Technical Report #1003

Technical Adequacy of the easyCBM Primary-Level Reading Measures (Grades K-1), 2009-2010 Version

Cheng-Fei Lai

Joseph F. T. Nese

Elisa M. Jamgochian

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.

Copyright © 2010. Behavioral Research and Teaching. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission.

The University of Oregon is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation. This document is available in alternative formats upon request.

Abstract

In this technical report, we provide the results of a series of studies on the technical adequacy of the early reading measures available on the easyCBM® assessment system. The results from the two-level hierarchical linear growth model analyses suggest that the reliability of the slope estimates for the easyCBM® reading measures are strong, with some exceptions. The growth reliabilities tend to decrease as grade-level increases, and were generally weaker for students who began the year as higher achieving, the latter of which could be explained by a ceiling effect. Results from the regression analyses examining the relation between the spring easyCBM® measures and the SAT-10 provide evidence supporting the concurrent and predictive validity of easyCBM® for grades K and 1. The models were statistically significant for both grades, explaining 73% and 58% of the variance in SAT-10 respectively. Finally, the overall predictive validity coefficients for the slope of improvement for all students in easyCBM® reading measures across grades were positive. Students with lower initial fall scores on the LS and WRF measures across grades (first and second quartile) had the highest rate of growth compared to other quartiles, with moderate predictive validity coefficients in the 0.50s and 0.60s.

Technical Adequacy of the easyCBM® Primary-Level Reading Measures (Grades K-1), 2009-2010 Version

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 Grades K - 1. 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 later grades (Jamgochian et al., 2010; 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 110,007 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

Although the specific sample and analytic technique vary somewhat across the different studies reported on in this technical report, the different studies have a great deal in common. Thus, the information on the measures applies to all the studies reported herein. In the areas which differ by study, we provide study-specific description.

The easyCBM® Early Reading Measures

At each grade level, there are 20 alternate forms of each of the early reading measures on easyCBM®, with 3 designated for seasonal benchmark screenings and the remaining 17 designating for progress monitoring. All easyCBM® forms were scaled to be of equivalent difficulty with a 1PL Rasch model (Alonzo & Tindal, 2007a, 2007b). The reading measures on easyCBM® were written to address "The Big Five" from the National Reading Panel: alphabetic principles, phonics, fluency, vocabulary, and comprehension.

The Kindergarten measures analyzed in our studies include phoneme segmenting, letter sounds, and word reading fluency. First-grade measures include all of the measures available at Kindergarten as well as passage reading fluency. All K-1 easyCBM® measures are designed to be individually-administered by a trained educator. Each of the measures is administered for a period of 60 seconds. Student performance is reported as the total number of correct phonemes, letter sounds, or words read per minute.

Sample

Data for this study were gathered in the 2009-2010 school year from a convenience sample of three school districts in the Pacific Northwest. Sixty-eight schools provided data for the Kindergarten measures, and 71 schools provided data for grade 1. All students in attendance at the schools during the testing windows participated in the study. The Kindergarten phoneme

segmenting sample consisted of 1,984 students, the letter sounds (LS) sample consisted of 1,987 students, and the word reading fluency (WRF) sample consisted of 1,985 students. Approximately 48% of the sample was female. The first grade phoneme segmenting sample consisted of 953 students, the LS sample consisted of 2,181 students, the WRF sample consisted of 2,179 students, and the passage reading fluency (PRF) sample consisted of 2,180 students. Approximately 48% of the sample was female. No other demographic data was available for the Kindergarten and Grade 1 samples.

Analysis Used to Evaluate Reliability of the Slope Measurements

We used a two-level hierarchical linear growth model to represent 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 grade level and ethnicity group. In addition, analyses were separated by quartile based on fall easyCBM® measure, 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 for each student group. The 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, that is

$$\lambda_{0j} = \frac{\tau_{00}}{\tau_{00} + \sigma^2 / n_j},$$

where τ_{00} represents level-2 variance of the growth estimate and σ / n_j represents the measurement error for the level-2 variance (Raudenbush & Bryk, 2002). All analyses were conducted using *R*, the free online statistical software (R Development Core Team, 2010).

Analyses Used to Examine Validity Information for the Performance Level Score

In this section of the report, we describe the analyses we conducted to gather the following types of validity evidence for the easyCBM® reading measures for grades K – 1: concurrent, predictive, and construct. For information regarding the content development and validity of the easyCBM® reading measures, see technical reports on Letter Names, Letter Sounds; and Phoneme Segmenting (Alonzo & Tindal, 2007a); and Word and Passage Reading Fluency (Alonzo & Tindal, 2007b).

Concurrent validity evidence. To examine the concurrent validity of easyCBM® for grades K and 1, we conducted regression analyses using the Spring easyCBM® measures and a published outcome measure (SAT-10). Total scores were used for the easyCBM® measures and standard scores were used for the SAT-10. Pairwise deletion was specified to remove missing variables in the analyses.

