | 60 | $.77^{* *}$ | 18 | $.85^{* *}$ |  |
| $12 / 13$ | 88 | $.80^{* *}$ | 17 | $.82^{* *}$ |  |
| $13 / 14$ |  |  | 37 | $.68^{* *}$ |  |
| $14 / 15$ |  |  | 37 | $.66^{* *}$ |  |
| $15 / 16$ |  |  | $81^{* *}$ | 20 | $.51^{*}$ |
| $16 / 17$ |  |  | 19 | $.63^{* *}$ |  |

Table 10. Alternate Form Reliability Coefficients for Basic Skills Measures.

|  | District A |  | District B |  |
| :---: | :---: | :---: | :---: | :---: |
| Week | $N$ | $r$ | $N$ | $r$ |
| 2/3 |  |  |  |  |
| 3/4 |  |  |  |  |
| 4/5 |  |  |  |  |
| 5/6 |  |  |  |  |
| 6/7 |  |  |  |  |
| 7/8 | 25 | .81** |  |  |
| 8/9 |  |  |  |  |
| 9/10 |  |  |  |  |
| 10/11 |  |  |  |  |
| 11/12 | 22 | .65** | 8 | .90** |
| 12/13 |  |  | 11 | .85** |
| 13/14 |  |  |  |  |
| 14/15 |  |  |  |  |
| 15/16 | 25 | .77** | 13 | .93** |
| 16/17 |  |  | 14 | .93** |

Criterion Validity. We used available data to examine the concurrent and predictive validity of the measures. These data included students' scores on the districtadministered achievement measures (ITED and ITBS, described previously in the Measures section) and their semester grades in Algebra. To represent students' performance early in the school year, we used students' Week 5 scores (corresponding to the final week of September) for the Content Analysis-Multiple Choice measure and their Week 3 scores (corresponding to the second week of September) for the Basic Skills measure. We selected these weeks because data were available for most teachers' classes during the same week. To represent students' performance at the end of the semester, we used Content Analysis-Multiple Choice data from Week 14 for Teachers 2 and 3 in District A (who were alternating between the two measures) and Teacher 3 in District B; data from Week 17 for the remaining teachers, with the exception of Teacher 2 in District B, who did not administer the Content Analysis-Multiple Choice measure after Week 9. End of the semester performance on the Basic Skills measure was drawn from Week 16. We computed correlation coefficients between the selected probe scores and the criterion measures. The results are reported, by district, in Tables 11 and 12.

Table 11. Concurrent and Predictive Validity for the Content Analysis-Multiple Choice Measure.

| Criterion Measure | Concurrent |  |  |  | Predictive |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | A |  | B |  | A |  | B |  |
|  | $N$ | $r$ | $N$ | $r$ | $N$ | $r$ | $N$ | $r$ |
| ITBS Math Total | 22 | ns |  |  | 21 | ns |  |  |
| ITED Computation | 62 | .48** | 43 | ns | 64 | . 30 * | 61 | ns |
| ITED Prob/Data | 62 | . $51 * *$ | 42 | ns | 64 | .32* | 60 | . $34^{* *}$ |
| Algebra Grade | 88 | . $65 * *$ | 42 | .53** | 88 | .38** | 53 | .31* |

Note. $\mathrm{ns}=$ non-significant; $*=p<.05 ; * *=p<.01$.
Concurrent validity coefficients in District A were in the moderate range, with the exception of the 8th Grade Algebra class' correlations between the algebra measures and the ITBS math total score. Predictive validity coefficients were much lower, in the .30 to .38 range. In District B, many of the validity coefficients were not statistically significant. The measures did tend to correlate more strongly with algebra grades than with ITED scores, which represent more general proficiency in mathematics (as opposed to algebra in particular). These results are roughly comparable (though perhaps slightly lower) to those obtained in earlier studies. In general, criterion validity coefficients for the algebra progress monitoring measures are in the moderate range.

