A Comparison of CBMmath Process Scoring Methods: Rapid Scoring and Error Analysis

Danielle M. Becker
Rachel Brown, Ph.D., NCSP
Danielle Dupuis, Ph.D.

520 Nicollet Mall
Suite #910
Minneapolis, MN 55402

612.254.2534
sales@fastbridge.org
www.fastbridge.org

FastBridge Learning
THIS WAY TO FASTER RESULTS.

Abstract

Math CBM has been available for over 30 years and provides an important tool for both screening and progress monitoring students' math skills. Although available, it is not used as widely as CBM for reading, perhaps because traditional formats (e.g., paper and pencil) are more time consuming to score. This study compared two types of math CBM scoring: Rapid and Error Analysis. Results indicated that these two methods yield scores that were highly correlated. This result is promising because the Rapid method is potentially faster and easier for teachers to use. Still, additional research is needed to replicate these results and to add information about actual scoring time, the benefits for instructional planning, and whether these results are true for progress data as well.

Recommended citation:

Becker, D., Brown, R., Dupuis, D. (2018). *A comparison of CBMmath process scoring methods: Rapid scoring and error analysis* [white paper]. Minneapolis, MN: FastBridge Learning.

A Comparison of CBMmath Process Scoring Methods:
Rapid Scoring and Error Analysis

Various forms of Curriculum-Based Measurement (CBM) have provided teachers with brief and easy to use formative assessments of students' academic skills. Although many of these forms are scored during administration, an exception is one type of CBMmath assessment that measures the steps used to solve multi-digit computation problems. Such problems require understanding and application of many specific steps such as regrouping based on place value and using additional operations. *FastBridge Learning* distributes a version of such a measure called CBMmath Process. This assessment is designed for students in grades 2 through 6 and can be used for both screening and progress monitoring, although different scoring methods are applied for each purpose. The types of problems included for each grade level vary in difficulty. There are both general outcome measure (GOM) and single skill versions of this assessment. The GOM version includes multiple problem types that reflect the standards at each grade level. The single-skill probes include only one type of problem and are designed to measure a student's mastery of the specific skill. All CBMmath Process probes are timed for 10 minutes in grades 2 through 4 and 15 minutes in grades 5 and 6.

An important benefit of using CBM is that these measures provide information about how students are likely to score on future assessments including annual state tests. Shapiro, Keller, Lutz, Santoro, and Hintze (2006) showed that a form of CBM for mathematics was predictive of later student performance on the Pennsylvania state assessment. Similarly, Sisco-Taylor, Fung, and Swanson (2014) found that math CBM assessments that included word problems predicted student performance on the California state test. The predictive validity of math CBM is encouraging but there remain questions about how well various forms identify individual student learning needs (Hintze, Christ, & Keller, 2002). Strait, Smith, Pender, Malone, Roberts, and Hall (2015) found that a free version which allows teachers to select the specific problem types was adequately reliable for screening purposes. These researchers, however, did not consider different types of scoring procedures for CBM for math.

Identifying the specific errors that students make when completing math problems has long been an important part of mathematics instruction. Dennis, Calhoon, Olson, and Williams, (2013) found that rates of improvement for students with math difficulties were better when error analysis procedures were used. Also important is the accuracy of the measures used. Research on math CBM forms has indicated that they can be reliable and valid indicators of students' math skills, however, variability across forms exists. In a study of the order and type of items on math CBM forms, Methe, Briesch, and Hulac (2015) found that a controlled item order (e.g., stratified) was less prone to measurement error than random problem order. These findings suggest that it is important for teachers to be aware of what skills are tested on math assessments and that the order and type of skills represented could affect students' scores.

Despite recent progress in documenting the benefits from math CBM, there remain limitations in these assessments. An ongoing frustration for teachers using assessments like CBMmath Process is that this type of assessment can be very time-consuming to score. Traditionally, such measures involve analyzing each problem in relation to the steps – or processes – that the student used to solve it. This procedure includes identifying and coding any and all errors that the student made and assigning point values in relation to the total correct steps. Although this method can provide details about each student's procedural fluency in math, it can take a very long time to complete this analysis for each student. As a result, many teachers have found using such assessments too time intensive.