Predictive validity evidence. Regression analyses were conducted for each grade level to analyze the predictive validity of the easyCBM® reading measures. Fall, Winter, and Fall and Winter easyCBM® scores and an outcome measure (SAT-10) were used in the analyses. As with the regression analyses performed to provide evidence of concurrent validity, 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.

Construct validity evidence. A confirmatory factor analysis (CFA) was conducted with Mplus version 5.21 (Muthén, & Muthén, 2009) using Maximum Likelihood (ML) iterative estimation procedure to investigate the construct validity of the easyCBM reading measures. ML was used to minimize the discrepancies between the sample variance/covariance matrix and the model implied estimate of the population variance/covariance matrix. Models are presented

below as they pertain to each grade level (Figures 1 and 2). We report a combination of absolute and incremental fit statistics, which we used to evaluate the model fit, including the root-mean-square error of approximation (RMSEA), the standardized root-mean square residual, the Comparative Fit Index (CFI), the Tucker-Lewis Index (TLI). RMSEA values in the vicinity of 0.05 indicate close fit, values near 0.08 suggest fair fit, and values above 0.10 indicate poor model fit (Hu & Bentler, 1999, & Kline, 2005). Hu and Bentler (1999) recommend TLI and CFI values of .95 or better as indicators of a model fit.

Analysis Conducted to Evaluate the Predictive Validity for the Slope of Improvement

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 the level-2 model by student. The easyCBM reading measures, collected in fall 2009, and winter and spring 2010, served as the dependent variables. Student ethnicity along with initial reading achievement grouped into quartiles was modeled at the intercept and slope. Data for each grade level were analyzed separately. The level-2 residuals from the final model were correlated with students' performance scores on the SAT-10.

Results

Only results of ethnic groups with sample sizes of 30 or more are reported here.

Reliability of the Slope Measurements

In this section we report the results of our analysis of the reliability of the slope measurements for the early reading measures by grade and individual measure.

Kindergarten. For the Kindergarten phoneme segmenting measure, the reliability of the growth slope for students in the first quartile was .67, the reliability for students in the second quartile was .66, the reliability for students in the third quartile was .46, and the reliability for

students in the fourth quartile was .09. For the Kindergarten LS measure, the reliability of the growth slope for students in the first quartile was .76, the reliability for students in the second quartile was .76, the reliability for students in the third quartile was .68, and the reliability for students in the fourth quartile was .53. For the Kindergarten WRF measure, the reliability of the growth slope for students in the lower 50th percentile was .82, and the reliability for students in the upper 50th percentile was .71 (see Table 1).

Grade 1. For the grade 1 phoneme segmenting measure, the reliability of the growth slope for students in the first quartile was .28. No other reliability estimates were possible for this measure in grade 1. For the grade 1 LS measure, the reliability of the growth slope for students in the first quartile was .57, the reliability for students in the second quartile was .66, the reliability for students in the third quartile was .72, and the reliability for students in the fourth quartile was .42.

For the grade 1 WRF measure, the reliability of the growth slope for students in the first quartile was .88, the reliability for students in the second quartile was .82, the reliability for students in the third quartile was .76, and the reliability for students in the fourth quartile was .63. For the grade 1 PRF measure, the reliability of the growth slope for students in the first quartile was .93, the reliability for students in the second quartile was.88, the reliability for students in the third quartile was .89, and the reliability for students in the fourth quartile was .60 (see Table 2).

Validity Information for the Performance Level Score

Concurrent validity evidence. The spring model for Grade K was significant, $F(3,179) = 158.80$, $p < .05$, and accounted for 73% of the variance in SAT-10. The spring Word Reading measure had the highest coefficient, $b = 1.98$, $t(189) = 11.49$, $p < .05$, and uniquely explained

20.16% of the variance in the regression model. The descriptive statistics and regression model summaries for Grade K are presented in Tables 5-8.

The spring model for Grade 1 was significant, $F(1,178) = 234.35$, $p < .05$, accounting for 57.8% of the variance in SAT-10. The spring Passage Reading Fluency benchmark had a coefficient, $b = .71$, $t(185) = 15.31$, $p < .05$, and uniquely explained 56.78% of the variance. The descriptive statistics and regression model summaries for Grade 1 are presented in Tables 9-12.

