Technical Report # 0906

Alternate Form and Test-Retest Reliability of easyCBM Reading Measures

Julie Alonzo

Gerald Tindal

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

This study was partially funded by a grant from the Office of Special Education Programs (OSEP): Model Demonstration Centers on Progress Monitoring (Award # H326M050003). The findings and conclusions, however, reflect those of the authors only and do not necessarily represent the perspectives of the U.S. Department of Education or OSEP.

Abstract

We report the results of a test-retest and alternate form reliability study of grade 1, 3, 5, and 8 reading measures from the easyCBM assessment system. Approximately 50 students in each grade participated in the study. In Grade 1, we studied the following measures: Phoneme Segmenting, Letter Sounds, Letter Names, Word Reading Fluency, and Passage Reading Fluency. In Grade 3, we studied Word and Passage Reading Fluency and Multiple Choice Reading Comprehension. In both Grades 5 and 8, we studied Passage Reading Fluency and Multiple Choice Reading Comprehension. Correlations for the test-retest analyses ranged from .45 to .97. Correlations for the alternate form analyses ranged from .76 to .97.

Introduction

The easyCBM™ assessment system (Alonzo, Tindal, Ulmer, & Glasgow, 2006) was developed by researchers at the University of Oregon. Initial funding for the assessments came in part through a 2006 federal grant for progress monitoring, with an emphasis on pre-K through grade 4. Over the years, the system has been expanded to include reading measures in grades K-8, with mathematics measures to be added in the future. Detailed descriptions of the measurement development can be found in the technical reports available on the easyCBM™ website. In this technical report, we describe a study of the alternate form and test-retest reliability of the easyCBM™ reading measures.

Methods

In this section, we describe the methods we used in conducting a study of the alternate form and test-retest reliability of a selection of measures from the easyCBM assessment system.

Setting and Participants

We conducted the reliability study in a mid-sized K-8 school in the Pacific Northwest in the spring of 2009. With 50% of the student body eligible for free or reduced-price meals, the school is comprised of 50% white, 16% Hispanic, 4% Asian, 2% Black, and 2% Native American students. In 2007, students at the participating school outperformed their peers in both the district and the state on the statewide reading assessment. Fully 90% of third-grade students at the participating school tested proficient on the state reading test, compared to 84% for the district overall and 82% for the state. In fifth grade, 83% of students at the participating school scored proficient, compared to 63% for the district and 71% for the state. In eighth grade, 70% of students at the participating school tested proficient, compared to 66% for the district and 68%

for the state (demographic information retrieved from www.schoolmatters.com on May 4, 2009).

Because we wanted representation from early elementary, late elementary, and middle school grades, we opted to recruit participants from grades 1, 3, 5, and 8. Teachers were recruited by their principal to participate in the study and were given $50 for classroom supplies and a $150 gift card for participating in the study. All students in attendance on the days the assessments were administered participated in the study.

Measures

We selected a sub-set of forms (roughly 40% of those available from selected grades on the easyCBM assessment system) to use in this study (see Table 1).

Table 1

Measures Used in This Study (# indicates how many unique forms of this measure were used)

Measure	Grade			
	1	3	5	8
Phoneme Segmenting	4	0	0	0
Letter Names	4	0	0	0
Letter Sounds	4	0	0	0
Word Reading Fluency	4	4	0	0
Passage Reading Fluency	4	4	6	6
Reading Comprehension	0	2	2	2

Each student participated in the testing on two separate testing sessions on two different days, one week apart. On day one, students in first grade completed three phoneme segmentation tests, three letter names tests, three letter sounds tests, three word reading fluency tests, and three passage reading fluency tests. Students in third grade completed three word reading fluency tests and three passage reading fluency tests. Students in fifth grade completed four passage reading fluency tests. Students in eighth grade completed four passage reading fluency tests.

One week later, students in first grade repeated two of each type of measure while students in all other grades repeated two of the measures administered in week one and some additional alternate forms of the measures, with the exact number depending on grade level. Thus, during the second testing session, Grade 3 students completed a total of three word reading fluency and three passage reading measures. Students in fifth and eighth grade repeated two of the week-one passage reading fluency measures and an additional two new forms of that measure.

