

Technical Report 2401

**Criterion Validity and Classification Accuracy of easyCBM:
Grades 3-8**

Gerald Tindal, PhD
Joseph F. T. Nese, PhD

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: This technical report was supported in part by Riverside Insight, the exclusive distributor of easyCBM. The report does not reflect any endorsement by any of these organizations.

Copyright © 2024. Behavioral Research and Teaching. All rights reserved. This publication or parts thereof, may not be used or reproduced in any manner without written permission.

APA Reference: Tindal, G. & Nese, J. F. T., & (2024). *Criterion validity and classification accuracy of easyCBM: Grades 3-8*. (Technical Report # 2401). Eugene, OR: Behavioral Research and Teaching, University of Oregon.

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 alternate formats upon request.

Abstract

We present two types of validity evidence to support inferences and decisions about use of easyCBMs in relation to state testing programs. The first type involves the use of Benchmarks in reading to use in making predictions of performance on the Smarter Balanced (SB) test. These predictions can be made both well in advance (several months) or nearly concurrent with administration of the SB test. The second type of validity evidence supports the use of Benchmarks in screening students for risk of problems in learning to read. With data from two states and four school districts, we analyzed well over 8,000 anonymized records, with students reflecting a broad range of demographics. In the **first section** of the report, we present both the grade level demographics of this sample and their performances on the following easyCBMs in the Fall, Winter, and Spring: Proficient Reading, Vocabulary, Oral (Passage) Reading Fluency, and Math. Criterion validity data (both predictive and concurrent) are first presented for grade level easyCBM raw scores for each measure correlated with the SB test. However, because different CBMs were used within and across grades, we also computed a composite that reflected a standard score (with z-scores computed for each grade level in reading and math). This score transformation allowed both correlations and regression analyses to be presented between easyCBMs and SB using this composite. In the **second section** of the paper, we present the results of a classification analysis and report a range of statistics: cut points for the 20th percentile rank (PR), classification data (true and false positives and negatives, area under the curve (with 95% confidence intervals for the lower and upper bounds), base rates, overall classification rates, sensitivity and specificity, false positive and negative rates, and finally positive and negative predictive power. The results from these classification analyses support the use easyCBMs as screeners with high AUCs, high values of sensitivity and specificity, and across the board high values of positive predictive power.

Introduction

This report complements previous BRT technical reports on the technical adequacy of easyCBMs in Grades 3 through 8 (Anderson et al., 2014), where various forms of reliability and validity information were summarized. In this report, however, we focus only on validity. According to the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), validity addresses the trustworthiness of the test outcomes and the degree to which inferences and decisions can be supported from the test outcomes. Validity is not of the test, but of the inferences and decisions made from the test.

Several types of validity are necessary, depending on the decisions, including content (item coverage and alignment with important content being measured), construct (the substance of the measure), criterion (relations of the measure with other, independent measures), and response processes (the degree to which the measure is accessible to the measure). In this research report, we focus on criterion validity (both predictive and concurrent), using the SB test, which is typically given at the end of the year to students in Grades 3 to 8.

We have divided the report into two major sections: (a) descriptive statistics on students' demographics and performance on all measures, as well as the relations among easyCBM reading and math measures and their composite scores with the SB test, and (b) the classification accuracy of easyCBM in recommending intensive interventions.

Note: The demographics for the descriptive section of this report do not exactly match the demographics for the classification accuracy because of pair-wise deletion. In the descriptive report, any student with a score was counted in both the demographics and in the descriptives statistics. In the criterion validity and classification accuracy parts of the report, students had to have scores for both individual easyCBMs and the SB test to be included in the analysis.

For each of these measures, we report on fall, winter, and spring descriptive statistics, including the count (n), minimum score (min), maximum score (max), mean (average performance), standard deviation (amount of variation around the mean), and kurtosis (degree to which the performance distribution is normal). We also report the same descriptive statistics for the SB test in English Language Arts and Mathematics.

Finally, we extend our analysis from grade level reading and math measures to a composite score in each subject, so that we can document predictive and concurrent validity relations using graphic displays to regress easyCBM performance on state tests. The reason for these regression analyses using only composites is to ensure comparability across measures and grades.

In the second section, we focus on classification accuracy, or the degree to which Benchmark (BM) easyCBMs can be used to classify students at risk of learning problems

and therefore in need of intensive services. Specifically, we investigate several components of this process as required by NCII.

1. Area under the curve (AUC), presents a comparison of two ratios. First, true positive rates are calculated: Number of students who were correctly identified at risk divided by this group + false negatives (students incorrectly identified as NOT at risk). Second, false positive rates are calculated: Number of students falsely identified as 'at risk divided by this group + true negatives (students correctly identified as NOT at risk).

2. If the AUC is perfect (1.0), high values occur with *sensitivity* (the ability of the measure to correctly identify students at risk who indeed are at risk) and *specificity* (the ability of the measure to correctly identify students NOT at risk who indeed are NOT at risk). Both values are important to ensure that students who need intensive intervention can receive it (sensitivity) and that students who do not need intensive services do not receive it (specificity). Following is a table displaying the relation between true and false positives and negatives, underlying calculations of ratios.

Table 1.

Sensitivity and Specificity in Decision Making: True and False Positives and Negatives

	Predicted Positive	Predicted Negative
Actual Positive	True Positive (TP)	False Negative (FN)
Actual Negative	False Positive (FP)	True Negative (TN)

$$\text{Precision} = \text{TP}/(\text{TP}+\text{FP})$$

These two dimensions often reflect tradeoffs, which are often differently valued, for example, in professions other than education. Medicine may emphasize sensitivity to ensure individuals who potentially have a condition are treated, even though some individuals may receive a treatment who did not need it. In contrast, in the criminal justice system, high specificity may be emphasized to ensure no individuals are convicted of a crime they did not commit. Personnel in educational systems may value true and false negatives/positives in varying ways. Some schools may have more tolerance for false negatives because the students identified at risk are burgeoning in Tier III, and an RTI system is best configured by providing services in Tier II. In contrast, a 'no child left behind' perspective may prevail in which case, personnel may have more tolerance for false positives to ensure maximum services regardless of risk levels.

To operationalize this research in practice, the findings from this investigation complement and update previous research on classification accuracy that have been conducted with easyCBMs. In this research, the following findings were reported from three earlier studies using easyCBM.

An early study of Grade 1 students by Clark et al. (2011) reported on the diagnostic efficiency using logistic regression of the Fall and Winter easyCBM Math Benchmark scores on the TerraNova Math at the 25th and 40th Percentiles. Their findings showed high sensitivity and specificity with "the AUC for the ROC of each season in the good range, increasing from .83 in fall, to .85 in winter, and .89 in spring" (p. 250).

Alonzo and Anderson (2018) reported on early reading easyCBMs for students in Grade K through Grade 2 in the Fall, Winter, and Spring using various grade level measures of letter names, letter sounds, phoneme segmenting, word reading fluency, passage reading fluency, vocabulary, and multiple-choice comprehension. They reported that “with the exception of the Phoneme Segmenting measure when administered in the Fall of Kindergarten, all measures demonstrated strong classification accuracy” (abstract).

Finally, Alonzo and Anderson (2018) conducted a follow up study of an earlier easyCBM investigation and administered alternate forms of four easyCBM measures (Letter Names (LN), Letter Sounds (LS), Phoneme Segmenting (PS), and Word Reading Fluency (WRF)) to all kindergarten and first-grade students in the sample” (pages 1-2), along with the SAT-10. They reported varied results by measure and grade, but most findings were generally within good or acceptable ranges for all measures and grade levels analyzed.

In summary, in the early grades, the research established easyCBM to be a reasonable measure in screening students for risk of learning to read. This study extends this research beyond the early grades and focuses on grade levels participating in state testing programs (Grade 3 – Grade 8).

Demographics of the Total Sample

Table 2.

Frequencies for Grade by Special Education Status

Grade	No	Yes
3	1178	326
4	1171	259
5	1221	255
6	1052	178
7	1078	189
8	1105	181
NA	214	87

Table 3.

Frequencies for Grade by Gender

Grade	Female	Male	NA
3	713	790	1
4	689	741	NA
5	714	759	3
6	578	651	1
7	598	665	4
8	599	676	11
NA	144	156	1

Table 4.
Frequencies for Grade by Race

Grade	AABlack	AIAN	Asian	Hisp	NHPI	TwoM	White
3	114	23	101	23	41	255	947
4	119	34	81	35	32	270	859
5	139	18	83	19	49	253	915
6	116	11	87	NA	37	236	743
7	125	21	96	NA	32	238	755
8	110	20	102	NA	50	256	748
NA	28	9	10	NA	12	54	188

Key: AABlack = African American-Black / AIAN = American Indian-Alaska Native / Asian = Asian / Hisp = Hispanic-Latino / NHPI = Native Hawaiian-Pacific Islander / TwoM = Two or more races / White = White

Table 5.
Frequencies for Grade by English Learner Status

Grade	No	Yes
3	1386	118
4	1338	92
5	1397	79
6	1160	70
7	1203	64
8	1239	47
NA	283	18

Table 6.
Frequencies for Grade by Ethnicity

Grade	Hispanic/Latino	Not Hispanic/Latino
3	311	1193
4	313	1117
5	324	1152
6	273	957
7	273	994
8	283	1003
NA	61	240

Descriptive Statistics of the Total Sample

Overview

All research conducted in schools needs to report on the basic descriptive statistics so that reviewers can analyze the values to determine their accuracy and possible comparability to both other research as well as values in school settings. In this section, we present tables with the results summarized by grade level with the following key for columns: the number of students (N), the minimum score (Min), the maximum score (Max), the average (Mean), standard deviation (SD), and distribution (Kurtosis). As noted earlier, the student samples in this section of the report had to have taken the only measure in the table, with no requirement that they had taken other measures (either easyCBMs or the SB test).