Predictive validity evidence. The fall model for Grade K was significant, $F(3,158) = 69.72$, $p < .05$, and accounted for 57% of the variance in SAT-10. The fall LS benchmark had the highest coefficient, $b = 2.48$, $t(2013) = 6.09$, $p < .05$, and uniquely explained 10.11% of the variance. The winter model for Grade K was significant, $F(3,169) = 112.11$, $p < .05$, and accounted for 66.6% of the variance in SAT-10. The winter WRF benchmark had the highest coefficient, $b = 2.48$, $t(1972) = 8.27$, $p < .05$, and uniquely explained 13.54% of the variance. The fall and winter model for Grade K was significant, $F(6,155) = 86.18$, $p < .05$, and accounted for 76.9% of the variance in SAT-10. The winter WRF benchmark had the highest coefficient, $b = 4.02$, $t(1972) = 8.09$, $p < .05$, and uniquely explained 9.73% of the variance. Descriptive statistics and regression model summaries for Grade K are presented in Tables 13-22.

The fall model for Grade 1 was significant, $F(3,155) = 48.18$, $p < .05$, accounting for 48.3% of the variance in SAT-10. The fall WRF benchmark had the highest coefficient, $b = .82$, $t(2179) = 2.54$, $p < .05$, and uniquely explained 2.16% of the variance. The winter model for Grade 1 was significant, $F(3,173) = 73.64$, $p < .05$, and accounted for 56.1% of the variance in SAT-10. The winter LS benchmark had the highest coefficient, $b = .92$, $t(2195) = 5.16$, $p < .05$, and uniquely explained 6.76% of the variance. The model in which we included both fall and winter results for Grade 1 was significant, $F(6,152) = 33.87$, $p < .05$, and accounted for 57.2% of

the variance in SAT-10. The winter LS benchmark had the highest coefficient, $b = 1.03$, $t(2195) = 4.47$, $p < .05$, and uniquely explained 5.62% of the variance. Descriptive statistics and regression model summaries for Grade 1 are presented in Tables 23-32.

Construct validity evidence. For grade K, the model fit for the hypothesized 1-factor model for all seasons was good, with CFI indices ranging from .997 to .999, TLI indices ranging from .992 to .997 and RMSEA indices ranging from .028 to .047. Descriptive statistics and full CFA model fit indices are presented in Tables 33-34.

For grade 1, CFI and TLI indices suggested good model fit for the hypothesized 1-factor model for all seasons with CFI indices ranging from .978 to .988 and TLI indices ranging from .979 to .982. RMSEA, on the other hand, while indicating a fair model fit for the fall model suggested poor model fit for winter and spring models, with RMSEA indices ranging from .106 to .136. Descriptive statistics and full CFA model fit indices are presented in Tables 35-36.

Predictive Validity Information for the Slope of Improvement

The correlation between slope random effect in Segmenting and Letter Sounds measures with SAT-10 for Kindergarteners were generally positive and moderate for all quartiles. In general, all students in this grade had a moderately strong rate of growth as predicted by WRF and the SAT-10.

No predictive validity coefficients could be obtained from the analysis for any students for the Segmenting measure. A moderate rate of growth was found for almost all students in the LS, WRF, and PRF measures. Some exceptions include the fourth quartiles in WRF and PRF where there were low and moderate negative rate of growth and the first quartile in PRF where the rate of growth was the highest compared to the rest (see Tables 37-38).

Conclusion

The results from the two-level hierarchical linear growth model analyses suggest that the reliability of the slope estimates for the easyCBM® reading measures are strong, with some exceptions. The growth reliabilities tend to decrease as grade-level increases, and were generally weaker for students who began the year as higher achieving, the latter of which could be explained by a ceiling effect. Evaluating students' academic growth has become a critical issue in education for both educators and researchers. Sound research, methods, and measures are needed to assess student growth and explore the practical implications of CBM. This study adds to research on the technical properties of the slope estimates produced by CBM and indicates strong reliability of the growth estimates of the easyCBM® reading assessments for students in grades K-1.

Results from the regression analyses examining the relation between the spring easyCBM® measures and the SAT-10 provide evidence supporting the concurrent validity of easyCBM® for grades K and 1. The models were statistically significant for both grades, explaining 73% and 58% of the variance in SAT-10 respectively. We also have evidence supporting the predictive validity of easyCBM® for grades K and 1 based on the regression analyses conducted using Fall, Winter, and Fall and Winter easyCBM® scores to predict student performance on the SAT-10. The CFA analysis for grade K provided strong evidence for the 4-factor model. Mixed results, however, were obtained from the CFA analysis for grade 1 5-factor model.

Finally, the overall predictive validity coefficients for the slope of improvement for all students in easyCBM® reading measures across grades were positive. Students with lower initial fall scores on the LS and WRF measures across grades (first and second 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 grade 1 PRF measures, similar rate of growth was observed, where students in the first quartile typically had a higher rate of growth compared to other quartiles.

References

Alonzo, J., & Tindal, G. (2007a). *Examining the Technical Adequacy of Early Literacy Measures in a Progress Monitoring Assessment System: Letter Names, Letter Sounds, and Phoneme Segmenting (Technical Report # 39)*. Eugene, OR: Behavioral Research and Teaching.