In an effort to make CBMmath Process easier and faster to score, an alternative method was developed. Known as the Rapid Scoring procedure, this alternative offers a simpler way to score student answers. Instead of coding each error and deducting points in relation to the types of errors, Rapid

Scoring involves scoring student answers according to the number of errors. Half of the points are awarded for problems with no more than two minor errors and zero points for problems with three or more errors. This study evaluated the consistency in score outcomes between the new Rapid Scoring method and the traditional Error Analysis method to determine if the Rapid Scoring method is a reliable alternative to full error analysis.

Method

Participants

A convenience sample of 221 fifth grade and 187 sixth grade students completed FAST CBMmath Process probes. The students attended a suburban elementary school in the upper Midwestern U.S. The sample included all probes completed by the students during spring universal screening.

Materials

FAST CBMmath process is a timed assessment of multistep computation skills. Both single skill and general outcome measures (GOM) are available. The GOM version is used for screening and was used in this analysis. Fifth grade GOM probes include multi-digit whole number multiplication and division problems. Sixth grade GOM probes include multi-digit addition, subtraction, multiplication and division problems with whole numbers and decimals. The probes are completed using paper and pencil.

Procedures

Student forms were distributed by the classroom teachers who instructed students to leave the papers turned over while directions were read. After reading the directions, the teacher instructed the students to turn over the paper and said “begin” and started a timer. Both the fifth and sixth grade probes were timed for 15 minutes. At the end of 15 minutes, the teacher instructed the students to stop and then collected the papers. All of the probes were scored by trained university graduate (?) students using two methods: (a) Rapid Scoring, and (b) Error Analysis. For both methods, all of the problems that a student attempted or skipped up through the last one attempted were scored. Each problem had a range of points possible based on the number of steps required to complete the problem correctly.

Rapid scoring. This method involves awarding all, half, or zero possible points for each completed problem. All points were awarded for problems there were 100% correct. Half of the points were awarded for problems with no more than two minor errors and zero points for problems with three or more errors.

Error analysis. This method involves examining each step the student completed as part of problem solution and coding all errors using the following error types:

- **Calculation-addition error (C+):** A student made an error in adding two numbers; a student wrote an incorrect answer for an addition problem. This includes addition errors made in all items involving addition, including multi-step multiplication items (1 point).
- **Calculation-subtraction error (C-):** A student made an error in subtracting two numbers; a student wrote an incorrect answer for a subtraction problem. This includes subtraction errors made in all items involving subtraction, including multi-step division items (1 point).
- **Calculation-multiplication error (Cx):** A student made an error in multiplying two numbers; a student wrote an incorrect answer for a multiplication problem (1 point).
- **Calculation-division error (C÷):** A student made an error in dividing two numbers; a student wrote an incorrect answer for a division problem (1 point).

- **Regrouping or Carrying error: (R/C)** A student made an error in adding numbers with a sum greater than 10 or subtracting numbers with a difference less than 0; the student made an error in "borrowing" from the next placeholder to solve the subtraction problem or "carrying" a number to the next placeholder to solve the addition problem (1 point).
- **Remainder error: (RE)** A student made an error in a division problem in regard to the remainder; the remainder in the solution is incorrect (1 point).
- **Primary Process error (PP):** The student missed one or more primary processes (steps) used to solve the problem (2 points).
- **Place Holder/Maintain Value error: (PH/V)** A student made an error in maintaining place value or using a place holder in calculating a multiplication or division item. This also includes failure to maintain correct decimal places or forgetting to include a decimal (1 point).
- **Primary Operation error (PO):** The student used the wrong operation at the first step in the problem, thus resulting in an incorrect solution (1 point).
- **Unknown error (UE):** Any error that does not match the other error definitions (1 point).
- **Skipped (S):** All skipped items are awarded zero points (all points).
- **Problem Setup Incorrectly (PSI):** All CBMmath Process problems are shown horizontally on the page. Students must rewrite the problem in the format matching the problem type (i.e., use a division "box" for division problems;1 point).
- **Wrong Operation error: (WO)** A student made an error due to misreading a primary operation sign and performing a different operation than specified in the item. This type of error does not include secondary operation signs (e.g., addition within a multi-step multiplication problem or subtraction within a multi-step division problem; all points).

With the error analysis method, any and all errors were coded and the corresponding number of points deducted from the total for that problem. Most error types result in a deduction of one point; the exceptions are Primary Process errors which deduct 2 points, Skipped problems which deduct all points, and Wrong Operations which deduct all points.