CBM Proficient Reading

The table below presents, for each grade (3-8), descriptive statistics for Fall (F), Winter (W), and Spring (S) Benchmark measures. Note: CBMF_readP=Fall Proficient Reading, CBMW_readP=Winter Proficient Reading, CBMS_readP=Spring Proficient Reading.

Table 7.

easyCBM Proficient Reading Statistics (by Season and Grade)

Measure-Season	Grade	N	Min	Max	Mean	SD	Kurtosis
CBMF_readP	3	1306	0	20	10.79	3.87	-0.74
CBMF_readP	4	1231	0	20	12.33	4.29	-0.79
CBMF_readP	5	1296	0	20	13.60	3.54	0.63
CBMF_readP	6	988	1	20	13.52	3.65	0.96
CBMF_readP	7	1003	1	20	12.63	3.77	-0.29
CBMF_readP	8	1058	0	20	12.81	3.80	0.21
CBMW_readP	3	1235	1	18	10.47	2.98	-0.51
CBMW_readP	4	1165	2	20	13.61	3.62	-0.37
CBMW_readP	5	1346	0	20	15.18	3.92	1.41
CBMW_readP	6	1007	1	20	12.67	3.55	0.20
CBMW_readP	7	1027	0	20	13.88	3.49	1.10
CBMW_readP	8	1062	0	20	12.06	3.41	0.38
CBMS_readP	3	1389	0	20	12.94	4.20	-0.59
CBMS_readP	4	1298	0	20	13.59	3.83	-0.08
CBMS_readP	5	1381	0	20	13.82	3.69	1.37
CBMS_readP	6	917	1	20	13.16	3.79	-0.01
CBMS_readP	7	996	1	18	10.97	3.26	-0.13
CBMS_readP	8	1000	1	19	11.78	3.50	-0.09

easyCBM Vocabulary

The table below presents, for each grade (3-8), descriptive statistics for Fall (F), Winter (W), and Spring (S) Benchmark measures. Note: CBMF_readVoc=Fall Vocabulary, CBMW_readVoc=Winter Vocabulary, CBMS_readVoc=Spring Vocabulary.

Table 8.
CBM Vocabulary Statistics (by Season and Grade)

Measure-Season	Grade	N	Min	Max	Mean	SD	Kurtosis
CBMF_readVoc	3	1318	0	20	14.31	4.60	0.01
CBMF_readVoc	4	1230	0	20	14.91	5.23	1.34
CBMF_readVoc	5	1298	0	20	16.02	3.59	4.35
CBMF_readVoc	6	975	1	20	17.01	2.87	5.47
CBMF_readVoc	7	1004	3	20	17.08	2.97	3.21
CBMF_readVoc	8	1060	1	20	17.16	2.98	4.03
CBMW_readVoc	3	1238	2	20	15.84	3.83	0.99
CBMW_readVoc	4	1164	2	20	16.77	3.25	2.30
CBMW_readVoc	5	1329	0	20	16.30	3.46	3.84
CBMW_readVoc	6	1004	1	20	16.70	3.30	3.66
CBMW_readVoc	7	1034	6	20	16.90	2.71	2.29
CBMW_readVoc	8	1075	2	20	17.05	2.75	4.59
CBMS_readVoc	3	1394	0	20	16.18	3.93	3.13
CBMS_readVoc	4	1286	0	20	16.16	4.99	3.66
CBMS_readVoc	5	1361	0	20	15.99	4.14	4.91
CBMS_readVoc	6	920	1	20	17.36	2.79	5.11
CBMS_readVoc	7	1010	3	20	16.67	2.97	2.06
CBMS_readVoc	8	1013	1	20	17.57	3.03	5.18

easyCBM Oral Passage Reading Fluency

The table below presents, for each grade (3-8), descriptive statistics for Fall (F), Winter (W), and Spring (S) Benchmark measures. Note: CBMF_readORF=Fall Oral Passage Reading Fluency, CBMW_readORF=Winter Oral Passage Reading Fluency, CBMS_readORF=Spring Oral Passage Reading Fluency.

Table 9.
CBM Oral Passage Reading Statistics (by Season and Grade)

Measure-Season	Grade	N	Min	Max	Mean	SD	Kurtosis
CBMF_readORF	3	1393	1	218	81.10	43.17	-0.43
CBMF_readORF	4	1309	0	246	107.71	42.24	0.18
CBMF_readORF	5	1369	1	262	143.33	49.76	0.42
CBMF_readORF	6	671	30	271	150.23	37.09	0.37
CBMF_readORF	7	760	18	304	156.42	35.11	0.85
CBMF_readORF	8	728	4	295	170.38	39.01	1.01
CBMW_readORF	3	1243	1	260	115.17	47.35	-0.20
CBMW_readORF	4	1166	3	255	132.40	41.94	0.96
CBMW_readORF	5	1419	1	281	149.03	51.57	0.44

Continued on next page.

Measure-Season	Grade	N	Min	Max	Mean	SD	Kurtosis
CBMW_readORF	6	823	9	269	163.26	36.64	0.46
CBMW_readORF	7	806	37	279	172.76	39.27	-0.09
CBMW_readORF	8	803	24	266	162.17	39.86	0.15
CBMS_readORF	3	1470	0	254	105.43	49.43	-0.32
CBMS_readORF	4	1386	1	256	131.69	50.49	0.00
CBMS_readORF	5	1436	1	309	160.06	52.59	0.76
CBMS_readORF	6	382	12	283	165.94	41.51	0.12
CBMS_readORF	7	476	32	285	161.66	42.12	-0.14
CBMS_readORF	8	688	14	302	163.69	37.18	0.86

CBM Math

The table below presents, for each grade (3-8), descriptive statistics for Fall (F), Winter (W), and Spring (S) Benchmark measures. Note: CBMF_math=Fall Math Performance, CBMW_math=Winter Math Performance, CBMS_math=Spring Math Performance.

Table 10.
easyCBM Math Statistics (by Season and Grade)

Measure-Season	Grade	N	Min	Max	Mean	SD	Kurtosis
CBMF_math	3	1207	9	39	24.52	5.23	-0.04
CBMF_math	4	1138	8	40	25.49	6.96	-0.76
CBMF_math	5	1203	1	39	24.52	6.29	-0.54
CBMF_math	6	980	1	44	23.23	7.32	-0.03
CBMF_math	7	1049	2	41	21.27	6.52	0.11
CBMF_math	8	1025	1	45	22.92	7.95	-0.21
CBMW_math	3	1242	10	40	28.76	5.64	-0.07
CBMW_math	4	1174	7	40	27.45	6.16	-0.25
CBMW_math	5	1238	8	40	27.18	6.53	-0.58
CBMW_math	6	934	1	45	25.78	7.71	-0.16
CBMW_math	7	1051	1	45	25.04	8.67	-0.57
CBMW_math	8	1052	1	45	27.17	8.25	-0.55
CBMS_math	3	1274	10	40	31.82	5.46	1.29
CBMS_math	4	1199	7	40	29.80	6.74	0.15
CBMS_math	5	1253	8	40	27.50	7.15	-0.56
CBMS_math	6	1022	1	45	26.62	8.26	-0.50
CBMS_math	7	1047	1	45	25.31	9.22	-0.53
CBMS_math	8	961	1	45	29.62	8.69	-0.54

ELA State Test Performance

The table below presents, for each grade (3-8), descriptive statistics for the Spring English Language Arts Test from Smarter Balanced (SB) test.

Table 11.*Spring English Language Arts SB test Statistics by Grade*

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1479	2126	2678	2412.84	95.71	-0.24
4	1409	2049	2739	2476.82	104.37	0.28
5	1423	2056	2807	2503.08	105.48	0.28
6	1205	2142	2780	2511.80	102.11	-0.33
7	1236	2192	2808	2547.79	101.81	-0.20
8	1253	2110	2863	2576.32	105.50	0.29

Math State Test Performance

The table below presents, for each grade (3-8), descriptive statistics for the Spring Math Test from Smarter Balanced (SB) test.

Table 12.*Spring Math SB test Statistics by Grade*

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1298	2076	2705	2439.88	84.92	0.08
4	1241	2090	2797	2487.96	88.64	0.69
5	1285	2095	2797	2504.87	95.27	0.49
6	1213	2103	2841	2503.03	112.45	0.42
7	1250	2108	2818	2523.55	114.11	-0.11
8	1260	2113	2905	2540.93	119.57	0.33

Reading Composite

Because we had three different reading measures, all with their own scale, students' performance had to be converted to a standard scale. We use z-score transformations for each measure by calculating the grade level mean and subtracting it from each student's score. That means, if the student was below the mean, they would receive a negative score (to -3.0) and if it was above the mean, they would receive a positive score (to 3.0). This comparison (difference) was divided by the standard deviation for each grade level measures. Once these individual measures were converted, they averaged.