In addition, students in grades three, five, and eight completed two reading comprehension tests on different days during the two-week testing window. The comprehension tests were group administered by the teacher. All other measures were individually administered by trained research assistants.

Phoneme segmentation measure. The phoneme segmentation measure is administered entirely orally. An assessor reads from a list of words and asks the student to segment each word into its constituent phonemes. Students receive one point for each phoneme segmented correctly. Self-corrections are scored as correct. This measure is administered individually for 60 seconds.

Letter names measure. In the letter names measure, students are presented with a sheet of paper on which letters in both their capital and lower case forms are printed in a table. Students

are given 30 seconds to name as many letters as they can, reading across the paper from left to right, then down to the next row. Errors and skipped letters are counted as incorrect while self-corrections and letters read correctly are counted as correct; the student receives one point for each correct response.

Letter sounds measure. The letter sounds measure is identical to the letter names measure except students are prompted to produce the sound the letter makes rather than its name. In addition, some common digraphs (ph, th, sh) are included on this measure. Students are again given 30 seconds to complete this measure, and the scoring rules are the same as for the letter names measure.

Word reading fluency measure. For the word reading fluency measure, students are shown a piece of paper with a variety of decodable and sight-words arranged in a table. They are instructed to read the words aloud, moving left to right and then down the rows. Errors and skipped words are counted as incorrect while self-corrections and words read correctly are counted as correct. The student receives one point for every correct response and has 30 seconds to complete the measure.

Passage reading fluency measure. On the passage reading fluency measure, students are given 60 seconds to read aloud a short (approximately 250 word) narrative passage presented to them on a single side of a sheet of paper. Assessors follow along on their own test protocol, marking as errors any words skipped or read incorrectly. If a student pauses more than three seconds on a word, the assessor supplies the word and marks it as incorrect. As in the other measures, self-corrections are counted as correct. The passages used are written to be at middle of the year reading level for each grade. The score, total words read correctly, is calculated by subtracting the number of errors from the total words read.

Reading comprehension measure. Unlike the other measures, the reading comprehension measures on easyCBM™ are designed for group administration. Students first read an original work of narrative fiction, approximately 1,500 words long, and then answer 20 multiple choice questions based on the story. Of the questions, seven sample literal comprehension, seven inferential comprehension, and six evaluative comprehension. Each question is comprised of the question stem and three possible answers: the correct answer and two incorrect but plausible distractors. Each comprehension measure has a total of 20 points possible; students earn one point for every question they answer correctly.

Analysis

We calculated descriptive statistics and correlation coefficients for each of the measures. To evaluate alternate-form reliability, we used correlations between alternate forms of the measures administered in the same grade. To evaluate test-retest reliability, we used correlations between each measure administered in session 1 and the same measure administered a week later in session 2. In our final analysis, we calculated the standard error of measure for each test type, using the median score from the different forms and the average standard deviation.

Results

Approximately 50 students per grade participated in the reliability study. In this section, we first present descriptive statistics, then the results of our alternate forms reliability study, ending with the results of our test-retest reliability study.

Descriptive Statistics

Descriptive statistics from the first testing session are presented in Table 2 for the students in first grade.

Table 2
Descriptive Statistics: Grade 1 Measures, Session 1

Measure	N	M	SD
SEG 1_1	48	46.69	11.99
SEG 1_3	48	48.44	13.00
SEG 1_5	48	48.44	13.30
LN 1_1	48	39.83	6.28
LN 1_3	48	40.96	7.13
LN 1_5	48	40.00	7.33
LS 1_1	48	25.09	6.84
LS 1_3	48	28.85	6.44
LS 1_5	48	30.60	7.47
WRF 1_1	48	25.14	12.63
WRF 1_3	48	25.52	13.22
WRF 1_5	48	26.36	12.69
PRF 1_1	48	71.84	40.27
PRF 1_3	48	62.63	36.20
PRF 1_5	48	65.75	39.14

Table 3 presents descriptive statistics from the second testing session for first-grade students.