Alonzo, J., & Tindal, G. (2007b). *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.

Alonzo, J., Tindal, G., & Ketterlin-Geller, L.R. (2006). General outcome measures of basic skills in reading and math. In L. Florian (Ed.), Handbook of Special Education. Thousand Oaks, CA: Sage.

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.

Deno, S. L. (2003). Developments in curriculum-based measurements. *The Journal of Special Education, 37*, 184-192.

Deno, S. (1987). Curriculum-based measurement. *Teaching Exceptional Children.* (Fall), 41-47.

Deno, S. L., & Mirkin, P. M. (1977). *Data based program modification.* Minneapolis, MN: University of Minnesota Leadership Training Institute/Special Education.

Good, R. H., Gruba, J., & Kaminski, R. A. (2002). Best practices in Using Dynamic Indicators of Basic Early Literacy Skills (DIBELS) in an Outcomes-Driven Model. In A. Thomas

and J. Grimes (Eds.). *Best Practices in School Psychology IV* (pp.679-700). Washington, DC: National Association of School Psychologists.

Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modeling*, *6*(1), 1-55.

Kline, R. B. (2005). Principles and practice of structural equation modeling (2nd ed.). New York: Guilford.

Muthén, L. K., & Muthén, B. O. (2009). Mplus version 5.21 [computer software]. Los Angeles: Authors.

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.

R Development Core Team. (2010). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Raudenbush, S., & Bryk, A. (2002). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage.

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.

Table 1
Kindergarten, Reliability of Phoneme Segmenting, Letter Sounds, and Word Reading Fluency Growth Slopes

Measure	Student Group	Fixed effect, Intercept	*SE*	Level-1 residual variance	Reliability, Intercept	Fixed, effect, slope	*SE*	Variance, slope	Reliability, Slope	*n*
Pho-neme Segmen-ting	Quartile 1	0.61	0.31	74.96	0.00	19.09	0.38	51.16	0.67	617
	Quartile 2	4.29	0.40	66.47	0.00	18.42	0.48	43.91	0.66	334
	Quartile 3	13.54	0.34	63.36	0.14	15.92	0.32	17.78	0.46	481
	Quartile 4	32.84	0.44	81.59	0.49	7.90	0.30	2.76	0.09	484
Letter Sounds	Quartile 1	-0.74	0.18	28.74	0.00	13.73	0.26	34.96	0.78	703
	Quartile 2	0.90	0.27	26.19	0.00	14.75	0.37	27.08	0.76	293
	Quartile 3	5.34	0.23	30.68	0.00	14.23	0.28	21.61	0.68	464
	Quartile 4	19.75	0.38	38.05	0.73	11.35	0.27	14.17	0.53	460
Word Reading Fluency	Lower 50%ile	-0.67	0.07	5.37	0.00	4.15	0.11	8.13	0.82	836
	Upper 50%ile	3.41	0.37	38.35	0.86	7.96	0.25	31.27	0.71	791

Table 2
Grade 1, Reliability of Phoneme Segmenting, Letter Sounds, Word Reading Fluency, and Passage Reading Fluency Growth Slopes

Measure	Student Group	Fixed effect, Intercept	*SE*	Level-1 residual variance	Reliability, Intercept	Fixed, effect, slope	*SE*	Variance, slope	Reliability, Slope	*n*
Phon-eme Segmen-ting	Quartile 1	16.96	0.47	107.05	0.30	16.41	0.49	14.62	0.28	277
	Quartile 2	-	-	-	-	-	-	-	-	-
	Quartile 3	-	-	-	-	-	-	-	-	-
	Quartile 4	-	-	-	-	-	-	-	-	-
Letter Sounds	Quartile 1	11.75	0.30	50.91	0.33	15.72	0.30	23.03	0.57	566
	Quartile 2	24.14	0.25	41.92	0.00	11.88	0.31	28.31	0.66	542
	Quartile 3	31.90	0.25	40.24	0.00	9.02	0.33	36.27	0.72	543
	Quartile 4	43.63	0.36	64.66	0.31	5.92	0.33	16.52	0.42	478
Word Reading Fluency	Quartile 1	0.25	0.16	20.07	0.00	12.24	0.31	48.73	0.88	626
	Quartile 2	3.50	0.23	30.05	0.00	16.11	0.38	48.72	0.82	460
	Quartile 3	9.25	0.29	50.98	0.00	20.00	0.41	56.76	0.76	510
	Quartile 4	45.34	0.89	72.19	0.94	15.96	0.39	43.00	0.63	531
Passage Reading Fluency	Quartile 1	-0.13	0.35	25.80	0.00	14.71	0.86	114.97	0.93	171
	Quartile 2	3.36	0.49	44.25	0.00	21.46	0.96	113.07	0.88	145
	Quartile 3	9.53	0.67	72.76	0.00	26.31	1.36	196.78	0.89	125
	Quartile 4	53.44	2.94	169.33	0.96	33.13	1.08	84.72	0.60	143

Note. Parameters could not be estimated for the phoneme segmenting quartiles 2, 3, and 4.