Table 13.*Fall Reading Composite Statistics by Grade*

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1280	-2.38	2.08	0.04	0.86	-0.72
4	1208	-2.76	1.93	0.04	0.81	-0.42
5	1245	-3.57	1.6	0.05	0.8	1.29
6	628	-2.82	1.56	0.06	0.78	1.41
7	714	-3.15	1.84	0.04	0.79	0.62
8	680	-3.07	1.57	0.01	0.8	0.52

Table 14.
Winter Reading Composite Statistics by Grade

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1226	-2.5	1.91	0.01	0.85	-0.16
4	1150	-2.85	1.74	0.01	0.85	0.68
5	1285	-3.66	1.45	0.06	0.81	2.1
6	771	-2.86	1.61	0.06	0.77	1.1
7	750	-3.03	1.56	0.04	0.77	0.76
8	746	-2.96	1.64	0.02	0.75	0.7

Table 15.
Spring Reading Composite Statistics by Grade

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1361	-2.87	1.79	0.05	0.83	0.31
4	1264	-3.07	1.46	0.05	0.77	0.23
5	1315	-3.07	1.49	0.05	0.78	1.6
6	315	-3.08	1.5	0.11	0.86	1.57
7	422	-2.74	1.86	0.03	0.82	0.17
8	598	-3.06	1.91	0.05	0.77	1.23

Math easyCBM Composites (across grades)

An easyCBM Math composite was calculated in a similar manner to that used with all the reading measures. Whereas the reading measures required a z-score to allow a composite score to be calculated across different measures, each with their own scale, in mathematics, only one type of Benchmark measure (BM) was used. However, to allow comparisons across the grades in a uniform correlation with the criterion measure at all three time periods, a z-score is necessary. In this way, each student's score is compared to the mean of that grade level (taking the difference) and dividing it by the standard deviation for that same grade level. We followed four steps in creating the composite: (a) calculated the mean and standard deviation for each measure by grade level, (b) calculated the z-scores for each measure and grade level student, and (c) averaged the z-score within and across grades.

Table 16.
Fall Math Composite Statistics by Grade

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1203	-2.97	2.77	0	1	-0.03
4	1134	-2.51	2.09	0	1	-0.76
5	1196	-3.74	2.3	0	1	-0.54
6	967	-3.04	2.84	0	1	-0.01
7	1033	-2.96	3.03	0.01	1	0.14
8	1003	-2.76	2.78	0.01	1	-0.21

Table 17.
Winter Math Composite Statistics by Grade

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1239	-3.33	1.99	0	1	-0.07
4	1173	-3.32	2.04	0	1	-0.25
5	1234	-2.94	1.96	0	1	-0.58
6	921	-3.21	2.49	0	1	-0.23
7	1039	-2.77	2.3	0.01	1	-0.58
8	1032	-3.17	2.16	0.01	1	-0.53

Table 18.
Spring Math Composite Statistics by Grade

Grade	N	Min	Max	Mean	SD	Kurtosis
3	1269	-4	1.5	0	0.99	1.28
4	1198	-3.38	1.51	0	1	0.17
5	1250	-2.73	1.75	0	1	-0.56
6	1011	-3.1	2.23	0.01	1	-0.5
7	1037	-2.64	2.13	0.01	1	-0.51
8	949	-3.29	1.77	0.01	1	-0.53

CBM Measures – Predictive and Concurrent Validity

Table 19.

*Correlations Among Grade **Three** easyCBM Raw Scores and SB*

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.60	0.64	0.48	0.63	0.53
ReadORF		0.72	0.47	0.65	0.58
ReadVoc			0.47	0.64	0.57
Math				0.59	0.69
SB ELA					0.76
WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.54	0.56	0.48	0.60	0.53
ReadORF		0.68	0.55	0.66	0.61
ReadVoc			0.57	0.69	0.61
Math				0.65	0.77
SB ELA					0.76
SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.55	0.60	0.53	0.69	0.59
ReadORF		0.66	0.57	0.65	0.60
ReadVoc			0.64	0.68	0.61
Math				0.66	0.78
SB ELA					0.76

Table 20.

*Correlations Among Grade **Four** easyCBM Raw Scores and SB*

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.61	0.69	0.58	0.70	0.59
ReadORF		0.67	0.53	0.65	0.56
ReadVoc			0.60	0.67	0.61
Math				0.67	0.80
SB ELA					0.76
WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.57	0.63	0.51	0.68	0.60
ReadORF		0.65	0.48	0.63	0.54
ReadVoc			0.55	0.67	0.62
Math				0.67	0.80
SB ELA					0.76

Continued on next page.

SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.52	0.64	0.59	0.69	0.60
ReadORF		0.61	0.55	0.64	0.55
ReadVoc			0.60	0.67	0.62
Math				0.71	0.84
SB ELA					0.76

Table 21.

*Correlations Among Grade **Five** easyCBM Raw Scores and SB*

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.52	0.63	0.46	0.62	0.54
ReadORF		0.64	0.48	0.63	0.51
ReadVoc			0.52	0.66	0.57
Math				0.66	0.77
SB ELA					0.76
WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.55	0.62	0.58	0.65	0.58
ReadORF		0.62	0.53	0.62	0.51
ReadVoc			0.60	0.66	0.59
Math				0.70	0.82
SB ELA					0.76
SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.55	0.63	0.57	0.66	0.58
ReadORF		0.56	0.53	0.61	0.53
ReadVoc			0.55	0.64	0.58
Math				0.74	0.84
SB ELA					0.75

Table 22.

*Correlations Among Grade **Six** easyCBM Raw Scores and SB*

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.41	0.56	0.38	0.56	0.48
ReadORF		0.45	0.31	0.48	0.37
ReadVoc			0.45	0.62	0.51
Math				0.64	0.74
SB ELA					0.76

Continued on next page.

WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.43	0.55	0.43	0.54	0.51
ReadORF		0.47	0.37	0.51	0.40
ReadVoc			0.50	0.58	0.56
Math				0.62	0.74
SB ELA					0.73
SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.58	0.66	0.54	0.68	0.60
ReadORF		0.59	0.50	0.64	0.56
ReadVoc			0.49	0.66	0.57
Math				0.75	0.88
SB ELA					0.81

Table 23.

Correlations Among Grade Seven easyCBM Raw Scores and SB

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.39	0.58	0.47	0.63	0.54
ReadORF		0.45	0.32	0.48	0.39
ReadVoc			0.42	0.58	0.53
Math				0.59	0.74
SB ELA					0.71
WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.39	0.53	0.51	0.60	0.53
ReadORF		0.44	0.39	0.49	0.39
ReadVoc			0.49	0.59	0.55
Math				0.66	0.79
SB ELA					0.72
SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.46	0.48	0.46	0.59	0.51
ReadORF		0.54	0.39	0.56	0.44
ReadVoc			0.55	0.62	0.61
Math				0.73	0.86
SB ELA					0.76

Table 24.*Correlations Among Grade **Eight** easyCBM Raw Scores and SB*

FALL	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.42	0.51	0.40	0.53	0.41
ReadORF		0.53	0.42	0.56	0.42
ReadVoc			0.49	0.63	0.54
Math				0.68	0.79
SB ELA					0.74
WINTER	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.34	0.44	0.46	0.54	0.47
ReadORF		0.45	0.43	0.53	0.42
ReadVoc			0.56	0.64	0.57
Math				0.66	0.80
SB ELA					0.72
SPRING	ReadORF	ReadVoc	Math	SB ELA	SB MTH
ReadProf	0.43	0.52	0.43	0.58	0.44
ReadORF		0.45	0.40	0.59	0.44
ReadVoc			0.54	0.64	0.53
Math				0.70	0.86
SB ELA					0.75

CBM Composite Measures – Predictive and Concurrent Validity Coefficients

Fall Reading correlation with the ELA state test	.64
Winter Reading correlation with the ELA state test	.57
Spring Reading correlation with the ELA state test	.66
Fall Math correlation with the Math state test	.57
Winter Math correlation with the Math state test	.59
Spring Math correlation with the Math state test	.61

CBM Composite Scores Regressed on ELA State Test Performance

Next, we present plot distributions for easyCBM Reading and Math Composite Performance and State SB Test Performance displaying their concurrent distributions. In these plots, the CBM performance is displayed along the x-axis and the SB performance is displayed along the y-axis. Note that all CBMs used the composite scores because we are collapsing across measures and grades. For each plot, we ran a regression line (in blue) that summarizes this concurrent distribution.

