Technical Report #1004

Technical Adequacy of the easyCBM Grade 2 Reading Measures

Elisa Jamgochian

Bitnara Jasmine Park

Joseph F.T. Nese

Cheng-Fei Lai

Leilani Sáez

Daniel Anderson

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

Note: Funds for the data set used to generate this report come from a federal grant awarded to the UO from the Institute for Education Sciences, U.S. Department of Education: *Postdoctoral Fellowships on Progress Monitoring in Reading and Math* (Award #R305B080004 funded from 2008-2012).

Support for this work was also provided by the Center on Teaching and Learning (CTL) and Behavioral Research and Teaching (BRT), College of Education at the University of Oregon.

Abstract

In this technical report, we provide reliability and validity evidence for the easyCBM® Reading measures for grade 2 (word and passage reading fluency and multiple choice reading comprehension). Evidence for reliability includes internal consistency and item invariance. Evidence for validity includes concurrent, predictive, and construct validities for performance level scores, as well as slope of improvement. Reliability of alternate forms and content validity were analyzed previously (references to previous technical reports are provided). Internal consistency, split-half reliability, and reliability of growth slopes were moderate. For concurrent and predictive validities, multiple choice reading comprehension was a better predictor of SAT-10 scores than either word or passage reading fluency. Construct validity was supported by strong model fit indices. Overall, predictive validity coefficients for all students on all measures were positive and low.

Technical Adequacy of the easyCBM® Grade 2 Reading Measures

Progress monitoring assessments are a key component of many school improvement efforts, including the Response to Intervention (RTI) approach to meeting students' academic needs. In an RTI approach, teachers first administer a screening or benchmarking assessment to identify students who need supplemental interventions to meet grade-level expectations, then use a series of progress monitoring measures to evaluate the effectiveness of the interventions they are using with the students. When students fail to show expected levels of progress (as indicated by 'flat line scores' or little improvement on repeated measures over time), teachers use this information to help them make instructional modifications with the goal of finding an intervention or combination of instructional approaches that will enable each student to make adequate progress toward achieving grade level proficiency and content standards. In such a system, it is critical to have reliable measures that assess the target construct and are sensitive enough to detect improvement in skill over short periods of time.

Conceptual Framework: Curriculum-Based Measurement and Progress Monitoring

Curriculum-based measurement (CBM), long a bastion of special education, is gaining support among general education teachers seeking a way to monitor the progress their students are making toward achieving grade-level proficiency in key skill and content areas. While reading in particular has received a great deal of attention in the CBM literature, a growing body of work is beginning to appear in the area of mathematics CBM.

By definition, CBM is a formative assessment approach. By sampling skills related to the curricular content covered in a given year of instruction yet not specifically associated with a particular textbook, CBMs provide teachers with a snapshot of their students'

current level of proficiency in a particular content area as well as a mechanism for tracking the progress students make in gaining desired academic skills throughout the year. Historically, CBMs have been very brief individually administered measures (Deno, 2003; Good, Gruba, & Kaminski, 2002), yet they are not limited to the 'one minute timed probes' that many people associate them with.

In one of the early definitions of curriculum-based measurement (CBM), Deno (1987) stated that "the term curriculum-based assessment, generally refers to any approach that uses direct observation and recording of a student's performance in the local school curriculum as a basis for gathering information to make instructional decisions…The term curriculum-based measurement refers to a specific set of procedures created through a research and development program … and grew out of the *Data-Based Program Modification* system developed by Deno and Mirkin (1977)" (p. 41). He noted that CBM is distinct from many teacher-made classroom assessments in two important respects: (a) the procedures reflect technically adequate measures ("they possess reliability and validity to a degree that equals or exceeds that of most achievement tests" (p. 41), and (b) "growth is described by an increasing score on a standard, or constant task. The most common application of CBM requires that a student's performance in each curriculum area be measured on a single global task repeatedly across time" (p. 41).

In the three decades since Deno and his colleagues introduced CBM, *progress monitoring probes,* as they have come to be called, have increased in popularity, and they are now a regular part of many schools' educational programs (Alonzo, Tindal, & Ketterlin-Geller, & 2006). However, CBMs – even those widely used across the United States – often lack the psychometric properties expected of modern technically-adequate assessments.

Although the precision of instrument development has advanced tremendously in the past 30 years with the advent of more sophisticated statistical techniques for analyzing tests on an item by item basis rather than relying exclusively on comparisons of means and standard deviations to evaluate comparability of alternate forms, the world of CBMs has not always kept pace with these statistical advances.

A key feature of assessments designed for progress monitoring is that alternate forms must be as equivalent as possible to allow meaningful interpretation of student performance data across time. Without such cross-form equivalence, changes in scores from one testing session to the next are difficult to attribute to changes in student skill or knowledge. Improvements in student scores may, in fact, be an artifact of the second form of the assessment being easier than the form that was administered first. The advent of more sophisticated data analysis techniques (such as the Rasch modeling used in the development of the easyCBM® progress monitoring and benchmarking assessments) have made it possible to increase the precision with which we develop and evaluate the quality of assessment tools.