Table 3

Kindergarten Descriptive Statistics by Subtest

	N	Minimum	Maximum	*M*	*SD*
Fall 2009					
Segmenting	2010	0	69	11.63	13.846
Letter Names	1744	0	88	19.54	14.596
Letter Sounds	2013	0	58	6.17	8.783
Word Reading	1719	0	107	2.33	7.566
Winter 2010					
Segmenting	1970	0	69	29.54	16.580
Letter Names	898	0	86	32.30	14.393
Letter Sounds	1972	0	90	18.38	12.632
Word Reading	1972	0	98	4.78	8.880
Spring 2010					
Segmenting	2083	0	70	41.85	14.299
Letter Names	863	0	100	43.46	14.595
Letter Sounds	2083	0	98	32.58	13.195
Word Reading	2083	0	115	13.67	14.441

Table 4

Grade 1 Descriptive Statistics by Subtest

	N	Minimum	Maximum	*M*	*SD*
Fall 2009					
Segmenting	1958	0	70	34.55	14.625
Letter Names	1959	0	100	40.41	15.535
Letter Sounds	2182	0	75	25.62	12.331
Word Reading	2179	0	101	15.98	20.409
Passage Reading	632	0	181	16.69	26.765
Winter 2010					
Segmenting	950	0	79	49.63	13.729
Letter Names	878	0	104	55.20	17.767
Letter Sounds	2195	0	95	41.39	13.599
Word Reading	2194	0	112	26.65	22.898
Passage Reading	2186	0	253	35.33	39.351
Spring 2010					
Segmenting	869	3	80	51.28	12.234
Letter Names	895	6	131	66.27	18.470
Letter Sounds	2072	0	96	46.12	13.330
Word Reading	2074	0	120	47.25	25.756
Passage Reading	2268	0	246	62.51	44.114

Table 5

Kindergarten Descriptive Statistics for Concurrent Validity

	M	*SD*	*N*
easyCBM			
Segmenting	41.85	14.299	2083
Letter Names	43.46	14.595	863
Letter Sounds	32.58	13.195	2083
Word Reading	13.67	14.441	2083
SAT10 Reading Total Standard Score	491.80	50.674	189

Table 6

Kindergarten Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.853[a]	.727	.722	26.704	.727	158.798	3	179	.000

a. Predictors: (Constant), Spr10WRF, Spr10Seg, Spr10LS

Table 7

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	339713.494	3	113237.831	158.798	.000[a]
	Residual	127643.493	179	713.092		
	Total	467356.987	182			

a. Predictors: (Constant), Spr10WRF, Spr10Seg, Spr10LS

b. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 8

Kindergarten Model Coefficients[a]

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	415.556	6.679		62.223	.000	402.377	428.735			
	Spr10 Seg	.062	.162	.017	.382	.703	-.257	.380	.396	.029	.015
	Spr10 LS	1.429	.208	.372	6.875	.000	1.019	1.840	.724	.457	.269
	Spr10 WRF	1.983	.173	.565	11.485	.000	1.642	2.323	.797	.651	.449

a. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 9

Grade 1 Descriptive Statistics

	M	*SD*	*N*
easyCBM			
Segmenting	51.28	12.234	869
Letter Names	66.27	18.470	895
Letter Sounds	46.12	13.330	2072
Word Reading	47.25	25.756	2074
Passage Reading Fluency	62.51	44.114	2268
SAT10 Reading Total Standard Score	561.36	41.287	185

Table 10

Grade 1 Model Summary

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.754[a]	.568	.566	27.202	.568	234.347	1	178	.000

a. Predictors: (Constant), Spr10PRF

Table 11

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	173407.974	1	173407.974	234.347	.000[a]
	Residual	131713.533	178	739.964		
	Total	305121.507	179			

a. Predictors: (Constant), Spr10PRF

b. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 12

Grade 1 Model Coefficients[a]

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)	517.250	3.523		146.817	.000	510.298	524.203			
	Spr10PRF	.706	.046	.754	15.308	.000	.615	.796	.754	.754	.754

a. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 13

Kindergarten Descriptive Statistics

	Mean	Std. Deviation	N
SAT10 Reading Total Standard Score - Kindergarten	491.80	50.674	189
Fall09Seg	11.63	13.846	2010
Fall09LN	19.54	14.596	1744
Fall09LS	6.17	8.783	2013
Fall09WRF	2.33	7.566	1719
Wint10Seg	29.54	16.580	1970
Wint10LN	32.30	14.393	898
Wint10LS	18.38	12.632	1972
Wint10WRF	4.78	8.880	1972