Figure 1

Regression of CBM Fall Reading Composite on SB

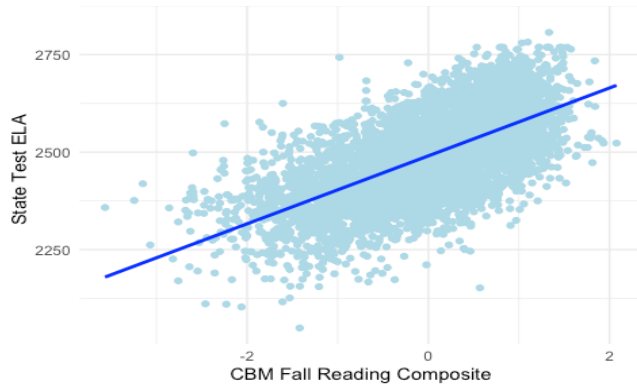


Figure 2

Regression of CBM Winter Reading Composite on SB

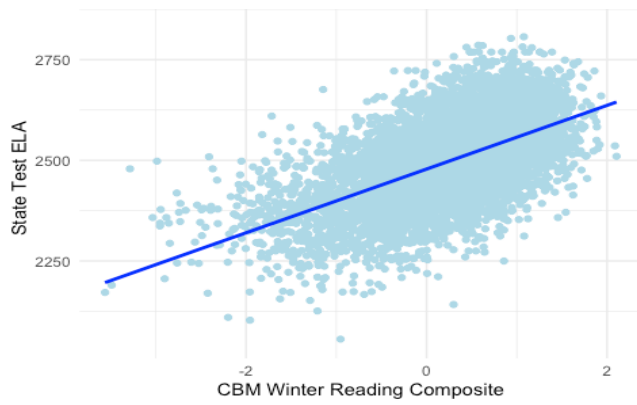


Figure 3

Regression of CBM Spring Reading Composite on SB

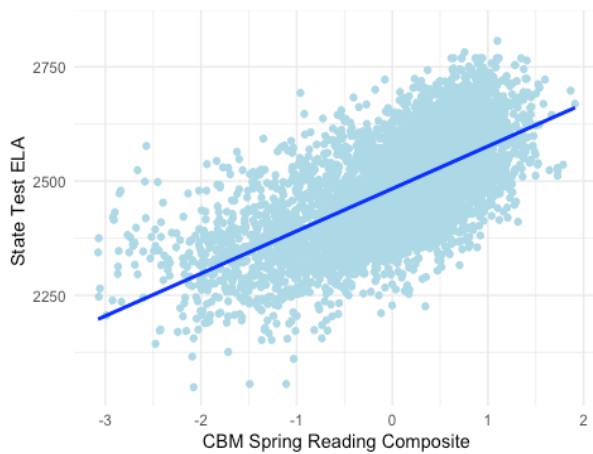


Figure 4

Regression of CBM Fall Math Composite on SB

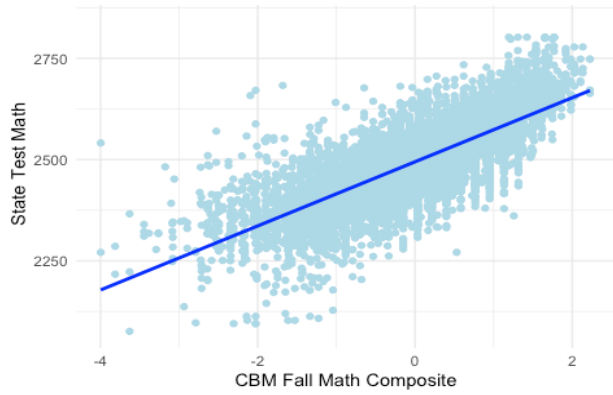


Figure 5

Regression of CBM Winter Math Composite on SB

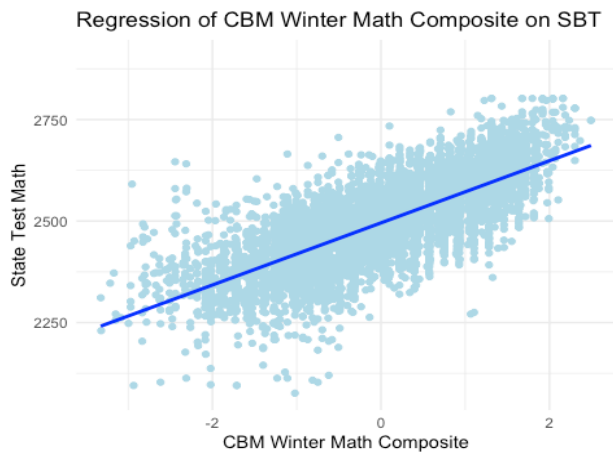
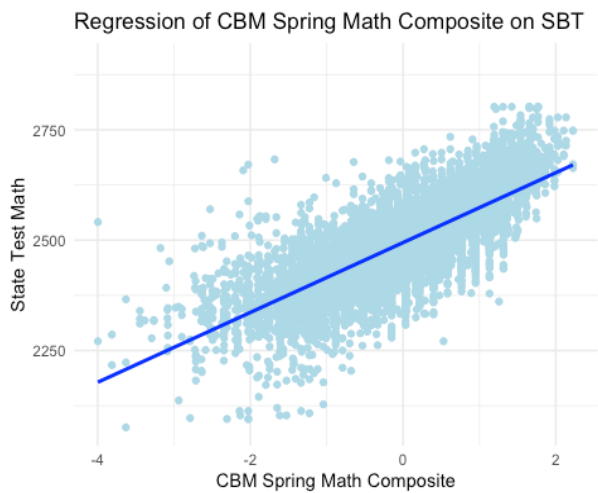


Figure 6

Regression of CBM Spring Math Composite on SB



Classification Accuracy: Evidence, Summary, and Demographics

The Classification Accuracy (CA) analysis presents important information on the use of easyCBMs to classify students with disabilities for all the reading and math measures in Grades 3 to 8. In that document, information is provided on (a) the percentile rank (PR) used in this classification (for most measures and grade, the 20th PR is optimal), (b) accuracy, including base rates, and (c) receiver operating characteristics with attention to area under the curve (AUC) and true/false positives and negatives.

In this section, we present three tables for each season/test displaying classification accuracy values reflecting (a) evidence, (b) summary (statistics), and (c) demographics. The rows with gray highlight show values required by NCII. Bold values reflect convincing evidence, non-bold reflects partially convincing evidence, and red font reflect unconvincing evidence.

Table 25.

Classification Accuracy Fall 2023: Proficient Reading CBM with SB

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	10	11	13	13	13	12
Classification Data - True Positive (a)	713	760	831	625	541	659
Classification Data - False Positive (b)	67	39	57	59	34	41
Classification Data - False Negative (c)	299	255	214	130	251	223
Classification Data - True Negative (d)	217	169	159	154	161	112
Area Under the Curve (AUC)	0.80	0.85	0.83	0.83	0.83	0.80
AUC Estimate's 95% Confidence Interval: Lower Bound	0.80	0.85	0.83	0.83	0.83	0.80
AUC Estimate's 95% Confidence Interval: Upper Bound	0.81	0.85	0.83	0.83	0.83	0.80

Table 26.*Classification Accuracy Summary Fall 2023: Proficient Reading CBM with SB*

Statistics	Grade3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1296	1223	1261	968	987	1035
Base Rate	0.22	0.17	0.17	0.22	0.20	0.15
Overall Classification Rate	0.72	0.76	0.79	0.80	0.71	0.74
Sensitivity	0.70	0.75	0.80	0.83	0.68	0.75
Specificity	0.76	0.81	0.74	0.72	0.83	0.73
False Positive Rate	0.24	0.19	0.26	0.28	0.17	0.27
False Negative Rate	0.30	0.25	0.20	0.17	0.32	0.25
Positive Predictive Power	0.91	0.95	0.94	0.91	0.94	0.94
Negative Predictive Power	0.42	0.40	0.43	0.54	0.39	0.33

Table 27.*Classification Accuracy Demographics Fall 2023: Proficient Reading CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	678	637	651	506	520	556
Female	618	586	607	461	466	470
Gender Unknown	13	22	15	1	1	9
American Indian/Alaska Native	17	29	15	9	14	17
Asian/Pacific Islander	90	70	78	66	77	81
Black	97	100	118	81	93	92
Multi-Racial	229	241	222	181	183	202
Native Hawaiian/Pacific Islander	35	28	45	33	24	42
White	815	733	771	598	596	601
Hispanic/Latino	267	276	279	207	220	230
Not Hispanic/Latino	1029	947	982	761	767	805
IEP or diagnosed disability	270	214	215	140	153	154
English Language Learner	105	79	67	55	52	41

Table 28.*Classification Accuracy Fall 2023: Oral Reading Fluency CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	66	90	129	134	151	148
Classification Data - True Positive (a)	819	862	844	411	367	472
Classification Data - False Positive (b)	63	63	65	42	35	49
Classification Data - False Negative (c)	244	202	237	102	222	113

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - True Negative (d)	250	163	175	102	115	73
Area Under the Curve (AUC)	0.85	0.84	0.81	0.82	0.77	0.75
AUC Estimate's 95% Confidence Interval: Lower Bound	0.85	0.84	0.81	0.82	0.77	0.75
AUC Estimate's 95% Confidence Interval: Upper Bound	0.85	0.84	0.82	0.82	0.77	0.75

Table 29.

Classification Accuracy Summary Fall 2023: Oral Reading Fluency CBM with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1376	1290	1321	657	739	707
Base Rate	0.23	0.18	0.18	0.22	0.20	0.17
Overall Classification Rate	0.78	0.79	0.77	0.78	0.65	0.77
Sensitivity	0.77	0.81	0.78	0.80	0.62	0.81
Specificity	0.80	0.72	0.73	0.71	0.77	0.60
False Positive Rate	0.20	0.28	0.27	0.29	0.23	0.40
False Negative Rate	0.23	0.19	0.22	0.20	0.38	0.19
Positive Predictive Power	0.93	0.93	0.93	0.91	0.91	0.91
Negative Predictive Power	0.51	0.45	0.42	0.50	0.34	0.39

Table 30.