In this technical report, we provide the results of a series of studies to evaluate the technical adequacy of the easyCBM® progress monitoring assessments in reading, designed for use with students in Grade 2. This assessment system was developed to be used by educators interested in monitoring the progress their students make in the area of acquiring skills in the constructs of oral reading fluency and comprehension. Additional technical reports report the results of similar studies of the easyCBM® assessments in mathematics (Anderson et al, 2010; Nese et al., 2010) and in reading with a focus on

Kindergarten and first grade measures (Lai et al., 2010) and grade three through eight measures (Saéz et al., 2010).

The easyCBM® Progress Monitoring Assessments

The online easyCBM® progress monitoring assessment system, launched in September 2006 as part of a Model Demonstration Center on Progress Monitoring, was funded by the Office of Special Education Programs (OSEP). At the time this technical report was published, there were 111,977 teachers with easyCBM® accounts, representing schools and districts spread across every state in the country. During the 2008-2009 school year, the system had an average of 305 new accounts registered each week, and the popularity of the system continues to grow. In the month of October 2010, alone, 11,885 new teachers registered for accounts. The online assessment system provides both universal screener assessments for fall, winter, and spring administration and multiple alternate forms of a variety of progress monitoring measures designed for use in K-8 school settings.

As part of state funding for Response to Intervention (RTI), states need technically adequate measures for monitoring progress. Given the increasing popularity of the easyCBM® online assessment system, it is imperative that a thorough analysis of the measures' technical adequacy be conducted and the results shared with research and practitioner communities. This technical report addresses that need directly, providing the results of a series of studies examining the technical adequacy of the 2009 / 2010 version of the easyCBM® assessments in reading.

Methods

In this section, we describe the setting and subjects, measures, and data analysis

procedures.

Setting and Subjects

The data were gathered during the 2009-2010 school year from 71 schools in three districts in the Pacific Northwest. All students in attendance at the schools during the assessment period participated in the testing. The second grade word reading fluency (WRF) sample ranged from 2,154-2207 students (Fall – Spring), the passage reading fluency (PRF) sample ranged from 2,205-2,236 students, and the multiple choice reading comprehension (MCRC) sample ranged from 2,144-2301 students; 205 students took the SAT-10. Approximately 49% of the sample was female. No other demographic data were available for the grade two sample.

Measures

Assessment data used in this study included scores from the fall, winter, and spring administrations of the easyCBM® reading measures for grade 2 and scores from the SAT-10.

easyCBM® word reading fluency (WRF). 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 60 seconds to complete the measure.

easyCBM® passage reading fluency (PRF). 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 in one minute.

easyCBM® Multiple-choice reading comprehension (MCRC). Unlike the other measures, the reading comprehension measures on easyCBM® are designed for computer-based administration in a group setting. Students first read an original work of narrative fiction, approximately 900 words long, and then answer 12 multiple choice questions based on the story. The questions sample literal, inferential, and evaluative comprehension. Each question is comprised of the question stem and three possible answers: the correct answer and two incorrect but plausible distractors. A total of 12 points are possible for the 2nd grade measure; students earn one point for every question they answer correctly. Questions to which students do not respond are counted as incorrect.

SAT-10 Reading. The SAT-10 Reading assessment contains subtests for phonemic awareness, decoding, phonics, vocabulary, and comprehension of literary, informational, and functional text.

Data Analysis Procedures

Reliablity. Reliability of alternate forms was analyzed previously (see below for references). Procedures for analyzing the reliability of performance level scores and growth slopes are discussed below.

Alternate forms. For information about the development of alternate forms for word and passage reading fluency, please see *The development of word and passage reading fluency measures in a progress monitoring assessment system* (Technical Report No. 41; Alonzo, Liu, & Tindal [2007]). For information about the development of MCRC alternate forms, please see *Examining the technical adequacy of second-grade reading comprehension measures in a*

progress monitoring assessment system (Technical Report No. 0808; Alonzo, Liu, & Tindal [2008]).

Reliability of performance level score. Reliability was evaluated for the MCRC measure only. (No item-level data were available for the WRF and PRF subtests). Cronbach's alpha and split-half reliability were calculated using SPSS. Listwise deletion was specified to remove missing variables in the analysis.