Table 12. Concurrent and Predictive Validity for the Basic Skills Measure.

| Criterion Measure | Concurrent |  |  |  | Predictive |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | A |  | B |  | A |  | B |  |
|  | $N$ | $r$ | $N$ | $r$ | $N$ | $r$ | $N$ | $r$ |
| ITBS Math Total |  | -- |  | -- |  | -- |  | -- |
|  |  |  |  |  |  |  |  |  |
| ITED Computation | 23 | ns | 15 | .81** | 22 | .47* |  | -- |
|  |  |  |  |  |  |  |  |  |
| ITED Prob/Data | 23 | ns | 15 | ns | 22 | ns |  | -- |
|  |  |  |  |  |  |  |  |  |
| Algebra Grade | 25 | ns | 13 | .78** | 24 | ns |  | -- |

Note. $\mathrm{ns}=$ non-significant; $*=p<.05 ; * *=p<.01$.
The results reported for the Basic Skills measure in Table 12 reveal that the correlation coefficients differed substantially by district. While the scores from students in District B showed strong relations to the ITED Computation subtest and students’ algebra grades, there were no significant relations for the concurrent coefficients in District A. The moderate predictive relation between Basic Skills scores in the fall and students' subsequent ITED Computation score in District A may suggest that the measure taps into a component of computational skill that may be predictive of later performance on a subtest of general mathematics computation.

Growth. We examined students' growth on the measures across the semester by computing an ordinary least squares regression coefficient for each student's weekly data. Readers should note that in District A, this analysis treats the data for two probes gathered in a single week as if they had been gathered in two consecutive weeks. Table 13 reports the ranges, means and standard deviations for the mean slopes and the corresponding standard errors for the Content Analysis-Multiple Choice and Basic Skills measures by district and class type.

It is important to note that throughout Project AAIMS, we have used a weekly growth rate of .5 as benchmark and goal in our research. We anticipate that in order for algebra progress monitoring measures to be useful to teachers on a practical level, they must be able to expect to see scores grow by at least one point every two weeks (hence a weekly growth rate of .5). The results reported in Table 13 reveal that this threshold was met for the Content Analysis-Multiple Choice measure in three classes (8th Grade Algebra and Algebra 1 in District A, and Algebra 1B in District B). The Algebra 1A classes in both districts had similar mean slopes, both around .35. The Algebra 1B class in District A had a negative mean slope.

Mean slope values on the Basic Skills measure ranged from .29 to .48. In District A, the values were .29 for both the Algebra 1A and the 1B classes. Slopes on this measure were higher in District B, where students in Algebra 1A gained nearly . 5 points per week on this measure. These rates are similar to previous research, where average growth across class types has ranged from .34 to .46 points per week.

Table 13. Descriptive Statistics for Weekly Slope and Standard Error on the Algebra Progress Monitoring Measures

|  | 8th Grade Algebra |  |  | Algebra 1 |  |  |  | Algebra 1B |  |  |  | Algebra 1A |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| N | Range | Mean | SD | N | Range | Mean | SD | N | Range | Mean | SD | N | Range | Mean | SD |
| Content Analysis-Multiple Choice |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| District A |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Slope 22 | 0.9-2.1 | 1.5 | 0.3 | 45 | -0.1-2.1 | 0.80 | 0.5 | 7 | -1.4-0.7 | -0.45 | 0.6 | 19 | -0.6-1.6 | 0.37 | 0.6 |
| SEE 22 | . $16-.32$ | . 24 | 0.1 | 45 | . $11-.56$ | 0.27 | 0.1 | 7 | . $51-1.2$ | 0.79 | 0.3 | 19 | . $13-.92$ | 0.51 |  |
| District B |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Slope |  |  |  |  |  |  |  | 26 | -0.7-2.0 | 0.79 | 0.6 | 48 | -1.5-2.3 | 0.33 | 0.8 |
| SEE |  |  |  |  |  |  |  | 26 | . $23-1.6$ | 0.51 | 0.4 | 48 | . $11-1.5$ | 0.52 | 0.3 |
| Basic Skills |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| District A |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Slope |  |  |  |  |  |  |  | 7 | -0.1-1.1 | . 29 | 0.4 | 19 | -0.5-. 92 | . 29 | 0.4 |
| SEE |  |  |  |  |  |  |  | 7 | 0.2-0.4 | . 30 | 0.1 | 19 | 0.1-6 | . 32 | 0.2 |
| District B |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Slope |  |  |  |  |  |  |  |  |  |  |  | 15 | -0.9-2.7 | . 48 | 1.0 |
| SEE |  |  |  |  |  |  |  |  |  |  |  | 15 | 0.1-0.8 | . 44 | 0.2 |