Classification Accuracy Demographics: Oral Reading Fluency CBM with SB

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	716	670	678	350	402	388
Female	660	620	640	306	336	314
Gender Unknown	20	34	21	1	1	5
American Indian/Alaska Native	18	29	16	7	9	10
Asian/Pacific Islander	92	70	77	42	57	60
Black	101	101	118	52	65	64
Multi-Racial	236	246	225	112	125	117
Native Hawaiian/Pacific Islander	36	28	45	21	18	28
White	873	782	822	423	465	428
Hispanic/Latino	279	289	286	139	161	153
Not Hispanic/Latino	1097	1001	1035	518	578	554
IEP or diagnosed disability	293	223	223	96	106	124
English Language Learner	106	83	68	38	40	30

Table 31.*Classification Accuracy Fall 2023: Vocabulary CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	14	15	16	17	17	17
Classification Data - True Positive (a)	799	787	809	610	632	695
Classification Data - False Positive (b)	74	36	52	70	61	40
Classification Data - False Negative (c)	212	228	232	138	158	186
Classification Data - True Negative (d)	223	173	168	140	135	117
Area Under the Curve (AUC)	0.83	0.84	0.84	0.82	0.81	0.84
AUC Estimate's 95% Confidence Interval: Lower Bound	0.83	0.84	0.84	0.81	0.81	0.83
AUC Estimate's 95% Confidence Interval: Upper Bound	0.83	0.84	0.84	0.82	0.81	0.84

Table 32.*Classification Accuracy Summary Fall 2023: Vocabulary CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1308	1224	1261	958	986	1038
Base Rate	0.23	0.17	0.17	0.22	0.20	0.15
Overall Classification Rate	0.78	0.78	0.77	0.78	0.78	0.78
Sensitivity	0.79	0.78	0.78	0.82	0.80	0.79
Specificity	0.75	0.83	0.76	0.67	0.69	0.75
False Positive Rate	0.25	0.17	0.24	0.33	0.31	0.25
False Negative Rate	0.21	0.22	0.22	0.18	0.20	0.21
Positive Predictive Power	0.92	0.96	0.94	0.90	0.91	0.95
Negative Predictive Power	0.51	0.43	0.42	0.50	0.46	0.39

Table 33.*Classification Accuracy Demographics Fall 2023: Vocabulary CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	685	637	649	494	524	552
Female	623	587	609	463	461	477
Gender Unknown	14	20	14	1	1	9
American Indian/Alaska Native	17	29	15	9	14	16
Asian/Pacific Islander	91	70	78	65	77	86
Black	98	101	119	79	92	90
Multi-Racial	233	241	221	179	183	202
Native Hawaiian/Pacific Islander	35	27	45	33	26	43
White	820	736	772	593	594	601
Hispanic/Latino	270	274	279	206	220	232
Not Hispanic/Latino	1038	950	982	752	766	806
IEP or diagnosed disability	276	215	219	133	155	150
English Language Learner	106	79	68	51	52	42

Table 34.*Classification Accuracy Winter 2024: Proficient Reading CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	10	12	15	12	13	11
Classification Data - True Positive (a)	710	808	849	606	672	683
Classification Data - False Positive (b)	66	40	55	81	68	55
Classification Data - False Negative (c)	264	173	222	158	129	205
Classification Data - True Negative (d)	190	142	183	146	138	100
Area Under the Curve (AUC)	0.80	0.85	0.85	0.79	0.81	0.76
AUC Estimate's 95% Confidence Interval: Lower Bound	0.80	0.85	0.85	0.79	0.81	0.76
AUC Estimate's 95% Confidence Interval: Upper Bound	0.80	0.85	0.85	0.79	0.82	0.77

Table 35.*Classification Accuracy Summary Winter 2024: Proficient Reading CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1230	1163	1309	991	1007	1043
Base Rate	0.21	0.16	0.18	0.23	0.20	0.15
Overall Classification Rate	0.73	0.82	0.79	0.76	0.80	0.75
Sensitivity	0.73	0.82	0.79	0.79	0.84	0.77
Specificity	0.74	0.78	0.77	0.64	0.67	0.65
False Positive Rate	0.26	0.22	0.23	0.36	0.33	0.35
False Negative Rate	0.27	0.18	0.21	0.21	0.16	0.23
Positive Predictive Power	0.91	0.95	0.94	0.88	0.91	0.93
Negative Predictive Power	0.42	0.45	0.45	0.48	0.52	0.33

Table 36.*Classification Accuracy Demographics Winter 2024: Proficient Reading CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	641	601	671	524	533	553
Female	589	562	635	466	473	482
Gender Unknown	0	0	15	1	1	8
American Indian/Alaska Native	18	31	15	7	17	18
Asian/Pacific Islander	94	71	78	66	76	84
Black	103	109	126	89	97	86
Multi-Racial	238	244	232	192	182	206
Native Hawaiian/Pacific Islander	37	30	46	33	29	40
White	740	678	800	604	606	609
Hispanic/Latino	270	264	293	211	219	233
Not Hispanic/Latino	960	899	1016	780	788	810
IEP or diagnosed disability	235	198	226	148	162	158
English Language Learner	103	74	69	50	54	40

Table 37.*Classification Accuracy Winter 2024: Oral Reading Fluency CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	92	122	150	156	169	134
Classification Data - True Positive (a)	807	768	723	452	400	537
Classification Data - False Positive (b)	68	35	38	53	38	50
Classification Data - False Negative (c)	170	211	390	177	229	130

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - True Negative (d)	192	150	219	127	119	72
Area Under the Curve (AUC)	0.85	0.86	0.82	0.78	0.77	0.77
AUC Estimate's 95% Confidence Interval: Lower Bound	0.85	0.86	0.82	0.78	0.77	0.77
AUC Estimate's 95% Confidence Interval: Upper Bound	0.86	0.86	0.82	0.78	0.77	0.77

Table 38.
Classification Accuracy Summary Winter 2024: Oral Reading Fluency CBM with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1237	1164	1370	809	786	789
Base Rate	0.21	0.16	0.19	0.22	0.20	0.15
Overall Classification Rate	0.81	0.79	0.69	0.72	0.66	0.77
Sensitivity	0.83	0.78	0.65	0.72	0.64	0.81
Specificity	0.74	0.81	0.85	0.71	0.76	0.59
False Positive Rate	0.26	0.19	0.15	0.29	0.24	0.41
False Negative Rate	0.17	0.22	0.35	0.28	0.36	0.19
Positive Predictive Power	0.92	0.96	0.95	0.90	0.91	0.91
Negative Predictive Power	0.53	0.42	0.36	0.42	0.34	0.36

Table 39.
Classification Accuracy Demographics Winter 2024: Oral Reading Fluency CBM with SB

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	645	606	702	433	417	428
Female	592	558	665	375	367	355
Gender Unknown	0	0	21	1	2	6
American Indian/Alaska Native	18	31	17	6	13	14
Asian/Pacific Islander	94	70	78	53	56	62
Black	103	109	126	65	68	58
Multi-Racial	239	245	236	157	145	146
Native Hawaiian/Pacific Islander	38	31	47	30	19	32
White	745	678	848	498	485	477
Hispanic/Latino	271	267	302	166	167	171
Not Hispanic/Latino	966	897	1068	643	619	618
IEP or diagnosed disability	240	199	237	101	115	118
English Language Learner	104	77	69	38	41	30

Table 40.*Classification Accuracy Winter 2024: Vocabulary CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	15	17	16	17	17	17
Classification Data - True Positive (a)	829	748	865	579	607	711
Classification Data - False Positive (b)	66	23	57	48	70	35
Classification Data - False Negative (c)	146	229	200	180	194	185
Classification Data - True Negative (d)	191	162	174	180	144	122
Area Under the Curve (AUC)	0.87	0.89	0.87	0.85	0.80	0.85
AUC Estimate's 95% Confidence Interval: Lower Bound	0.87	0.89	0.87	0.85	0.80	0.85
AUC Estimate's 95% Confidence Interval: Upper Bound	0.87	0.89	0.87	0.85	0.81	0.85

Table 41.*Classification Accuracy Summary Winter 2024: Vocabulary CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1232	1162	1296	987	1015	1053
Base Rate	0.21	0.16	0.18	0.23	0.21	0.15
Overall Classification Rate	0.83	0.78	0.80	0.77	0.74	0.79
Sensitivity	0.85	0.77	0.81	0.76	0.76	0.79
Specificity	0.74	0.88	0.75	0.79	0.67	0.78
False Positive Rate	0.26	0.12	0.25	0.21	0.33	0.22
False Negative Rate	0.15	0.23	0.19	0.24	0.24	0.21
Positive Predictive Power	0.93	0.97	0.94	0.92	0.90	0.95
Negative Predictive Power	0.57	0.41	0.47	0.50	0.43	0.40

Table 42.*Classification Accuracy Demographics Winter 2024: Vocabulary CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	641	601	661	521	536	557
Female	591	561	632	465	478	487
Gender Unknown	0	0	14	1	1	9
American Indian/Alaska Native	18	31	15	8	17	17
Asian/Pacific Islander	93	71	78	66	76	84
Black	102	109	125	85	98	93
Multi-Racial	238	244	232	190	186	209
Native Hawaiian/Pacific Islander	38	30	47	32	28	40
White	743	677	788	606	610	610
Hispanic/Latino	270	262	292	211	223	236
Not Hispanic/Latino	962	900	1004	776	792	817
IEP or diagnosed disability	237	199	223	149	164	156
English Language Learner	103	73	68	51	54	40

Table 43.*Classification Accuracy Spring 2024: Proficient Reading CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	13	13	13	12	11	11
Classification Data - True Positive (a)	770	832	925	596	546	637
Classification Data - False Positive (b)	40	41	68	55	46	29
Classification Data - False Negative (c)	284	235	162	113	246	215
Classification Data - True Negative (d)	284	185	187	144	144	109
Area Under the Curve (AUC)	0.85	0.87	0.85	0.85	0.78	0.83
AUC Estimate's 95% Confidence Interval: Lower Bound	0.85	0.86	0.85	0.85	0.78	0.82
AUC Estimate's 95% Confidence Interval: Upper Bound	0.86	0.87	0.85	0.85	0.78	0.83

Table 44.*Classification Accuracy Summary Spring 2024: Proficient Reading CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1378	1293	1342	908	982	990
Base Rate	0.24	0.17	0.19	0.22	0.19	0.14
Overall Classification Rate	0.76	0.79	0.83	0.81	0.70	0.75
Sensitivity	0.73	0.78	0.85	0.84	0.69	0.75
Specificity	0.88	0.82	0.73	0.72	0.76	0.79
False Positive Rate	0.12	0.18	0.27	0.28	0.24	0.21
False Negative Rate	0.27	0.22	0.15	0.16	0.31	0.25
Positive Predictive Power	0.95	0.95	0.93	0.92	0.92	0.96
Negative Predictive Power	0.50	0.44	0.54	0.56	0.37	0.34