Reliability of slope. A two-level hierarchical linear growth model represented student reading growth within one academic year, with *time* at level-1 and *student* at level-2. The easyCBM® reading measures were collected in a multiple-time-point design during the fall, winter, and spring, and were used as the criterion variable. Each student's reading growth was represented by an individual growth trajectory over time. Analyses were separated by quartile based on fall easyCBM® measure (PRF or MCRC), in effect conditioning the results on fall score status. The fixed and random effects for the intercept and slope and the reliability of the growth estimates were reported. The growth reliability of the growth estimates was defined as the ratio between the level-2 variance component and the sum of the level-2 and level-1 components, with the latter divided by the number of students within that particular group (Raudenbush & Bryk, 2002). All analyses were conducted using *R*, the free online statistical software (R Development Core Team, 2010).

Validity of performance level score. Analyses of content, concurrent, predictive, and construct validities are discussed below.

Content validity. For information regarding the content development and validity of the easyCBM® reading measures, see technical reports on Word and Passage Reading Fluency (Alonzo & Tindal, 2007); and Reading Comprehension (Alonzo, Liu, & Tindal, 2007).

***Concurrent validity*.** Concurrent validity was examined two ways. First, correlations between the Spring easyCBM® scores and the SAT-10 standard score were calculated. Next, a regression analysis was conducted using the Spring easyCBM® measures and an outcome measure (SAT-10). Total scores were used for the easyCBM® measures and standard scores were used for SAT-10. Pairwise deletion was specified to remove missing variables in the analyses.

Predictive validity. Predictive validity was examined in a number of ways. First, correlations between the Fall and Winter measures were calculated. Next, regression analyses were conducted to analyze the predictive validity of the Fall and Winter easyCBM® reading measures individually and combined by season. Total scores were used for the easyCBM® measures and standard scores were used for SAT-10 (outcome measure). Pairwise deletion was specified to remove missing variables in the analyses.

***Construct validity*.** Confirmatory factor analysis was used to investigate the construct validity of the easyCBM® reading measures. Correlation coefficients were calculated between each of the easyCBM® measures at each time point (Fall, Winter, Spring) and the SAT-10 using SPSS. A hypothesized 3-factor model, PRF, WRF, and MCRC, was analyzed using Mplus (Muthen & Muthen, 2002; Figure 1). Because of the structure of the available data, total scores were used for PRF and WRF, and item-level scores were used for MCRC. Model fit indices (Comparative Fit Index [CFI], Tucker-Lewis Index [TLI], and Root Mean Square Error of Approximation [RMSEA]) are reported in Table 33.

Predictive validity of the slope. We examined students' rate of growth (slope) in a year using a two-level hierarchical linear growth model (HLM; Raudenbush & Bryk, 2002). The level-1 model was represented by time and with the level-2 model by student. The easyCBM®

reading scores at Fall, Winter, and Spring were used as the dependent variable. The level-2 residuals from the final model were correlated with students' performance scores on the SAT-10.

Results

To aid in interpretation, we present the results of each of our different analyses separately.

Reliability

In this section, we report the reliability of the grade 2 easyCBM® reading measures (WRF, PRF, and MCRC). Reliability of alternate forms can be found in previous technical reports, referenced below. Reliability of performance level score was evaluated for the MCRC measure only, and reliability of the growth slope was analyzed for all three subtests.

Alternate forms. (See technical reports on the development of WRF and PRF [Alonzo, Liu, & Tindal, 2007] and MCRC [Alonzo, Liu, & Tindal, 2008]).

Reliability of performance level score. Data for this analysis included scores from the MCRC measure only. There are 12 items on the grade 2 version of this assessment. The sample size ranged from 1,696 in Fall to 2,039 in Spring. The mean ranged from 6.70 (*SD* = 2.72*)* in Fall to 9.11 (*SD* = 2.58) in Spring. (See table 1 for descriptive statistics). Across the three time points (Fall, Winter, Spring), Cronbach's alpha ranged from .679 to .748 (Table 2).

Split-half reliability coefficients were calculated by comparing the first 6 items of the MCRC measure to the last 6 items. Across the three time points, the correlation between forms ranged from .499 to .599. (See tables 3 - 5).

Reliability of growth slope. For the grade 2 word reading fluency (WRF) measure, the reliability of the growth slope for the students in the first quartile was .77, the reliability for students in the second quartile was .61, the reliability for students in the third quartile was .60,

and the reliability for students in the fourth quartile was .50. For the grade 2 passage reading fluency (PRF) measure, the reliability of the growth slope for the students in the first quartile was .75, the reliability for students in the second quartile was .70, the reliability for students in the third quartile was .68, and the reliability for students in the fourth quartile was .25. For the grade 2 multiple choice reading comprehension measure (MCRC), the reliability of the growth slope for the student in the first quartile was .48, the reliability for the student in the second quartile was .62, and the reliability for the student in the fourth quartile was .14 (Table 6).

Validity

In this section, we report the results of the validity analyses for the grade 2 easyCBM® reading measures (WRF, PRF, and MCRC). Content validity evidence can be found in previous technical reports, referenced below. Evidence for concurrent, predictive and construct validities are presented below, as well as the predictive validity of the slope.