## Effects of Teachers’ Access to Student Data on Student Performance

In addition to replicating previous research on the technical adequacy of the algebra progress monitoring measures, a primary purpose of this study was to examine the effects of teachers' access to the data on students' levels of performance. As described in the Method section, we assigned two of each teacher's classes (of the same type) to a condition in which they had access to their students' data (data classes) or student data was withheld until the end of the semester (no data classes); these assignments were reported in Table 4.

We first examined the comparability of the classes at the beginning of the semester by conducting t -tests of the beginning of the semester probes scores (the same values used in the criterion validity analyses). The results revealed no significant differences in initial performance levels for all six pairs of classes.

We next conducted t-tests for each pair of classes, using students' slopes and their end of semester performance levels as the outcome variables. The results of these analyses are reported in Table 14.

Table 14. T-test Results for Comparison Classes

|  | District | Teacher (Classes) | $\frac{\text { Mean for Data }}{\text { Class }}$ | $\frac{\text { Mean for No }}{\text { Data Class }}$ | $\underline{t}$ | $p$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| CA-MC |  |  |  |  |  |  |
| Slope | A | $1(3,2)$ | . $66(\mathrm{~N}=13$ ) | . $70(\mathrm{~N}=8)$ | -0.23 | . 82 |
|  |  | $1(7,6)$ | . 73 ( $\mathrm{N}=13$ ) | $1.12(\mathrm{~N}=11)$ | -1.81 | . 09 |
|  |  | 2 \& $3(5,7)$ | . $32(\mathrm{~N}=11$ ) | . $44(\mathrm{~N}=8)$ | -0.40 | . 70 |
|  | B | $1(1,2)$ | . $76(\mathrm{~N}=12$ ) | . $81(\mathrm{~N}=14$ ) | -0.17 | . 86 |
|  |  | $2(4,2)$ | -. $11(\mathrm{~N}=6)$ | . 54 ( $\mathrm{N}-10)$ | -1.13 | . 29 |
|  |  | $3(2,3)$ | . $34(\mathrm{~N}=18)$ | . $35(\mathrm{~N}=14$ ) | -0.004 | . 99 |
| EOS | A | $1(3,2)$ | $19.60(\mathrm{~N}=13)$ | $27.00(\mathrm{~N}=8)$ | -1.79 | . 09 |
|  |  | $1(7,6)$ | 21.77 ( $\mathrm{N}=13$ ) | 28.73 ( $\mathrm{N}=11$ ) | -1.71 | . 10 |
|  |  | 2 \& $3(5,7)$ | $8.09(\mathrm{~N}=11)$ | 14.00 ( $\mathrm{N}=8)$ | -2.43* | . 03 |
|  | B | $1(1,2)$ | 30.11 ( $\mathrm{N}=9$ ) | $28.54(\mathrm{~N}=11)$ | 0.51 | . 62 |
|  |  | $2(4,2)$ | -- | -- | -- | -- |
|  |  | $3(2,3)$ | 13.3 ( $\mathrm{N}=13$ ) | $15.2(\mathrm{~N}=10)$ | -0.61 | . 55 |
| BS |  |  |  |  |  |  |
| Slope | A | 2 \& $3(5,7)$ | . $36(\mathrm{~N}=11$ ) | . $19(\mathrm{~N}=8)$ | 0.84 | . 42 |
|  | B | $2(4,2)$ | . 43 ( $\mathrm{N}=5$ ) | . $50(\mathrm{~N}=10)$ | -0.11 | . 92 |
| EOS | A | 2 \& $3(5,7)$ | $15.36(\mathrm{~N}=11)$ | $17.88(\mathrm{~N}=8)$ | -0.66 | . 52 |
|  | B | $2(4,2)$ | 8.80 ( $\mathrm{N}=5$ ) | 13.40 ( $\mathrm{N}=10$ ) | -0.78 | . 46 |