Table 14

Kindergarten Model Summary (Fall easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.755[a]	.570	.561	33.557	.570	69.715	3	158	.000

a. Predictors: (Constant), Fall09WRF, Fall09Seg, Fall09LS

Table 15

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	235512.168	3	78504.056	69.715	.000[a]
	Residual	177919.013	158	1126.070		
	Total	413431.180	161			

a. Predictors: (Constant), Fall09WRF, Fall09Seg, Fall09LS

b. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 16

Kindergarten Model Coefficients[a] *(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)	459.327	3.562		128.935	.000	452.291	466.364			
	Fall09Seg	1.330	.223	.363	5.956	.000	.889	1.771	.620	.428	.311
	Fall09LS	2.479	.407	.430	6.092	.000	1.675	3.283	.681	.436	.318
	Fall09WRF	.738	.426	.110	1.730	.086	-.104	1.579	.471	.136	.090

a. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 17

Kindergarten Model Summary (Winter easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.816[a]	.666	.660	29.564	.666	112.114	3	169	.000

a. Predictors: (Constant), Wint10WRF, Wint10Seg, Wint10LS

Table 18

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	293969.241	3	97989.747	112.114	.000[a]
	Residual	147708.790	169	874.017		
	Total	441678.031	172			

a. Predictors: (Constant), Wint10WRF, Wint10Seg, Wint10LS

b. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 19

Kindergarten Model Coefficients[a] (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)	429.688	4.929		87.168	.000	419.957	439.419			
	Wint10Seg	1.128	.156	.369	7.248	.000	.820	1.435	.618	.487	.322
	Wint10LS	.923	.227	.230	4.059	.000	.474	1.372	.635	.298	.181
	Wint10WRF	2.475	.300	.434	8.265	.000	1.884	3.067	.672	.536	.368

a. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 20

Kindergarten Model Summary (Fall + Winter easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.877[a]	.769	.760	24.802	.769	86.178	6	155	.000

a. Predictors: (Constant), Wint10WRF, Wint10Seg, Fall09Seg, Wint10LS, Fall09LS, Fall09WRF

Table 21

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	318081.289	6	53013.548	86.178	.000[a]
	Residual	95349.891	155	615.161		
	Total	413431.180	161			

a. Predictors: (Constant), Wint10WRF, Wint10Seg, Fall09Seg, Wint10LS, Fall09LS, Fall09WRF

b. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 22

Kindergarten Model Coefficients[a] (Fall + 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)	433.864	4.331		100.178	.000	425.308	442.419			
	Fall09Seg	.870	.179	.238	4.860	.000	.516	1.223	.620	.364	.187
	Fall09LS	1.298	.350	.225	3.710	.000	.607	1.989	.681	.286	.143
	Fall09WRF	-2.854	.548	-.426	-5.209	.000	-3.937	-1.772	.471	-.386	-.201
	Wint10Seg	.731	.147	.239	4.969	.000	.440	1.022	.618	.371	.192
	Wint10LS	.308	.224	.077	1.377	.171	-.134	.751	.635	.110	.053
	Wint10WRF	4.015	.497	.704	8.081	.000	3.033	4.996	.672	.544	.312

a. Dependent Variable: SAT10 Reading Total Standard Score - Kindergarten

Table 23

Grade 1 Descriptive Statistics (Fall easyCBM and SAT-10)

	Mean	Std. Deviation	N
SAT10 Reading Total Standard Score - 1st Grade	561.36	41.287	185
Fall09Seg	34.55	14.625	1958
Fall09LN	40.41	15.535	1959
Fall09LS	25.62	12.331	2182
Fall09WRF	15.98	20.409	2179
Fall09PRF	16.69	26.765	632
Wint10Seg	49.63	13.729	950
Wint10LN	55.20	17.767	878
Wint10LS	41.39	13.599	2195
Wint10WRF	26.65	22.898	2194
Wint10PRF	35.33	39.351	2186

Table 24

Grade 1 Model Summary (Fall easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.695[a]	.483	.473	29.985	.483	48.184	3	155	.000

a. Predictors: (Constant), Fall09PRF, Fall09LS, Fall09WRF

Table 25

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	129965.031	3	43321.677	48.184	.000[a]
	Residual	139360.098	155	899.097		
	Total	269325.129	158			

a. Predictors: (Constant), Fall09PRF, Fall09LS, Fall09WRF

b. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 26

Grade 1 Model Coefficients[a] (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)	524.995	5.638		93.114	.000	513.857	536.132			
	Fall09LS	.765	.240	.228	3.186	.002	.291	1.239	.535	.248	.184
	Fall09WRF	.822	.324	.406	2.540	.012	.183	1.461	.669	.200	.147
	Fall09PRF	.218	.231	.141	.942	.348	-.238	.673	.628	.075	.054

a. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 27

Grade 1 Model Summary (Winter easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.749[a]	.561	.553	27.597	.561	73.641	3	173	.000