Validity of the performance level score. Information related to four types of validity comprise evidence for the validity of the performance level score: content, concurrent, predictive, and construct. Descriptive statistics are presented in Table 7. Results for each type of validity evidence are described below.

Content validity. (See technical reports on Word and Passage Reading Fluency [Alonzo & Tindal, 2007]; and Reading Comprehension [Alonzo, Liu, & Tindal, 2007]).

Concurrent validity. None of the grade 2 easyCBM® measures were significantly correlated with the SAT-10. Correlations ranged from .036 - .052 (Table 8). Regression analyses were conducted for each of the measures separately and combined (easyCBM® model). Individually, the measures were weak predictors of SAT-10 scores. In the combined model, MCRC was the best predictor. Descriptive statistics are presented in Table 7 and regression

model summaries are presented in Tables 9 - 16.

Predictive validity. Fall WRF and PRF were significantly correlated with SAT-10 (.154 [$p < .05$] and .194 [$p < .01$], respectively). Winter WRF and PRF were also significantly correlated with SAT-10 (.151 [$p < .05$] and .221 [$p < .01$], respectively). Neither Fall nor Winter MCRC was significantly correlated with SAT-10. Regression analyses were conducted for each of the measures separately and combined by season. For Fall, the individual measures were weak predictors of SAT-10. In the combined model, WRF and MCRC were better predictors than PRF. For Winter, individually, MCRC was the best predictor of SAT-10. In the combined model, MCRC was also the best predictor, WRF was moderate, and PRF was negative. Descriptive statistics are presented in Table 7 and regression model summaries are presented in Tables 17 – 24 (Fall) and 25 - 32 (Winter).

Construct validity. Correlation coefficients were calculated between each of the easyCBM® measures at three time points (Fall, Winter, Spring) and the SAT-10 using SPSS (Table 8). Correlations between easyCBM® measures for each time point and across time points were all significant (.799 - .999; $p < .01$). As noted above, correlations between the Fall and Winter WRF measure and SAT-10 were significant ($p < .05$) as were correlations between Fall and Winter PRF ($p < .01$). Confirmatory factor analysis was used to investigate the construct validity of the easyCBM® reading measures. The hypothesized 3-factor model, PRF, WRF, and MCRC, was analyzed using Mplus (Muthen & Muthen, 2002; Figure 1). Model fit indices indicate a favorable fit (CFI/TLI > .95; RMSEA <.05) and are reported in Table 33.

Predictive validity of the slope. Rate of growth for students in grade two in WRF was low but positive across the first through the third quartiles. A low and negative predictive validity coefficient was reported for students in the fourth quartile in WRF. A similar trend was

observed for the fourth quartile in PRF as well. Students in quartiles one through three had positive rate of growth, with low to moderate predictive validity coefficients. In the multiple choice reading comprehension measures (MCRC), students in the first two quartiles had a moderate rate of growth. No predictive validity coefficients could be obtained from the analysis for any students in the third quartile due to convergence errors. The rate of growth for the fourth quartile was positive, but relatively low (see Table 34).

Discussion

In this report, we presented data to support the technical adequacy of the grade 2 easyCBM® measures (word and passage reading fluency [WRF, PRF] and multiple choice reading comprehension [MCRC]). Internal consistency and split-half reliability for the MCRC measure were moderate. In general, reliability of the growth slopes was moderate as well (.50 - .77), with the exception of low reliability coefficients for growth slopes for students in the 4th quartile on the PRF and MCRC measures (.25 and .14, respectively). That is, students who began the year as higher achieving demonstrated less growth, which could be the result of a ceiling effect.

For concurrent and predictive validities, MCRC was a better predictor of SAT-10 scores than either WRF or PRF. Model fit indices support the hypothesized 3-factor model (WRF, PRF, MCRC) of reading at 2nd grade, providing evidence for construct validity. Overall, predictive validity coefficients for all students on all measures were positive and low. Students with the lowest initial fall scores on the PRF and MCRC measures (first quartile) had the highest rate of growth compared to other quartiles with moderate predictive validity coefficients in the 0.50s and 0.60s. Predictive validity coefficients for other quartiles were generally in the lower range,

suggesting lower rate of growth. On the MCRC measures, overall predictive validity coefficients were higher at the first and second quartiles compared to the other measures.

References

Alonzo, J., Liu, K., & Tindal, G. (2007). *Examining the Technical Adequacy of Reading Comprehension Measures in a Progress Monitoring Assessment System (Technical Report # 41).* Eugene, OR: Behavioral Research and Teaching.

Alonzo, J., Liu, K., & Tindal, G. (2008). *Examining the technical adequacy of second-grade reading comprehension measures in a progress monitoring assessment system* (Technical Report No. 0808). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

Alonzo, J., & Tindal, G. (2007). *Examining the Technical Adequacy of Word and Passage Reading Fluency Measures in a Progress Monitoring Assessment System (Technical Report # 40*). Eugene, OR: Behavioral Research and Teaching.