Note. $\mathrm{EOS}=$ End of semester level of performance
The data in Table 14 reveal that no significant differences in the expected direction were obtained in any of the comparisons. The single statistically significant result was found for the End of Semester scores on the Content Analysis-Multiple Choice
measure, but in the opposite direction (with the No Data class having a higher final mean than the Data class). These results are likely influenced by several factors. First, the treatment condition ("Data") was relatively weak because the data conferences occurred only monthly and the data that were reviewed had often been collected three to four weeks earlier and therefore did not represent current student performance levels. The fact that graphs were not shared with the students until the end of the semester was also a concern. Teachers commented that students seemed to lack motivation to put forth their best efforts on the measures. An additional factor that may have influenced the results was the fact that the Data and No Data conditions were different class periods taught by the teachers. When teachers learned of specific concerns about student performance (particularly issues that arose when reviewing the performance summaries), they implemented changes across all of their class periods. As a result, students in both conditions would have benefited from the supplemental instruction designed as a result of information from the Data condition. While this situation is not ideal, we found it unethical to ask teachers to withhold what they saw as necessary instruction from any of their students for research purposes.

## Summary and Future Research

The results of this study provide additional data documenting the alternate form reliability and criterion validity of the Basic Skills and Content Analysis-Multiple Choice measures for algebra progress monitoring. As an initial effort to explore the effects of teachers' use of algebra progress monitoring data to effect increased levels of student achievement, we found that providing student data and summaries of student performance on a monthly basis was not sufficient to result in observable differences between two class sections taught by the same teacher. Future research should examine methods to further engage teachers with students' data and to share data with students on a more regular basis.

## References

Foegen, A., Olson, J., \& Perkmen, S. (2005). Reliability and criterion validity of four revised algebra measures in Districts B and C. (Technical Report 10). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaims/Technical.html.

Perkmen, S., Foegen, A., \& Olson, J. (2006a). Reliability, criterion validity, and sensitivity to growth: Extending work on two algebra progress monitoring measures. (Technical Report 12). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaims/Technical.html.

Perkmen, S., Foegen, A., \& Olson, J. (2006b). A replication study of the reliability, criterion validity and sensitivity to growth of two algebra progress monitoring measures. (Technical Report 13). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaims/Technical.html.

Perkmen, S., Foegen, A., \& Olson, J. (2006c). Technical characteristics of two algebra progress monitoring measures: Reliability, criterion validity, and sensitivity to growth. (Technical Report 14). Project AAIMS, Department of Curriculum and Instruction, Iowa State University, Ames, IA. Available: http://www.ci.hs.iastate.edu/aaims/Technical.html.

## Appendix

# Algebra Progress Monitoring Measures 

Basic Skills<br>Content Analysis-Multiple Choice

## Sample Summary of Student Performance

Error Analysis Report for Content Analysis-Multiple Choice

| Algebra Probe A-1 |  | age 24 |
| :---: | :---: | :---: |
| Solve: $\begin{aligned} & 9+a=15 \\ & a= \end{aligned}$ | Solve: $\begin{aligned} & 10-6=g \\ & g= \end{aligned}$ |  |
| Evaluate: $12+(-8)+3$ | Simplify: $9-4 d+2+7 d$ |  |
| Simplify: $2 x+4+3 x+5$ | Simplify: $5(b-3)-b$ |  |
| Solve: $\begin{aligned} & 12-e=4 \\ & e= \end{aligned}$ | Solve: $\begin{aligned} & q \cdot 5=30 \\ & q= \end{aligned}$ |  |
| Simplify: $4(3+s)-7$ | Evaluate: $8-(-6)-4$ |  |
| Simplify: $b+b+2 b$ | Simplify: $2+w(w-5)$ |  |
| Solve: $\frac{r}{6}=\frac{12}{18}$ | Solve: <br> 1 foot $=12$ inches <br> 5 feet $=$ $\qquad$ inches |  |
| Simplify: $7-3(f-2)$ | Simplify: $4-7 b+5(b-1)$ |  |
| Evaluate: $-5+(-4)-1$ | Simplify: $s+2 s-4 s$ |  |
| Solve: $\begin{aligned} & 63 \div c=9 \\ & c= \end{aligned}$ | Solve: $\begin{aligned} & x+4=7 \\ & x= \end{aligned}$ |  |
| Simplify: $2(s-1)+4+5 s$ | Simplify: $-5(q+3)+9$ |  |
| Simplify: $8 m-9(m+2)$ | Evaluate: $9+(-3)-8$ |  |
| Solve: <br> 3 feet $=1$ yard $\qquad$ feet $=9$ yards | Solve: $\frac{12}{2}=\frac{48}{e}$ |  |
| Evaluate: $4-(-2)+8$ | Simplify: $y^{2}+y-4 y+3 y^{2}$ |  |
| Simplify: $2 k+3-5(k+7)$ | Simplify: $3(c+2)-2 c$ |  |