Interobserver agreement (IOA). In order to confirm the accuracy of scoring, 30% of probes were scored both ways by two researchers. Scoring accuracy was calculated using Cohen's Kappa where $K = p_o – p_e / 1 -- p_e$. For the Error Analysis method, Kappa for grade 5 was .84 and for grade 6 was .83. The Kappa values for Rapid Scoring were .93 at grade 5 and .89 at grade 6. All of these values indicate adequate agreement between raters.

Results

The two score types for each problem were compared using Pearson (product-moment) correlation coefficients. Tables 1 and 2 display each problem, the correlation coefficient for scores on each problem, and average correlation for each grade. Most correlations were strong and indicated probability (*p*) of statistical significance for grades 5 and 6. The average grade 5 coefficient was .81; grade 6 was .79. The fifth-grade correlations ranged from .496 to .918. The range for sixth grade was .315 to a perfect correlation of 1.0. The fifth-grade problem with the lower correlation was 9846 ÷ 76. The sixth-grade problem with the much lower correlation between scores was 5.74 ÷ 2.23. At both grade levels, the lowest correlation was an outlier and the vast majority of correlations were all .70 and higher. The problem with the perfect correlation was 81 ÷ 6. The correct answer is 13.5. There are very few digits in this calculation and when the Rapid method was used, the resulting student answers were all either right or wrong. This made agreement on the score for the item easy to assign, resulting in the perfect correlation.

Table 1

Grade 5 correlations between error analysis and rapid scoring methods

#	Problem	Correlation*
1	588 ÷ 49	.815853019
2	124 x 44	.819505645
3	6123 ÷ 26	.660199019
4	6352 ÷ 16	.74174751
5	985 x 567	.609116597
6	60 ÷ 7	.753058688
7	595 ÷ 76	.821532852
8	570 ÷ 15	.86330372
9	909 x 28	.874591978
10	6549 ÷ 11	.77926538
11	610 ÷ 10	.900409887
12	66 x 87	.904164111
13	345 x 66	.911159274
14	775 ÷ 77	.896054229
15	9265 ÷ 8	.72964361
16	989 x 43	.682301409
17	987 ÷ 32	.904278043
18	877 x 8	.885150303
19	5856 ÷ 96	.815853019
20	930 x 55	.918553338
21	9846 ÷ 76	.496579207
22	260 ÷ 10	.912560573
23	906 x 480	.768951094
24	595 ÷ 3	.918155297
	AVERAGE	.807582825

*$p < .000$

Table 2

Grade 6 correlations between error analysis and rapid scoring methods

#	Problem	Correlation*
1	55.7 + 34.1	.752192258
2	277 x 61	.757729223
3	75.133 - 49.378	.774556814
4	1035 ÷ 45	.808469765
5	1.9 x 4.4	.722600974
6	3984 ÷ 35	.730212655
7	3.296 ÷ 1.3	.681171101
8	3301 x 3	.941284705
9	755.72 + 769.22	.90868628
10	723 x 553	.795348951
11	134.54 - 102.58	.899153595
12	3200 ÷ 50	.87855975
13	8.22 x 9.40	.781038585
14	1657 ÷ 15	.743126195
15	5.74 ÷ 2.23	.315368264
16	86 x 37	.904411881
17	97.562 + 74.582	.972478255
18	607 x 820	.852971108
19	23.4 - 12.6	.956613153
20	5691 ÷ 21	.85981605
21	7.347 x 7.400	.665608634
22	2567 ÷ 20	.820329422
23	4.6 ÷ 1.3	.438295328
24	81 ÷ 6	1.0
	AVERAGE	.790000956

*$p < .000$

Discussion

The correlations reported here indicate that the Rapid Scoring and Error Analysis methods yielded highly similar score results. This finding is important because it suggests that Rapid Scoring is equally reliable to Error Analysis of CBMmath Process assessments. Teachers have long complained about the time consuming and difficult procedures needed to score the Process assessment and having the Rapid method available could reduce both scoring time and teacher frustration. Although these results are promising, there are some limitations to Rapid Scoring use. Specifically, Rapid Scoring is based on counting the number of errors that a student makes on each problem, but it does not provide details about the types of errors made or what additional instruction a student needs in order to develop improved math skills. In contrast, Error Analysis provides details of a student's error types and can be used to plan additional instruction based on those errors. As noted by Anderson, Lai, Alonzo, and Tindal (2011) and Dennis et al. (2013), using CBM data to drive individualized instruction was a design component and there might be cases when Rapid Scoring will provide insufficient information about a student's instructional needs.