Anderson, D., Lai, C.F., Nese, J.F.T., Park, B.J., Sáez, L , Jamgochian, E.M., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM primary-level mathematics measures (Grades K-2), 2009-2010 version.* (Technical Report #1006). Eugene, OR: Behavioral Research and Teaching.

Muthen. B. & Muthen, L. (2002). MPlus [computer software], Muthén & Muthén, Los Angeles, CA.

Nese, J.F.T., Lai, C. F., Anderson, D., Jamgochian, E.M., Kamata, A., Sáez, L , Park, B.J., Alonzo, J., & Tindal, G. (2010). *Technical adequacy of the easyCBM mathematics measures, (Grades 3-8), 2009-2010 version.* (Technical Report #1007). Eugene, OR: Behavioral Research and Teaching.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models. Applications and Data Analysis Methods*. Second edition. Thousand Oaks, CA: Sage Publications.

R Development Core Team (2010). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. http://www.R-project.org

Sáez, L , Park, B.J., Nese, J.F.T, Jamgochian, E.M., Lai, C.F., Anderson, D., Alonzo, J., & Tindal, G. (2010) *Technical adequacy of the easyCBM reading measures (Grades 3-8), 2009-2010 Version.* (Technical Report #1005). Eugene, OR: Behavioral Research and Teaching.

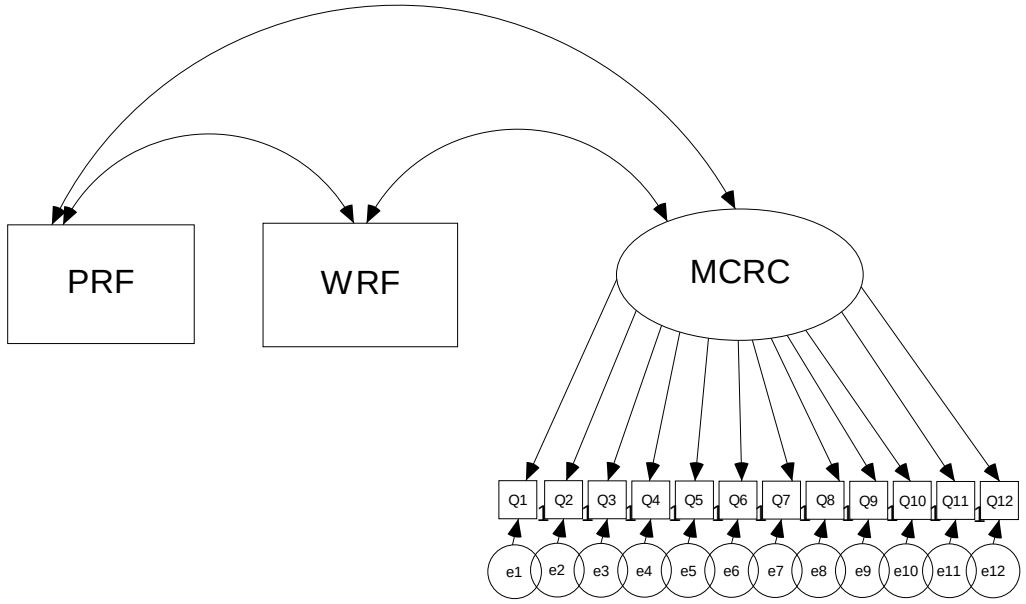

Figure 1. Hypothesized 3-Factor Model for easyCBM® Reading Grade 2
Note. WRF = Word Reading Fluency; PRF = Passage Reading Fluency; MCRC = Multiple Choice Reading Comprehension

Table 1

Grade 2 Descriptive Statistics – Reliability of Performance Level Score Analyses (MCRC)

	N	*M*	Variance	*SD*	*n of items*
Fall	1696	6.70	7.37	2.72	12
Winter	1983	8.07	7.51	2.74	12
Spring	2039	9.11	6.66	2.58	12

Table 2

Grade 2 Reliability Statistics – Internal Consistency: Multiple Choice Reading Comprehension (MCRC)

	N of Items	Cronbach's Alpha
Fall	12	.679
Winter	12	.733
Spring	12	.748

Table 3

Fall Grade 2 Reliability Statistics – Split Half: Multiple Choice Reading Comprehension (MCRC)

		Scale Statistics			
		Mean	Variance	Std. Deviation	N of Items
	Part 1	3.74	2.533	1.592	6[a]
	Part 2	2.96	2.383	1.544	6[b]
	Both Parts	6.70	7.369	2.715	12
Cronbach's Alpha	Part 1				.568
	Part 2				.472
Correlation Between Forms					.499
Spearman-Brown Coefficient	Equal Length				.666
	Unequal Length				.666
Guttman Split-Half Coefficient					.666

a. The items are: FallMCRCQ1C, FallMCRCQ2C, FallMCRCQ3C, FallMCRCQ4C, FallMCRCQ5C, FallMCRCQ6C.

b. The items are: FallMCRCQ7C, FallMCRCQ8C, FallMCRCQ9C, FallMCRCQ10C, FallMCRCQ11C, FallMCRCQ12C.