| Solve: $\begin{aligned} & 3 x+4=19 \\ & x= \end{aligned}$ <br> a) 8 <br> b) 22 <br> c) 15 <br> d) 5 | Evaluate $a^{2}-b \div 2$ when $a=4$ and $b=6$ <br> a) 1 <br> b) 5 <br> c) 10 <br> d) 13 | Which line on the graph is $y+2 x=4$ ? <br> a) Line A <br> b) Line B <br> c) Line C <br> d) Line D | Simplify: $3(m+2)+2(m-1)$ <br> a) $5 m+4$ <br> b) $5 m+1$ <br> c) $6 m+8$ <br> d) $6 m-8$ |
| :---: | :---: | :---: | :---: |
| Evaluate the expression: $4^{-2}$ <br> a) -16 <br> c) $\frac{1}{16}$ <br> b) $\frac{1}{8}$ <br> d) -8 | Solve the linear system: $\begin{aligned} & x-y=4 \\ & x+2 y=19 \end{aligned}$ <br> a) $(-1,-5)$ <br> b) $(5,8)$ <br> c) $(-2,19)$ <br> d) $(9,5)$ | This graph shows the solution for which equation? <br> a) $x>-3$ <br> b) $2 x \leq-6$ <br> c) $-3 x>9$ <br> d) $3 x \geq 9$ | Write the equation in slopeintercept form if $m=\frac{1}{2}$ and $b=3$ <br> a) $y=2 x+3$ <br> b) $y=3 x+\frac{1}{2}$ <br> c) $x=\frac{1}{2} y-3$ <br> d) $y=\frac{1}{2} x+3$ |


| Evaluate $d+3 e^{2}$ when $d=5$ and $e=2$ <br> a) 11 <br> b) 23 <br> c) 17 <br> d) 10 | Solve: $\begin{aligned} & 6 c+4=-3 c-14 \\ & c= \end{aligned}$ <br> a) $-\frac{10}{3}$ <br> b) -2 <br> c) 2 <br> d) 6 | Find the slope of a line through $(1,-1)(5,2)$ <br> a) $\frac{1}{5}$ <br> b) $\frac{3}{4}$ <br> c) -6 <br> d) $-\frac{4}{3}$ | Simplify: $6(2 b-3)-3(2-b)$ <br> a) $15 b-24$ <br> b) $9 b-9$ <br> c) $9 b+12$ <br> d) $15 b+12$ |
| :---: | :---: | :---: | :---: |
| Simplify the expression: $\frac{a^{2}}{a b^{3}} \cdot \frac{b^{4}}{a^{3}}$ <br> a) $\frac{a^{8}}{a^{3} b^{3}}$ <br> b) $\frac{a b^{8}}{a^{4} b^{3}}$ <br> c) $\frac{b}{a^{2}}$ <br> d) $\frac{b}{a}$ | Solve the linear system: $\begin{aligned} -6 x+3 y & =-6 \\ 2 x+6 y & =30 \end{aligned}$ <br> a) $(6,3)$ <br> b) $(3,4)$ <br> c) $(2,6)$ <br> d) $(4,-3)$ | Simplify: $b^{2}-4 b+2 b^{2}+7-5$ <br> a) $3 b^{2}-4 b+2$ <br> b) $2 b+2$ <br> c) $-b^{2}-4 b+12$ <br> d) $3 b^{2}-4 b+12$ | Write the equation of a line through $(5,3)(4,9)$. Use pointslope form. <br> a) $y+1=2(x-4)$ <br> b) $y+4=-6(x-1)$ <br> c) $y-3=-6(x-5)$ <br> d) $y=-6 x+30$ |