Table 3
Descriptive Statistics: Grade 1 Measures, Session 2

Measure	N	M	SD
SEG 1_3	52	48.13	15.78
SEG 1_5	52	49.17	15.96
LN 1_3	52	40.63	8.03
LN 1_5	52	40.89	8.60
LS 1_3	52	29.74	7.36
LS 1_5	52	31.98	7.27
WRF 1_3	52	27.90	12.67
WRF 1_5	52	27.62	12.86
PRF 1_3	52	69.27	37.75
PRF 1_5	52	73.39	43.95

Of the first-grade sample, four students were in attendance on the second day of testing who had been absent during the first testing session.

Descriptive statistics for the third-grade sample are presented in Table 4 (Session 1) and Table 5 (Session 2).

Table 4
Descriptive Statistics: Grade 3 Measures, Session 1

Measure	N	M	SD
WRF 3_1	48	37.13	12.96
WRF 3_3	48	38.19	12.45
WRF 3_5	48	37.25	13.97
PRF 3_1	48	130.92	40.23
PRF 3_3	48	126.73	38.88
PRF 3_5	48	126.48	40.05
MCRC 3_14	47	14.09	2.76

Table 5
Descriptive Statistics: Grade 3 Measures, Session 2

Measure	N	M	SD
WRF 3_3	48	38.50	12.37
WRF 3_5	48	40.13	13.87
WRF 3_7	48	41.19	13.74
PRF 3_3	48	137.48	40.52
PRF 3_5	48	137.79	44.82
PRF 3_7	48	121.13	40.51
MCRC 3_15	44	13.27	2.59

Of the third-grade sample, one student failed to complete the first comprehension measure, and four students failed to complete the second comprehension measure. These students were absent the days the comprehension tests were administered to the rest of the class.

Table 6 presents descriptive statistics for the fifth-grade students (Session 1), while Table 7 provides this information for the same group of students on the second testing session.

Table 6
Descriptive Statistics: Grade 5 Measures, Session 1

Measure	N	M	SD
PRF 5_1	54	150.19	38.07
PRF 5_3	54	169.50	45.81
PRF 5_5	54	166.50	43.74
PRF 5_7	54	162.54	43.68
MCRC 5_14	54	14.74	3.65

Table 7
Descriptive Statistics: Grade 5 Measures, Session 2

Measure	N	M	SD
PRF 5_5	49	184.46	38.77
PRF 5_6	49	183.83	39.27
PRF 5_8	49	168.65	38.15
PRF 5_9	49	187.35	44.49
MCRC 5_15	57	15.11	3.31

In the fifth-grade sample, five students who had been present for the first individual testing session were absent for the second testing session. In all, three students who had been absent for all of the individually-administered measures were present when their teacher group-administered the second comprehension test, thus the number of students varies by measure.

Descriptive statistics for the eighth-grade sample are presented in Tables 8 and 9.

Table 8
Descriptive Statistics: Grade 8 Measures, Day 1

Measure	N	M	SD
PRF 8_1	59	186.12	41.51
PRF 8_3	59	192.12	36.45
PRF 8_5	59	185.34	40.47
PRF 8_7	59	197.17	39.17
MCRC 8_7	48	13.73	3.13

Table 9
Descriptive Statistics: Grade 8 Passage Reading Fluency Measures, Day 2

Measure	N	M	SD
PRF 8_5	58	199.22	40.17
PRF 8_6	58	207.70	37.29
PRF 8_8	58	192.02	34.87
PRF 8_9	58	196.81	37.64
MCRC 8_9	48	15.42	2.36

In all, 11 eighth-grade students were absent from the first group-administered testing session and so did not complete the first comprehension test, while 10 were absent from the second group-administered testing session.

Alternate Form Reliability

Alternate form reliability was analyzed using bi-variate correlations. We present the correlations between the different forms of each measure, organized by grade.