Table 45.*Classification Accuracy Demographics Spring 2024: Reading Proficiency CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	721	663	686	482	513	534
Female	656	630	653	426	467	449
Gender Unknown	15	20	17	0	2	7
American Indian/Alaska Native	18	31	15	8	14	16
Asian/Pacific Islander	97	73	78	62	73	85
Black	106	115	131	79	102	86
Multi-Racial	249	253	237	170	185	189
Native Hawaiian/Pacific Islander	39	30	48	28	23	34
White	855	771	819	561	585	580
Hispanic/Latino	289	288	304	192	211	219
Not Hispanic/Latino	1089	1005	1038	716	771	771
IEP or diagnosed disability	293	229	229	133	143	155
English Language Learner	108	82	75	46	51	35

Table 46.*Classification Accuracy Spring 2024: Oral Reading Fluency CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	99	110	147	146	144	156
Classification Data - True Positive (a)	783	924	871	237	285	390
Classification Data - False Positive (b)	57	68	64	21	32	19
Classification Data - False Negative (c)	321	197	257	59	71	180

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - True Negative (d)	289	179	195	57	77	85
Area Under the Curve (AUC)	0.84	0.83	0.82	0.83	0.81	0.80
AUC Estimate's 95% Confidence Interval: Lower Bound	0.84	0.83	0.82	0.83	0.81	0.79
AUC Estimate's 95% Confidence Interval: Upper Bound	0.84	0.84	0.82	0.83	0.81	0.80

Table 47.

Classification Accuracy Summary Spring 2024: Oral Reading Fluency CBM with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1450	1368	1387	374	465	674
Base Rate	0.24	0.18	0.19	0.21	0.23	0.15
Overall Classification Rate	0.74	0.81	0.77	0.79	0.78	0.70
Sensitivity	0.71	0.82	0.77	0.80	0.80	0.68
Specificity	0.84	0.72	0.75	0.73	0.71	0.82
False Positive Rate	0.16	0.28	0.25	0.27	0.29	0.18
False Negative Rate	0.29	0.18	0.23	0.20	0.20	0.32
Positive Predictive Power	0.93	0.93	0.93	0.92	0.90	0.95
Negative Predictive Power	0.47	0.48	0.43	0.49	0.52	0.32

Table 48.

Classification Accuracy Demographics Spring 2024: Oral Reading Fluency CBM with SB

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	757	702	711	195	251	372
Female	692	666	673	179	211	294
Gender Unknown	23	34	21	0	3	8
American Indian/Alaska Native	21	31	17	3	9	11
Asian/Pacific Islander	97	73	79	17	26	51
Black	109	115	130	23	34	63
Multi-Racial	252	259	238	45	66	112
Native Hawaiian/Pacific Islander	39	32	48	7	7	27
White	910	824	857	279	323	410
Hispanic/Latino	303	306	308	69	88	150
Not Hispanic/Latino	1147	1062	1079	305	377	524
IEP or diagnosed disability	309	248	230	68	88	103
English Language Learner	111	88	75	21	20	25

Table 49.*Classification Accuracy Spring 2024: Vocabulary CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	17	16	17	17	16	18
Classification Data - True Positive (a)	826	896	760	605	668	642
Classification Data - False Positive (b)	68	58	46	65	69	34
Classification Data - False Negative (c)	227	164	318	109	135	217
Classification Data - True Negative (d)	260	164	202	129	125	107
Area Under the Curve (AUC)	0.87	0.85	0.82	0.83	0.82	0.83
AUC Estimate's 95% Confidence Interval: Lower Bound	0.87	0.85	0.82	0.83	0.82	0.83
AUC Estimate's 95% Confidence Interval: Upper Bound	0.87	0.85	0.82	0.84	0.82	0.84

Table 50.*Classification Accuracy Summary Spring 2024: Vocabulary CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1381	1282	1326	908	997	1000
Base Rate	0.24	0.17	0.19	0.21	0.19	0.14
Overall Classification Rate	0.79	0.83	0.73	0.81	0.80	0.75
Sensitivity	0.78	0.85	0.71	0.85	0.83	0.75
Specificity	0.79	0.74	0.81	0.66	0.64	0.76
False Positive Rate	0.21	0.26	0.19	0.34	0.36	0.24
False Negative Rate	0.22	0.15	0.29	0.15	0.17	0.25
Positive Predictive Power	0.92	0.94	0.94	0.90	0.91	0.95
Negative Predictive Power	0.53	0.50	0.39	0.54	0.48	0.33

Table 51.*Classification Accuracy Demographics Spring 2024: Vocabulary CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	723	660	678	469	527	545
Female	657	622	645	439	468	448
Gender Unknown	17	17	15	0	2	7
American Indian/Alaska Native	20	30	15	8	16	15
Asian/Pacific Islander	96	72	78	61	73	85
Black	107	114	129	84	103	83
Multi-Racial	250	253	236	172	187	188
Native Hawaiian/Pacific Islander	39	31	48	29	26	37
White	853	765	808	554	592	592
Hispanic/Latino	292	287	301	193	216	219
Not Hispanic/Latino	1089	995	1025	715	781	781
IEP or diagnosed disability	291	232	231	117	139	148
English Language Learner	109	82	74	48	51	38

Table 52.*Classification Accuracy Fall 2023: Math CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	23	23	23	21	21	21
Classification Data - True Positive (a)	743	739	716	573	511	560
Classification Data - False Positive (b)	45	13	13	42	29	19
Classification Data - False Negative (c)	247	228	303	169	298	270
Classification Data - True Negative (d)	168	154	164	183	195	154
Area Under the Curve (AUC)	0.84	0.92	0.89	0.86	0.81	0.84
AUC Estimate's 95% Confidence Interval: Lower Bound	0.84	0.91	0.89	0.86	0.81	0.84
AUC Estimate's 95% Confidence Interval: Upper Bound	0.85	0.92	0.89	0.86	0.81	0.84

Table 53.*Classification Accuracy Summary Fall 2023: Math CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1203	1134	1196	967	1033	1003
Base Rate	0.18	0.15	0.15	0.23	0.22	0.17
Overall Classification Rate	0.76	0.79	0.74	0.78	0.68	0.71
Sensitivity	0.75	0.76	0.70	0.77	0.63	0.67
Specificity	0.79	0.92	0.93	0.81	0.87	0.89
False Positive Rate	0.21	0.08	0.07	0.19	0.13	0.11
False Negative Rate	0.25	0.24	0.30	0.23	0.37	0.33
Positive Predictive Power	0.94	0.98	0.98	0.93	0.95	0.97
Negative Predictive Power	0.40	0.40	0.35	0.52	0.40	0.36

Table 54.*Classification Accuracy Demographics Fall 2023: Math CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	630	592	615	484	536	523
Female	573	542	578	482	495	470
Gender Unknown	0	0	3	1	2	10
American Indian/Alaska Native	16	28	13	7	15	13
Asian/Pacific Islander	92	71	79	61	79	79
Black	101	101	122	87	100	83
Multi-Racial	228	239	223	185	192	195
Native Hawaiian/Pacific Islander	36	28	45	33	25	40
White	730	667	714	594	622	593
Hispanic/Latino	260	256	273	217	233	210
Not Hispanic/Latino	943	878	923	750	800	793
IEP or diagnosed disability	237	194	194	146	158	146
English Language Learner	105	77	64	61	55	36

Table 55.*Classification Accuracy Winter 2024: Math CBM with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	27	25	24	23	22	24
Classification Data - True Positive (a)	816	796	851	574	615	661
Classification Data - False Positive (b)	21	26	23	36	24	26
Classification Data - False Negative (c)	199	200	195	134	195	193

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - True Negative (d)	203	151	165	177	205	152
Area Under the Curve (AUC)	0.92	0.91	0.92	0.88	0.88	0.88
AUC Estimate's 95% Confidence Interval: Lower Bound	0.91	0.91	0.92	0.89	0.88	0.88
AUC Estimate's 95% Confidence Interval: Upper Bound	0.92	0.92	0.93	0.89	0.88	0.88

Table 56.*Classification Accuracy Summary Winter 2024: Math CBM with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1239	1173	1234	921	1039	1032
Base Rate	0.18	0.15	0.15	0.23	0.22	0.17
Overall Classification Rate	0.82	0.81	0.82	0.82	0.79	0.79
Sensitivity	0.80	0.80	0.81	0.81	0.76	0.77
Specificity	0.91	0.85	0.88	0.83	0.90	0.85
False Positive Rate	0.09	0.15	0.12	0.17	0.10	0.15
False Negative Rate	0.20	0.20	0.19	0.19	0.24	0.23
Positive Predictive Power	0.97	0.97	0.97	0.94	0.96	0.96
Negative Predictive Power	0.50	0.43	0.46	0.57	0.51	0.44

Table 57.*Classification Accuracy Demographics Winter 2024: Math CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	646	609	633	469	530	542
Female	593	564	598	451	506	481
Gender Unknown	0	0	3	1	3	9
American Indian/Alaska Native	18	31	13	6	16	16
Asian/Pacific Islander	93	72	80	56	75	78
Black	106	111	130	82	106	85
Multi-Racial	239	247	234	178	197	200
Native Hawaiian/Pacific Islander	38	31	47	27	26	41
White	745	681	730	572	619	612
Hispanic/Latino	274	268	290	203	229	222
Not Hispanic/Latino	965	905	944	718	810	810
IEP or diagnosed disability	239	201	202	137	160	155
English Language Learner	107	79	67	46	54	35