Table 4

Winter Grade 2 Reliability Statistics – Split Half: Multiple Choice Reading Comprehension (MCRC)

		Scale Statistics			
		Mean	Variance	Std. Deviation	N of Items
Part 1		4.36	2.298	1.516	6[a]
Part 2		3.71	2.471	1.572	6[b]
Both Parts		8.07	7.511	2.741	12
Cronbach's Alpha	Part 1				.606
	Part 2				.555
Correlation Between Forms					.575
Spearman-Brown Coefficient	Equal Length				.730
	Unequal Length				.730
Guttman Split-Half Coefficient					.730

a. The items are: WintMCRCQ1C, WintMCRCQ2C, WintMCRCQ3C, WintMCRCQ4C, WintMCRCQ5C, WintMCRCQ6C.

b. The items are: WintMCRCQ7C, WintMCRCQ8C, WintMCRCQ9C, WintMCRCQ10C, WintMCRCQ11C, WintMCRCQ12C.

Table 5

Spring Grade 2 Reliability Statistics – Split Half: Multiple Choice Reading Comprehension (MCRC)

		Scale Statistics			
		Mean	Variance	Std. Deviation	N of Items
	Part 1	4.62	1.827	1.352	6[a]
	Part 2	4.49	2.348	1.532	6[b]
	Both Parts	9.11	6.659	2.580	12
Cronbach's Alpha	Part 1				.544
	Part 2				.642
Correlation Between Forms					.599
Spearman-Brown Coefficient	Equal Length				.750
	Unequal Length				.750
Guttman Split-Half Coefficient					.746

a. The items are: SprMCRCQ1C, SprMCRCQ2C, SprMCRCQ3C, SprMCRCQ4C, SprMCRCQ5C, SprMCRCQ6C.

b. The items are: SprMCRCQ7C, SprMCRCQ8C, SprMCRCQ9C, SprMCRCQ10C, SprMCRCQ11C, SprMCRCQ12C.

Table 6

Grade 2 Reliability of Growth Slope: Word Reading Fluency, Passage Reading Fluency, and Multiple Choice Reading Comprehension

Measure	Student Group	Fixed effect, Intercept	*SE*	Level-1 residual variance	Reliability, Intercept	Fixed, effect, slope	*SE*	Variance, slope	Reliability, Slope	*n*
Word Reading Fluency	Quartile 1	8.78	0.26	27.89	0.58	11.88	0.30	32.37	0.77	546
	Quartile 2	27.11	0.27	37.41	0.40	15.05	0.28	20.15	0.61	531
	Quartile 3	45.26	0.30	54.47	0.14	13.71	0.33	28.26	0.60	537
	Quartile 4	70.95	0.65	134.32	0.72	8.80	0.47	45.50	0.50	534
Passage Reading Fluency	Quartile 1	17.73	0.45	81.34	0.63	15.47	0.48	84.13	0.75	550
	Quartile 2	42.53	0.37	72.82	0.32	22.84	0.42	56.22	0.70	531
	Quartile 3	66.28	0.40	85.41	0.34	23.07	0.45	61.28	0.68	536
	Quartile 4	115.97	1.30	267.51	0.88	16.62	0.56	30.04	0.25	530
Multiple Choice Reading Comprehension	Quartile 1	3.09	0.07	3.03	0.18	2.32	0.07	0.96	0.48	554
	Quartile 2	5.99	0.05	2.44	0.00	1.42	0.06	1.33	0.62	693
	Quartile 3	-	-	-	-	-	-	-	-	466[+]
	Quartile 4	10.44	0.05	1.00	0.12	0.22	0.04	0.05	0.14	398

[+] Parameters could not be estimated for this model.