Grade 1. Results for the Grade 1 measures are presented in Tables 10 – 14. We found a strong positive relationship between the alternate forms of the Phoneme Segmenting measure, with correlations ranging from .86 to .91.

Table 10
Correlation between Alternate Forms of Phoneme Segmenting Measure

	Seg 1_3	Seg 1_5
Seg 1_1	.91	.86
Seg 1_3		.91

We also found a strong positive relationship between the alternate forms of the Letter Names measure, with correlations ranging from .82 to .89.

Table 11
Correlation between Alternate Forms of Letter Names Measure

	LN 1_3	LN 1_5
LN 1_1	.89	.82
LN 1_3		.85

Although the relationship between the alternate forms of the Letter Sounds measure was not quite as strong (with correlations ranging from .76 to .88), it was still a strong positive correlation.

Table 12
Correlation between Alternate Forms of Letter Sounds Measure

	LS 1_3	LS 1_5
LS 1_1	.76	.76
LS 1_3		.88

The relationship between the alternate forms of the Word Reading Fluency measure at grade 1 was extremely strong, with correlations ranging from .95 to .96.

Table 13
Correlation between Alternate Forms of Word Reading Fluency Measure, Day 1

	WRF 1_3	WRF 1_5
WRF 1_1	.96	.95
WRF 1_3		.95

Similarly, we found an extremely strong positive relationship between the alternate forms of the Passage Reading Fluency measure, with correlations ranging from .95 to .97.

Table 14
Correlation between Alternate Forms of Passage Reading Fluency Measure, Day 1

	PRF 1_3	PRF 1_5
PRF 1_1	.95	.96
PRF 1_3		.97

Grade 3. In third grade, we found a similar pattern of results (see Tables 15-16). There was a strong positive relationship between the alternate forms of the Word Reading Fluency measure, with correlations ranging from .87 to .93, and an extremely strong positive relationship between the alternate forms of the Passage Reading Fluency measure, with correlations ranging from .94 to .95.

Table 15
Correlation between Alternate Forms of Grade 3 Word Reading Fluency Measures

	WRF 3_3	WRF 3_5	WRF 3_7
WRF 3_1	0.92	0.90	.87
WRF 3_3		0.93	.89
WRF 3_5			.93

Table 16
Correlation between Alternate Forms of Grade 3 Passage Reading Fluency Measures

	PRF 3_3	PRF 3_5	PRF 3_7
PRF 3_1	0.95	0.94	.94
PRF 3_3		0.94	.95
PRF 3_5			.94

We found a weak positive correlation (R = .26) between the two comprehension measures administered to this third-grade sample.

Grade 5. We found a similar pattern of results in our analyses of the fifth-grade measures, with strong positive relationships between the alternate forms of the Passage Reading Fluency measure (correlations ranging from .87 to .96) and a weaker, but still positive, relationship between the alternate forms of the comprehension measure (R = .59).

Table 17
Correlation between Alternate Forms of Passage Reading Fluency Measure, Grade 5

	PRF 5_3	PRF 5_5	PRF 5_6	PRF 5_8	PRF 5_9
PRF 5_1	.91	.93	.94	.87	.90
PRF 5_3		.96	.94	.88	.91
PRF 5_5			.95	.90	.92
PRF 5_6				.88	.92
PRF 5_8					.95

Grade 8. As in the other grades, we found a strong positive relationship between the alternate forms of the Passage Reading Fluency measure in grade 8, with correlations ranging from .87 to .95. We found a weak positive relationship between the two alternate forms of the comprehension measure (R = .35).

Table 18
Correlation between Alternate Forms of Passage Reading Fluency Measure, Grade 8

	PRF 8_3	PRF 8_5	PRF 8_6	PRF 8_8	PRF 8_9
PRF 8_1	.92	.92	.90	.89	.92
PRF 8_3		.95	.91	.91	.89
PRF 8_5			.94	.92	.90
PRF 8_6				.89	.87
PRF 8_8					.92

Test-Retest Reliability

To evaluate test-retest reliability, we correlated performance on two forms of each measure type administered in both testing sessions. Table 19 presents results of these analyses.