The strong correlations found in this study indicate that the steps involved in scoring math CBM problems with a Rapid method are similar enough to Error Analysis for either method to be used in order to compare student performance in relation to either benchmarks or norms. Nonetheless, the Rapid Scoring method will show teachers only which students have scores below, at, or above grade level expectations. There is no direct pathway to convert the number of errors from the Rapid Scoring method to a full Error Analysis. In other words, the Rapid method will show teachers which students have difficulty completing multi-step problems, but not what specific problems they have. In cases where detailed information about the types of errors that a student made is necessary for instructional planning, the Error Analysis method might be superior.

Limitations and Additional Research

This study has limitations that require additional research. The student sample was limited to students attending one school in the Midwest, and not representative of all students. In addition, only selected grade levels were included, so the results might not reflect how well the two scoring methods match at other grade levels. The types of problems that the students completed might not be representative of the mathematics expectations and standards in other regions. This analysis examined data from one time point and did not compare student performance over time. The results should be considered with these limitations in mind, however, the findings suggest that additional research is warranted.

Future research about different scoring methods could include data concerning how long each scoring method takes per student. This would address teachers' concerns that any form of math CBM requiring paper and pencil completion is prohibitively time-consuming. Additional studies that examine how well the score information obtained from each method contributes to intervention planning would helpful as well. These initial findings suggest that Rapid Scoring results in scores on a par with Error Analysis, but it is unknown whether the more detailed error data are essential for intervention planning. Finally, analysis of students' progress monitoring scores over time using both scoring methods could reveal whether both methods are viable for tracking student math improvement, or if only one method is sensitive to such change. Currently, the FastBridge system allows the Rapid Method for screening only, and the Error Analysis for progress monitoring only. Research that evaluates how well each scoring method works when applied to each student's data over time would show whether changes in these scoring limitations are justified.

Conclusion

Math CBM has been available for over 30 years and provides an important tool for both screening and progress monitoring students' math skills. Although available, it is not used as widely as CBM for reading, perhaps because traditional formats (e.g., paper and pencil) are more time consuming to score. This study compared two types of math CBM scoring: Rapid and Error Analysis. Results indicated that these two methods yield scores that were highly correlated. This result is promising because the Rapid method is potentially faster and easier for teachers to use. Still, additional research is needed to replicate these results and to add information about actual scoring time, the benefits for instructional planning, and whether these results are true for progress data as well.

References

Anderson, D., Lai, C., Alonzo, J., & Tindal, G. (2011). Examining a grade-level math CBM designed for persistently low-performing students. *Educational Assessment, 16*, 15-34. 10.1080/10627197.2011.551084

Dennis, M. S., Calhoon, M. B., Olson, C. L., & Williams, C. (2014;2013;). Using computation curriculum-based measurement probes for error pattern analysis. *Intervention in School and Clinic, 49*, 281-289. 10.1177/1053451213513957

Hintze, J. M., Christ, T. J., & Keller, L. A. (2002). The generalizability of CBM survey-level mathematics assessments: Just how many samples do we need? *School Psychology Review, 31*, 514-528.

Methe, S. A., Briesch, A. M., & Hulac, D. (2015). Evaluating procedures for reducing measurement error in math curriculum-based measurement probes. *Assessment for Effective Intervention, 40*, 99-113. 10.1177/1534508414553295

Shapiro, E. S., Dennis, M. S., & Fu, Q. (2015). Comparing computer adaptive and curriculum-based measures of math in progress monitoring. *School Psychology Quarterly, 30*, 470-487. 10.1037/spq0000116

Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24*, 19-35. 10.1177/0734282905285237

Sisco-Taylor, D., Fung, W., & Swanson, H. L. (2015;2014;). Do curriculum-based measures predict performance on word-problem-solving measures? *Assessment for Effective Intervention, 40*, 131-142. 10.1177/1534508414556504

Strait, G. G., Smith, B. H., Pender, C., Malone, P. S., Roberts, J., & Hall, J. D. (2015). The reliability of randomly generated math curriculum-based measurements. *Assessment for Effective Intervention, 40*, 247-253. 10.1177/1534508415588075