Table 58.
Classification Accuracy Spring 2024: Math CBM with SB

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	30	28	22	23	24	25
Classification Data - True Positive (a)	892	815	953	659	580	667
Classification Data - False Positive (b)	42	13	37	38	16	18
Classification Data - False Negative (c)	141	201	99	106	234	135
Classification Data - True Negative (d)	194	169	161	208	207	129
Area Under the Curve (AUC)	0.92	0.94	0.94	0.92	0.88	0.92
AUC Estimate's 95% Confidence Interval: Lower Bound	0.92	0.94	0.94	0.91	0.87	0.92
AUC Estimate's 95% Confidence Interval: Upper Bound	0.92	0.94	0.94	0.92	0.88	0.92

Table 59.
Classification Accuracy Summary Spring 2024: Math CBM with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1269	1198	1250	1011	1037	949
Base Rate	0.19	0.15	0.16	0.24	0.22	0.15
Overall Classification Rate	0.86	0.82	0.89	0.86	0.76	0.84
Sensitivity	0.86	0.80	0.91	0.86	0.71	0.83
Specificity	0.82	0.93	0.81	0.85	0.93	0.88
False Positive Rate	0.18	0.07	0.19	0.15	0.07	0.12
False Negative Rate	0.14	0.20	0.09	0.14	0.29	0.17
Positive Predictive Power	0.96	0.98	0.96	0.95	0.97	0.97
Negative Predictive Power	0.58	0.46	0.62	0.66	0.47	0.49

Table 60.*Classification Accuracy Demographics Spring 2024: Math CBM with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	664	620	637	517	540	501
Female	604	578	610	493	493	438
Gender Unknown	1	0	3	1	4	10
American Indian/Alaska Native	19	31	13	10	18	12
Asian/Pacific Islander	98	73	79	62	78	74
Black	111	116	135	100	106	78
Multi-Racial	244	250	238	189	196	174
Native Hawaiian/Pacific Islander	39	31	47	32	25	34
White	758	697	738	618	614	577
Hispanic/Latino	284	270	297	237	235	199
Not Hispanic/Latino	985	928	953	774	802	750
IEP or diagnosed disability	247	209	199	150	147	145
English Language Learner	113	80	74	64	54	29

Table 61.*Classification Accuracy Fall 2023: Reading Composite with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	-0.330475	-0.436023	-0.131578	-0.014358	-0.116780	-0.289957
Classification Data - True Positive (a)	801	847	824	397	444	448
Classification Data - False Positive (b)	54	35	37	16	23	27
Classification Data - False Negative (c)	202	159	211	102	130	121
Classification Data - True Negative (d)	223	167	173	113	117	84
Area Under the Curve (AUC)	0.87	0.90	0.88	0.90	0.87	0.83
AUC Estimate's 95% Confidence Interval: Lower Bound	0.87	0.90	0.88	0.90	0.87	0.83
AUC Estimate's 95% Confidence Interval: Upper Bound	0.88	0.90	0.88	0.90	0.87	0.84

Table 62.*Classification Accuracy Summary Fall 2023: Reading Composite with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1280	1208	1245	628	714	680
Base Rate	0.22	0.17	0.17	0.21	0.20	0.16
Overall Classification Rate	0.80	0.84	0.80	0.81	0.79	0.78
Sensitivity	0.80	0.84	0.80	0.80	0.77	0.79
Specificity	0.81	0.83	0.82	0.88	0.84	0.76
False Positive Rate	0.19	0.17	0.18	0.12	0.16	0.24
False Negative Rate	0.20	0.16	0.20	0.20	0.23	0.21
Positive Predictive Power	0.94	0.96	0.96	0.96	0.95	0.94
Negative Predictive Power	0.52	0.51	0.45	0.53	0.47	0.41

Table 63.*Classification Accuracy Demographics Fall 2023: Reading Composite with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	667	628	640	331	387	372
Female	613	580	602	296	326	303
Gender Unknown	12	18	14	1	1	5
American Indian/Alaska Native	17	29	14	7	9	9
Asian/Pacific Islander	90	70	77	41	55	58
Black	96	100	117	46	62	58
Multi-Racial	229	240	221	106	122	108
Native Hawaiian/Pacific Islander	35	27	45	20	16	28
White	801	724	760	408	450	419
Hispanic/Latino	264	271	277	131	153	144
Not Hispanic/Latino	1016	937	968	497	561	536
IEP or diagnosed disability	260	207	210	90	103	120
English Language Learner	105	77	67	37	38	29

Table 64.*Classification Accuracy Winter 2024: Reading Composite with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score	-0.289215	-0.369622	-0.122265	-0.072003	-0.260689	-0.370871

Evidence (numeric) on screener measure	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - True Positive (a)	788	806	871	456	507	526
Classification Data - False Positive (b)	40	23	34	28	28	23
Classification Data - False Negative (c)	184	164	188	153	103	114
Classification Data - True Negative (d)	214	157	192	134	112	83
Area Under the Curve (AUC)	0.90	0.92	0.90	0.87	0.87	0.87
AUC Estimate's 95% Confidence Interval: Lower Bound	0.90	0.92	0.90	0.86	0.87	0.87
AUC Estimate's 95% Confidence Interval: Upper Bound	0.90	0.92	0.90	0.87	0.88	0.87

Table 65.
Classification Accuracy Summary Winter 2024: Reading Composite with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1226	1150	1285	771	750	746
Base Rate	0.21	0.16	0.18	0.21	0.19	0.14
Overall Classification Rate	0.82	0.84	0.83	0.77	0.83	0.82
Sensitivity	0.81	0.83	0.82	0.75	0.83	0.82
Specificity	0.84	0.87	0.85	0.83	0.80	0.78
False Positive Rate	0.16	0.13	0.15	0.17	0.20	0.22
False Negative Rate	0.19	0.17	0.18	0.25	0.17	0.18
Positive Predictive Power	0.95	0.97	0.96	0.94	0.95	0.96
Negative Predictive Power	0.54	0.49	0.51	0.47	0.52	0.42

Table 66.
Classification Accuracy Demographics Winter 2024: Reading Composite with SB

Sample	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	638	595	655	412	400	399
Female	588	555	627	358	349	341
Gender Unknown	0	0	14	1	1	6
American Indian/Alaska Native	18	31	15	6	12	13
Asian/Pacific Islander	93	70	77	52	55	62
Black	102	107	125	60	62	54

Sample	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Multi-Racial	237	242	230	155	135	134
Native Hawaiian/Pacific Islander	37	30	46	25	18	29
White	739	670	781	473	468	454
Hispanic/Latino	269	260	290	157	153	155
Not Hispanic/Latino	957	890	995	614	597	591
IEP or diagnosed disability	234	195	217	94	107	113
English Language Learner	102	72	67	32	35	26

Table 67.*Classification Accuracy Spring 2024: Reading Composite with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2342	2378	2411	2432	2456	2468
Cut Points - Corresponding performance score (numeric) on screener measure	-0.27360	-0.15965	-0.14045	-0.16345	-0.35961	-0.33043
Classification Data - True Positive (a)	909	836	891	220	275	428
Classification Data - False Positive (b)	52	24	41	9	21	14
Classification Data - False Negative (c)	139	213	184	41	55	93
Classification Data - True Negative (d)	261	191	199	45	71	63
Area Under the Curve (AUC)	0.91	0.90	0.89	0.92	0.88	0.89
AUC Estimate's 95% Confidence Interval: Lower Bound	0.91	0.90	0.89	0.92	0.87	0.89
AUC Estimate's 95% Confidence Interval: Upper Bound	0.92	0.90	0.89	0.92	0.88	0.89

Table 68.*Classification Accuracy Summary Spring 2024: Reading Composite with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1361	1264	1315	315	422	598
Base Rate	0.23	0.17	0.18	0.17	0.22	0.13
Overall Classification Rate	0.86	0.81	0.83	0.84	0.82	0.82
Sensitivity	0.87	0.80	0.83	0.84	0.83	0.82
Specificity	0.83	0.89	0.83	0.83	0.77	0.82
False Positive Rate	0.17	0.11	0.17	0.17	0.23	0.18
False Negative Rate	0.13	0.20	0.17	0.16	0.17	0.18
Positive Predictive Power	0.95	0.97	0.96	0.96	0.93	0.97
Negative Predictive Power	0.65	0.47	0.52	0.52	0.56	0.40

Table 69.