Table 19

Test-Retest Results

Test Form	Measure Type				
	Phoneme Segmenting	Letter Names	Letter Sounds	Word Reading Fluency	Passage Reading Fluency
1_3	.45	.79	.68	.95	.96
1_5	.47	.82	.64	.94	.97
3_3				.92	.94
3_5				.94	.93
5_5					.92
5_6					.94
8_5					.91
8_6					.91

Table 20 presents the results of our analysis of the standard error of measure for each test type.

Table 20
Standard Error of Measure for Each Test Type

Session 1	Measure-grade	AVE SD	SEM
Median r=.91	SEG-grade 1	12.76	3.83
Median r=.85	LN-grade1	6.91	2.68
Median r=.76	LS-grade1	6.92	3.39
Median r=.95	WRF-grade 1	12.85	2.87
Median r=.96	PRF-grade 1	38.54	7.71
Session 2	Measure-grade	AVE SD	SEM
Median r=.91	SEG-grade 1	15.87	4.76
Median r=.85	LN-grade1	8.315	3.22
Median r=.76	LS-grade1	7.315	3.58
Median r=.95	WRF-grade 1	12.765	2.85
Median r=.96	PRF-grade 1	40.85	8.17
Session 1	Measure-grade	AVE SD	SEM
Median r=.91	WRF-grade 3	13.13	3.94
Median r=.94	PRF-grade 3	39.72	9.73
Median r=.26	MCRC-grade 3	2.76	2.37
Session 2	Measure-grade	AVE SD	SEM
Median r=.91	WRF-grade 3	13.33	4.00
Median r=.94	PRF-grade 3	41.95	10.28
Median r=.26	MCRC-grade 3	2.59	2.23
Session 1	Measure-grade	AVE SD	SEM
Median r=.92	PRF-grade 5	42.83	12.11
Median r=.59	MCRC-grade 5	3.65	2.34
Session 2	Measure-grade	AVE SD	SEM
Median r=.92	PRF-grade 5	40.17	11.36
Median r=.59	MCRC-grade 5	3.31	2.12
Session 1	Measure-grade	AVE SD	SEM
Median r=.92	PRF-grade 5	39.40	11.14
Median r=.35	MCRC-grade 5	3.13	2.52
Session 2	Measure-grade	AVE SD	SEM
Median r=.92	PRF-grade 5	37.49	10.60
Median r=.35	MCRC-grade 5	2.36	1.90

Discussion

The results of this study provide evidence of the reliability of the easyCBM reading measures. Although we tested only a small proportion of the total measures and forms available on easycbm.com, our findings suggest that the measures are reliable. The vast majority of our alternate form reliability analyses indicate extremely strong correlations between the alternate forms of the measures, especially the Word and Passage Reading Fluency measures. Across all forms of all measures (with the exception of Grade 1 Letter Sounds and the reading comprehension measure), we found reliability co-efficients in the range that would normally be expected of statewide large-scale tests. The standard error of measure also appear to fall within the range of acceptable values.

With an alternate-form reliability co-efficient of .76 to .88, the Letter Sounds measure is still well within the bounds of acceptable reliability for progress monitoring. The reading comprehension measures, however, present more reason for concern. Additional research addressing the reliability of these measures is needed. It is possible that the administration of these measures in the current study introduced additional sources of variance. All other measures in this study were individually-administered by a small team of trained research assistants, under the direct supervision of the first author. Due to time constraints, however, the reading comprehension tests were group-administered by students' regular classroom teachers. Perhaps these teachers would have benefitted from more structured training in test administration or more specific instructions on expectations for establishing a consistent and appropriate testing environment, free of unnecessary distractions.

References

Alonzo, J., Tindal, G., Ulmer, K., & Glasgow, A. (2006). *easyCBM online progress monitoring assessment system.* Eugene, OR: Center for Educational Assessment Accountability. Available at http://easycbm.com.

www.schoolmatters.com, demographic information about participating school retrieved on May 4, 2009.