Table 7

Grade 2 Descriptive Statistics – Validity of the Performance Level Score Analyses

	N	Minimum	Maximum	*M*	*SD*
Fall 2009					
Word Reading	2207	0	230	39.02	24.852
Passage Reading	2208	0	254	61.04	40.945
Comprehension	2144	0	12	6.57	2.914
Winter 2010					
Word Reading	2205	0	117	47.73	24.677
Passage Reading	2205	0	219	77.36	40.246
Comprehension	2189	0	12	7.99	2.846
Spring 2010					
Word Reading	2154	0	194	62.97	26.526
Passage Reading	2236	0	242	98.41	44.497
Comprehension	2301	0	12	9.07	2.707
SAT10 Reading Total Standard Score	205			608.80	39.770

Table 8

Grade 2 Correlation Matrix (easyCBM® N = 3675; SAT-10 N = 205)

		Fall 2009			Winter 2010			Spring 2010			
		WRF	PRF	MCRC	WRF	PRF	MCRC	WRF	PRF	MCRC	SAT-10
Fall 2009											
WRF	Pearson Correlation		.999**	.936**	.913**	.912**	.907**	.801**	.836**	.880**	.154*
	Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.028
PRF	Pearson Correlation			.934**	.912**	.912**	.906**	.799**	.835**	.877**	.194**
	Sig. (2-tailed)			.000	.000	.000	.000	.000	.000	.000	.005
MCRC	Pearson Correlation				.916**	.914**	.917**	.803**	.838**	.866**	.070
	Sig. (2-tailed)				.000	.000	.000	.000	.000	.000	.318
Winter 2010											
WRF	Pearson Correlation					.998**	.971**	.848**	.891**	.905**	.151*
	Sig. (2-tailed)					.000	.000	.000	.000	.000	.031
PRF	Pearson Correlation						.971**	.848**	.892**	.904**	.221**
	Sig. (2-tailed)						.000	.000	.000	.000	.001
MCRC	Pearson Correlation							.834**	.872**	.901**	-.081
	Sig. (2-tailed)							.000	.000	.000	.251
Spring 2010											
WRF	Pearson Correlation								.885**	.889**	.047
	Sig. (2-tailed)								.000	.000	.502
PRF	Pearson Correlation									.925**	.052
	Sig. (2-tailed)									.000	.458
MCRC	Pearson Correlation										.036
	Sig. (2-tailed)										.608

**Correlation is significant at the 0.01 level. *Correlation is significant at the 0.05 level.

Table 9

Grade 2 Model Summary – Concurrent Validity (Spring easyCBM® and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	*F* Change	df1	df2	Sig. *F* Change
1	.742[a]	.550	.529	27.291	.550	25.718	3	63	.000

a. Predictors: (Constant), Spr10MCRC, Spr10WRF, Spr10PRF

Table 10

Grade 2 Model Coefficients[a] – Concurrent Validity (Spring easyCBM® and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for *B*		Correlations		
		B	Std. Error	Beta	*t*	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	515.735	12.160		42.413	.000	491.436	540.034			
	Spr10WRF	.077	.307	.051	.249	.804	-.537	.690	.604	.031	.021
	Spr10PRF	.293	.195	.327	1.504	.138	-.096	.682	.657	.186	.127
	Spr10MCRC	6.552	1.606	.446	4.080	.000	3.342	9.761	.683	.457	.345

a. Dependent Variable: SAT10 Reading Total Standard Score - 2nd Grade

Table 11

Grade 2 Model Summary – Concurrent Validity (Spring WRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.047[a]	.002	-.003	39.824

Table 12

Grade 2 Model Coefficients[a] – Concurrent Validity (Spring WRF and SAT-10)

		Unstandardized Coefficients		Standardized Coefficients		
Model		*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	605.609	5.502		110.080	.000
	Spr10WRF	.004	.007	.047	.672	.502

Table 13

Grade 2 Model Summary – Concurrent Validity (Spring PRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.052[a]	.003	-.002	39.814

Table 14

Grade 2 Model Coefficients[a] – Concurrent Validity (Spring PRF and SAT-10)

		Unstandardized Coefficients		Standardized Coefficients		
Model		*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	606.895	3.781		160.512	.000
	Spr10PRF	.005	.007	.052	.744	.458

Table 15

Grade 2 Model Summary – Concurrent Validity (Spring MCRC and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.036[a]	.001	-.004	39.842

Table 16

Grade 2 Model Coefficients[a] – Concurrent Validity (Spring MCRC and SAT-10)

		Unstandardized Coefficients		Standardized Coefficients		
Model		*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	608.474	2.854		213.215	.000
	Spr10MCRC	.009	.017	.036	.514	.608

Table 17

Grade 2 Model Summary – Predictive Validity (Fall easyCBM® and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	*F* Change	df1	df2	Sig. *F* Change
1	.683[a]	.466	.458	29.286	.466	55.585	3	191	.000

a. Predictors: (Constant), Fall09MCRC, Fall09PRF, Fall09WRF

Table 18

Grade 2 Model Coefficients[a] – Predictive Validity (Fall easyCBM® and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for *B*		Correlations		
		B	Std. Error	Beta	*t*	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	556.451	5.194		107.140	.000	546.207	566.695			
	Fall09WRF	.655	.177	.410	3.708	.000	.307	1.004	.656	.259	.196
	Fall09PRF	.119	.106	.123	1.122	.263	-.090	.329	.617	.081	.059
	Fall09MCRC	2.970	.962	.218	3.088	.002	1.073	4.866	.559	.218	.163

a. Dependent Variable: SAT10 Reading Total Standard Score - 2nd Grade

Table 19

Grade 2 Model Summary – Predictive Validity (Fall WRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.154[a]	.024	.019	39.393