a. Predictors: (Constant), Wint10PRF, Wint10LS, Wint10WRF

Table 28

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	168253.150	3	56084.383	73.641	.000[a]
	Residual	131754.589	173	761.587		
	Total	300007.739	176			

a. Predictors: (Constant), Wint10PRF, Wint10LS, Wint10WRF

b. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 29

Grade 1 Model Coefficients[a] (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)	500.881	6.734		74.382	.000	487.590	514.172			
	Wint10LS	.916	.178	.302	5.158	.000	.566	1.267	.557	.365	.260
	Wint10WRF	.243	.277	.135	.878	.381	-.303	.789	.692	.067	.044
	Wint10PRF	.455	.155	.434	2.938	.004	.149	.761	.691	.218	.148

a. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 30

Grade 1 Model Summary (Fall + Winter easyCBM and SAT-10)

Model					Change Statistics				
	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.756[a]	.572	.555	27.536	.572	33.865	6	152	.000

a. Predictors: (Constant), Wint10PRF, Wint10LS, Fall09LS, Fall09PRF, Wint10WRF, Fall09WRF

Table 31

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	154070.218	6	25678.370	33.865	.000[a]
	Residual	115254.911	152	758.256		
	Total	269325.129	158			

a. Predictors: (Constant), Wint10PRF, Wint10LS, Fall09LS, Fall09PRF, Wint10WRF, Fall09WRF

Table 32

Grade 1 Model Coefficients[a] (Fall + 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)	499.861	7.158		69.830	.000	485.718	514.004			
	Fall09LS	-.033	.268	-.010	-.122	.903	-.562	.496	.535	-.010	-.006
	Fall09WRF	.399	.464	.197	.858	.392	-.519	1.316	.669	.069	.046
	Fall09PRF	.189	.260	.123	.727	.468	-.325	.703	.628	.059	.039
	Wint10LS	1.027	.230	.338	4.467	.000	.573	1.481	.557	.341	.237
	Wint10WRF	.104	.385	.057	.269	.788	-.657	.864	.692	.022	.014
	Wint10PRF	.213	.217	.203	.980	.329	-.217	.643	.691	.079	.052

a. Dependent Variable: SAT10 Reading Total Standard Score - 1st Grade

Table 33

Kindergarten Descriptive Statistics for CFA

	N	Minimum	Maximum	Mean	Std. Deviation
Fall09Seg	2010	0	69	11.63	13.846
Fall09LN	1744	0	88	19.54	14.596
Fall09LS	2013	0	58	6.17	8.783
Fall09WRF	1719	0	107	2.33	7.566
Wint10Seg	1970	0	69	29.54	16.580
Wint10LN	898	0	86	32.30	14.393
Wint10LS	1972	0	90	18.38	12.632
Wint10WRF	1972	0	98	4.78	8.880
Spr10Seg	2083	0	70	41.85	14.299
Spr10LN	863	0	100	43.46	14.595
Spr10LS	2083	0	98	32.58	13.195
Spr10WRF	2083	0	115	13.67	14.441
Valid N (listwise)	452				

Table 34

Kindergarten CFA Model Fit Indices

	n	CFI	TLI	RMSEA
Fall	1449	0.999	0.997	0.030
Winter	897	0.997	0.992	0.047
Spring	862	0.999	0.997	0.028

Table 35

Grade 1 Descriptive Statistics for CFA

	N	Minimum	Maximum	Mean	Std. Deviation
Fall09Seg	1958	0	70	34.55	14.625
Fall09LN	1959	0	100	40.41	15.535
Fall09LS	2182	0	75	25.62	12.331
Fall09WRF	2179	0	101	15.98	20.409
Fall09PRF	632	0	181	16.69	26.765
Wint10Seg	950	0	79	49.63	13.729
Wint10LN	878	0	104	55.20	17.767
Wint10LS	2195	0	95	41.39	13.599
Wint10WRF	2194	0	112	26.65	22.898
Wint10PRF	2186	0	253	35.33	39.351
Spr10Seg	869	3	80	51.28	12.234
Spr10LN	895	6	131	66.27	18.470
Spr10LS	2072	0	96	46.12	13.330
Spr10WRF	2074	0	120	47.25	25.756
Spr10PRF	2268	0	246	62.51	44.114
Valid N (listwise)	400				

Table 36

Grade 1 CFA Model Fit Indices

	n	CFI	TLI	RMSEA
Fall	412	0.993	0.982	0.084
Winter	876	0.988	0.970	0.106
Spring	865	0.978	0.944	0.136