| Chapter | "Big Idea" | Samples | $\begin{aligned} & \% \\ & (3) \\ & \hline \end{aligned}$ | $\begin{gathered} \% \\ (2) \\ \hline \end{gathered}$ | $\begin{gathered} \% \\ (1) \\ \hline \end{gathered}$ | $\begin{gathered} \% \\ \mathbf{( 0 )} \\ \hline \end{gathered}$ | $\begin{gathered} \% \\ (\mathbf{G}) \\ \hline \end{gathered}$ | $\begin{gathered} \% \\ (\mathbf{B}) \\ \hline \end{gathered}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | E5.2 Write equation for line through 2 points | E5.2a Write the equation of a line through $(5,3)(4,9)$. <br> Use point-slope form. <br> E5.2b Write the equation of a line through $(4,2)(6,3)$. <br> Use slope-intercept form. <br> (1 item) | 24 | 0 | 5 | 0 | 10 | 62 |
|  |  |  | 24 | 0 | 0 | 10 | 14 | 52 |
|  |  |  | 9 | 4 | 39 | 13 | 0 | 35 |
|  |  |  | 17 | 0 | 13 | 13 | 9 | 48 |
| 6 <br>  <br> Graphing <br> Linear <br> Inequalities | E6 Interpret a graph of an inequality | This graph shows the solution for which equation? <br> (1 item) | 33 | 5 | 5 | 5 | 10 | 43 |
|  |  |  | 67 | 0 | 5 | 0 | 5 | 24 |
|  |  |  | 30 | 0 | 0 | 13 | 17 | 39 |
|  |  |  | 26 | 0 | 0 | 17 | 13 | 43 |
| 7 <br> Systems of <br> Linear <br>  <br> Inequalities | E7.1 Solve linear system by substitution | Solve the linear system: $\begin{aligned} & x-y=4 \\ & x+2 y=19 \end{aligned}$ <br> (1 item) | 24 | 0 | 29 | 0 | 5 | 43 |
|  |  |  | 14 | 0 | 0 | 19 | 10 | 57 |
|  |  |  | 48 | 13 | 9 | 0 | 13 | 17 |
|  |  |  | 39 | 13 | 0 | 0 | 9 | 39 |
|  | E7.2 Solve linear system by linear combination | Solve the linear system:$\begin{aligned} & -6 x+3 y=-6 \\ & 2 x+6 y=30 \\ & (1 \text { item }) \end{aligned}$ | 71 | 0 | 0 | 0 | 10 | 19 |
|  |  |  | 38 | 19 | 0 | 5 | 5 | 33 |
|  |  |  | 26 | 0 | 4 | 9 | 13 | 48 |
|  |  |  | 35 | 0 | 4 | 4 | 9 | 48 |
| 8 <br>  <br> Exponential Functions | E8.1 Evaluate expressions with negative exponents | Evaluate the expression: $4^{-2}$ <br> (1 item) | 62 | 0 | 24 | 0 | 5 | 10 |
|  |  |  | 86 | 0 | 10 | 0 | 5 | 0 |
|  |  |  | 61 | 0 | 0 | 13 | 13 | 13 |
|  |  |  | 61 | 0 | 0 | 13 | 17 | 9 |
|  | E8.2 Simplify expressions with exponents | Simplify the expression:$(1 \mathrm{item}) \frac{a^{2}}{a b^{3}} \cdot \frac{b^{4}}{a^{3}}$ | 38 | 5 | 14 | 0 | 10 | 33 |
|  |  |  | 33 | 0 | 19 | 0 | 10 | 38 |
|  |  |  | 52 | 0 | 13 | 0 | 0 | 35 |
|  |  |  | 43 | 13 | 9 | 0 | 4 | 30 |