Table 20

Grade 2 Model Coefficients[a] – Predictive Validity (Fall WRF and SAT-10)

		Unstandardized Coefficients		Standardized Coefficients		
Model		*B*	Std. Error	Beta	*t*	Sig.
1	(Constant)	606.053	3.017		200.890	.000
	Fall09WRF	.030	.013	.154	2.220	.028

Table 21

Grade 2 Model Summary – Predictive Validity (Fall PRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.194[a]	.038	.033	39.109

Table 22

Grade 2 Model Coefficients[a] – Predictive Validity (Fall PRF and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	604.377	3.149		191.907	.000
	Fall09PRF	.038	.013	.194	2.822	.005

Table 23

Grade 2 Model Summary – Predictive Validity (Fall MCRC and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.070[a]	.005	.000	39.770

Table 24

Grade 2 Model Coefficients[a] – Predictive Validity (Fall MCRC and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	608.081	2.869		211.954	.000
	Fall09MCRC	.013	.013	.070	1.001	.318

Table 25

Grade 2 Model Summary – Predictive Validity (Winter easyCBM® and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	*F* Change	df1	df2	Sig. *F* Change
1	.685[a]	.470	.461	29.187	.470	57.541	3	195	.000

a. Predictors: (Constant), Wint10MCRC, Wint10WRF, Wint10PRF

Table 26

Grade 2 Model Coefficients[a] - Predictive Validity (Winter easyCBM® and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for *B*		Correlations		
		B	Std. Error	Beta	*t*	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part
1	(Constant)	537.208	6.185		86.856	.000	525.010	549.407			
	Wint10WRF	.660	.233	.409	2.834	.005	.201	1.119	.617	.199	.148
	Wint10PRF	-.066	.145	-.067	-.455	.650	-.351	.220	.589	-.033	-.024
	Wint10MCRC	5.658	1.001	.405	5.653	.000	3.684	7.632	.633	.375	.295

a. Dependent Variable: SAT10 Reading Total Standard Score - 2nd Grade

Table 27

Grade 2 Model Summary – Predictive Validity (Winter WRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.151[a]	.023	.018	39.412

Table 28

Grade 2 Model Coefficients[a] – Predictive Validity (Winter WRF and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	605.150	3.224		187.684	.000
	Wint10WRF	.052	.024	.151	2.174	.031

Table 29

Grade 2 Model Summary – Predictive Validity (Winter PRF and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.221[a]	.049	.044	38.879

Table 30

Grade 2 Model Coefficients[a] – Predictive Validity (Winter PRF and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	600.743	3.685		163.010	.000
	Wint10PRF	.076	.024	.221	3.234	.001

Table 31

Grade 2 Model Summary – Predictive Validity (Winter MCRC and SAT-10)

Model	*R*	*R Square*	Adjusted R Square	Std. Error of the Estimate
1	.081[a]	.006	.002	39.739

Table 32

Grade 2 Model Coefficients[a] – Predictive Validity (Winter MCRC and SAT-10)

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	*t*	Sig.
1	(Constant)	609.511	2.843		214.363	.000
	Wint10MCRC	-.019	.017	-.081	-1.151	.251

Table 33

Grade 2 CFA Model Fit Indices – Construct Validity

	n	CFI	TLI	RMSEA
Fall	1685	0.996	0.997	0.035
Winter	1973	0.995	0.998	0.031
Spring	1821	0.998	0.999	0.019

Table 34

Grade 2 Predictive Validity of the Slope

Measure	Quartile	Fixed effect point estimate of intercept	*SE*	Reliability of intercept	Level-1 residual variance	Random effect variance estimate of intercept	Predictive validity coefficient	*n*
Word Reading Fluency	1	8.777	0.256	0.583	27.886	13.403	0.225	563
	2	27.111	0.271	0.398	37.412	8.442	0.311	544
	3	45.257	0.299	0.141	54.473	3.09	0.112	547
	4	70.946	0.647	0.717	134.316	117.106	-0.119	547
Passage Reading Fluency	1	17.729	0.449	0.632	81.343	47.073	0.597	573
	2	42.527	0.366	0.32	72.823	11.562	0.313	540
	3	66.283	0.398	0.336	85.409	14.779	0.275	546
	4	115.967	1.3	0.883	267.507	693.273	-0.114	541
Multiple Choice Reading Comprehension	1	3.092	0.07	0.179	3.026	0.224	0.681	568
	2	5.986	0.054	0	2.437	0	0.665	696
	3	--*	--*	--*	--*	--*	--*	466
	4	10.439	0.047	0.123	0.997	0.047	0.179	400

*Singular convergence (error).