Table 37
Predictive Validity Results for All Students in Grade K

Measure	Quartile/ Percentile	Fixed effect point estimate of intercept	SE	Reliability of intercept	Level-1 residual variance	Random effect variance estimate of intercept	Predictive validity coefficient	N
Segmenting	1	0.613	0.31	0	74.962	0	0.512	658
	2	4.29	0.399	0	66.466	0	0.512	352
	3	13.542	0.336	0.136	63.355	3.344	0.504	500
	4	32.84	0.436	0.491	81.585	26.373	0.414	500
Letter Sounds	1	-0.735	0.179	0	28.737	0	0.691	759
	2	0.902	0.268	0	26.192	0	0.609	306
	3	5.34	0.233	0	30.683	0	0.487	472
	4	19.747	0.376	0.733	38.054	35.197	0.558	476
Word Reading Fluency	Lower 50%ile	-0.671	0.071	0	5.366	0	0.796	904
	Upper 50%ile	3.409	0.365	0.856	38.345	76.666	0.818	815

Table 38
Predictive Validity Results for All Students in Grade 1

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
Seg-meting	1	16.955	0.469	0.299	107.047	15.974	NA	515
	2	--*	--*	--*	--*	--*	--*	--*
	3	--*	--*	--*	--*	--*	--*	--*
	4	--*	--*	--*	--*	--*	--*	--*
Letter Sounds	1	11.751	0.299	0.331	50.905	8.569	0.553	573
	2	24.14	0.253	0	41.921	0	0.481	551
	3	31.899	0.246	0	40.239	0	0.380	564
	4	43.631	0.362	0.31	64.656	10.141	0.405	494
Word Reading Fluency	1	0.246	0.163	0	20.071	0	0.645	636
	2	3.5	0.23	0	30.05	0	0.510	482
	3	9.248	0.288	0	50.983	0	0.385	519
	4	45.343	0.89	0.936	72.194	368.843	-0.51	542
Passage Reading Fluency	1	-0.131	0.347	0	25.795	0	0.821	181
	2	3.358	0.485	0	44.25	0	0.519	159
	3	9.526	0.671	0	72.761	0	0.519	137
	4	53.44	2.943	0.955	169.327	1199.056	-0.364	155

*Singular convergence (error).

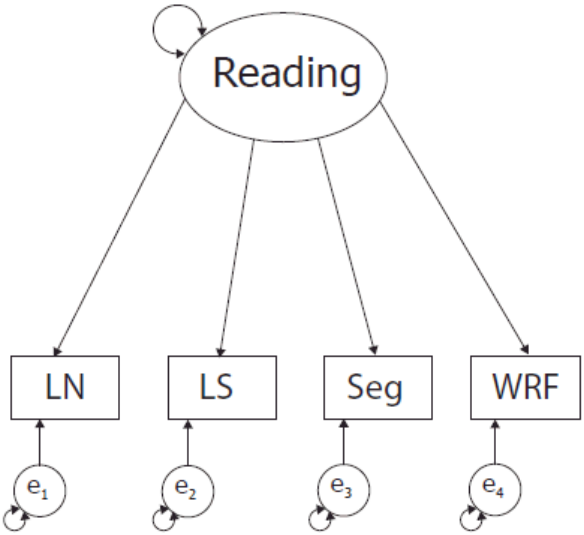

Figure 1. Hypothesized 1-Factor Model for easyCBM Reading Kindergarten

Note. LN = Letter Naming; LS = Letter Sounds; Seg = Phoneme Segmenting; WRF = Word Reading Fluency

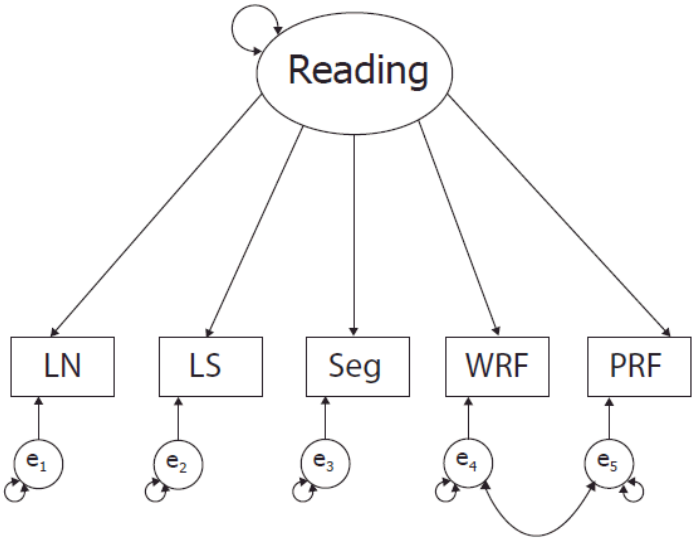

Figure 2. Hypothesized 1-Factor Model for easyCBM Reading Grade 1

Note. LN = Letter Naming; LS = Letter Sounds; Seg = Phoneme Segmenting; WRF = Word Reading Fluency; PRF = Passage Reading Fluency