Classification Accuracy Demographics Spring 2024: Reading Composite with SB

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	711	647	673	168	231	335
Female	649	617	639	147	189	258
Gender Unknown	15	17	15	0	2	5
American Indian/Alaska Native	18	30	15	2	6	9
Asian/Pacific Islander	96	72	77	14	22	51
Black	105	113	128	16	32	49
Multi-Racial	248	248	234	39	55	95
Native Hawaiian/Pacific Islander	39	30	48	4	7	21
White	841	754	801	240	300	373
Hispanic/Latino	285	282	299	53	72	125
Not Hispanic/Latino	1076	982	1016	262	350	473
IEP or diagnosed disability	283	224	223	61	77	97
English Language Learner	106	79	74	17	14	18

Table 70.*Classification Accuracy Fall 2023: Math Composite with SB with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	-0.291377	-0.357777	-0.241909	-0.304458	-0.040674	-0.241607
Classification Data - True Positive (a)	743	739	716	573	511	560
Classification Data - False Positive (b)	45	13	13	42	29	19
Classification Data - False Negative (c)	247	228	303	169	298	270
Classification Data - True Negative (d)	168	154	164	183	195	154
Area Under the Curve (AUC)	0.84	0.92	0.89	0.86	0.81	0.84
AUC Estimate's 95% Confidence Interval: Lower Bound	0.84	0.91	0.89	0.86	0.81	0.84
AUC Estimate's 95% Confidence Interval: Upper Bound	0.85	0.92	0.89	0.86	0.81	0.84

Table 71.*Classification Accuracy Summary Fall 2023: Math Composite with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1203	1134	1196	967	1033	1003
Base Rate	0.18	0.15	0.15	0.23	0.22	0.17
Overall Classification Rate	0.76	0.79	0.74	0.78	0.68	0.71
Sensitivity	0.75	0.76	0.70	0.77	0.63	0.67
Specificity	0.79	0.92	0.93	0.81	0.87	0.89
False Positive Rate	0.21	0.08	0.07	0.19	0.13	0.11
False Negative Rate	0.25	0.24	0.30	0.23	0.37	0.33
Positive Predictive Power	0.94	0.98	0.98	0.93	0.95	0.97
Negative Predictive Power	0.40	0.40	0.35	0.52	0.40	0.36

Table 72.*Classification Accuracy Demographics Fall 2023: Math Composite with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	630	592	615	484	536	523
Female	573	542	578	482	495	470
Gender Unknown	0	0	3	1	2	10
American Indian/Alaska Native	16	28	13	7	15	13
Asian/Pacific Islander	92	71	79	61	79	79
Black	101	101	122	87	100	83
Multi-Racial	228	239	223	185	192	195
Native Hawaiian/Pacific Islander	36	28	45	33	25	40
White	730	667	714	594	622	593
Hispanic/Latino	260	256	273	217	233	210
Not Hispanic/Latino	943	878	923	750	800	793
IEP or diagnosed disability	237	194	194	146	158	146
English Language Learner	105	77	64	61	55	36

Table 73.*Classification Accuracy Winter 2024: Math Composite with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	-0.31240	-0.39760	-0.48654	-0.36064	-0.35093	-0.38393
Classification Data - True Positive (a)	816	796	851	574	615	661
Classification Data - False Positive (b)	21	26	23	36	24	26
Classification Data - False Negative (c)	199	200	195	134	195	193
Classification Data - True Negative (d)	203	151	165	177	205	152
Area Under the Curve (AUC)	0.92	0.91	0.92	0.88	0.88	0.88
AUC Estimate's 95% Confidence Interval: Lower Bound	0.92	0.91	0.92	0.88	0.88	0.88
AUC Estimate's 95% Confidence Interval: Upper Bound	0.92	0.92	0.93	0.89	0.88	0.88

Table 74.*Classification Accuracy Summary Winter 2024: Math Composite with SB*

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1239	1173	1234	921	1039	1032
Base Rate	0.18	0.15	0.15	0.23	0.22	0.17
Overall Classification Rate	0.82	0.81	0.82	0.82	0.79	0.79
Sensitivity	0.80	0.80	0.81	0.81	0.76	0.77
Specificity	0.91	0.85	0.88	0.83	0.90	0.85
False Positive Rate	0.09	0.15	0.12	0.17	0.10	0.15
False Negative Rate	0.20	0.20	0.19	0.19	0.24	0.23
Positive Predictive Power	0.97	0.97	0.97	0.94	0.96	0.96
Negative Predictive Power	0.50	0.43	0.46	0.57	0.51	0.44

Table 75.*Classification Accuracy Demographics Winter 2024: Math with SB*

Sample Demographics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	646	609	633	469	530	542
Female	593	564	598	451	506	481
Gender Unknown	0	0	3	1	3	9
American Indian/Alaska Native	18	31	13	6	16	16
Asian/Pacific Islander	93	72	80	56	75	78
Black	106	111	130	82	106	85
Multi-Racial	239	247	234	178	197	200
Native Hawaiian/Pacific Islander	38	31	47	27	26	41
White	745	681	730	572	619	612
Hispanic/Latino	274	268	290	203	229	222
Not Hispanic/Latino	965	905	944	718	810	810
IEP or diagnosed disability	239	201	202	137	160	155
English Language Learner	107	79	67	46	54	35

Table 76.*Classification Accuracy Spring 2024: Math Composite with SB*

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Cut Points - Percentile rank on criterion measure	20	20	20	20	20	20
Cut Points - Performance score on criterion measure	2362	2397	2410	2423	2427	2434
Cut Points - Corresponding performance score (numeric) on screener measure	-0.33328	-0.26693	-0.76846	-0.43827	-0.14227	-0.53191
Classification Data - True Positive (a)	892	815	953	659	580	667
Classification Data - False Positive (b)	42	13	37	38	16	18

Evidence	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Classification Data - False Negative (c)	141	201	99	106	234	135
Classification Data - True Negative (d)	194	169	161	208	207	129
Area Under the Curve (AUC)	0.92	0.94	0.94	0.92	0.88	0.92
AUC Estimate's 95% Confidence Interval: Lower Bound	0.92	0.94	0.94	0.92	0.88	0.92
AUC Estimate's 95% Confidence Interval: Upper Bound	0.92	0.94	0.94	0.92	0.88	0.92

Table 77.

Classification Accuracy Summary Spring 2024: Math Composite with SB

Statistics	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Sample Size	1269	1198	1250	1011	1037	949
Base Rate	0.19	0.15	0.16	0.24	0.22	0.15
Overall Classification Rate	0.86	0.82	0.89	0.86	0.76	0.84
Sensitivity	0.86	0.80	0.91	0.86	0.71	0.83
Specificity	0.82	0.93	0.81	0.85	0.93	0.88
False Positive Rate	0.18	0.07	0.19	0.15	0.07	0.12
False Negative Rate	0.14	0.20	0.09	0.14	0.29	0.17
Positive Predictive Power	0.96	0.98	0.96	0.95	0.97	0.97
Negative Predictive Power	0.58	0.46	0.62	0.66	0.47	0.49

Table 78.

Classification Accuracy Demographics Spring 2024: Math Composite with SB

Sample	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Male	664	620	637	517	540	501
Female	604	578	610	493	493	438
Gender Unknown	1	0	3	1	4	10
American Indian/Alaska Native	19	31	13	10	18	12
Asian/Pacific Islander	98	73	79	62	78	74
Black	111	116	135	100	106	78
Multi-Racial	244	250	238	189	196	174
Native Hawaiian/Pacific Islander	39	31	47	32	25	34
White	758	697	738	618	614	577
Hispanic/Latino	284	270	297	237	235	199
Not Hispanic/Latino	985	928	953	774	802	750

Sample	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
IEP or diagnosed disability	247	209	199	150	147	145
English Language Learner	113	80	74	64	54	29

Conclusion: NCII Classification Accuracy

Each criterion measure in NCII is first evaluated if (a) the criterion measure is appropriate, (b) risk defined within an RTI approach to screening, and (c) classification analyses and cut points are adequately performed? In this Technical Report, the answer is yes to all three questions:

- The criterion measure is Smarter Balanced (SB).
- Risk is defined as the 20th percentile rank.
- Classification analyses and cut points were adequately performed with results on true and false positives and negatives reported for each measure, grade, and season.

According to the NCII website, “Area Under the Curve (AUC) is an overall indication of the diagnostic accuracy of a Receiver Operating Characteristic (ROC) curve. ROC curves are a generalization of the set of potential combinations of sensitivity and specificity possible for predictors. AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory reading performance, whereas values at 0.50 indicate the predictor is no better than chance.”





“Full Bubble: All of Q1 – Q3 (below) rated as YES and the lower bound of the confidence interval around the Area Under the Curve (AUC) estimate ≥ 0.80 and Sensitivity ≥ 0.80 and Specificity ≥ 0.80 .

Half Bubble: Q1-Q3 rated as YES (below) and either (a) the lower bound of the confidence interval around the AUC estimate ≥ 0.70 but < 0.80 or (b) Sensitivity ≥ 0.70 and Specificity ≥ 0.70 .

Empty Bubble: Does not meet full or half bubble.”











































Area Under the Curve (AUC) Statistic: an overall indication of the diagnostic accuracy of a Receiver Operating Characteristic (ROC) curve. ROC curves are a generalization of the set of potential combinations of sensitivity and specificity possible for predictors. AUC values closer to 1 indicate the screening measure reliably distinguishes among students with satisfactory and unsatisfactory reading performance, whereas values at 0.50 indicate the predictor is no better than chance.

In replicating the chart from NCII, we used their symbols to reflect ratings and criteria.

Rating	Criteria
	Convincing evidence
	Partially convincing evidence
	Unconvincing evidence
	Data unavailable

In applying these criteria, however, it is important to note that the previous tables displaying values reflect the following caveats:

- For AUC, all but a very few of the values were **well above .80**.
- For sensitivity and specificity, all but a very few were **well above .70s and many >.80**.
- Base rates ranged from **.15** (minimum) to **.24** (maximum) and averaged .18, .19, or .20
- Overall classification values were typically **above .70 or .75** and often in the **.80 range**.
- False positive and false negative values were typically in the low **.10 to .20** range.
- The positive predictive power was most often in the mid to **high .90s** which indicates that easyCBM is very accurate in predicting risk at the 20th PR.

READING MEASURES	Grade	Fall	Winter	Spring
<i>Proficient Reading Benchmark</i>	3			
<i>Proficient Reading Benchmark</i>	4			
<i>Proficient Reading Benchmark</i>	5			
<i>Proficient Reading Benchmark</i>	6			
<i>Proficient Reading Benchmark</i>	7			
<i>Proficient Reading Benchmark</i>	8			
<i>Oral Reading Fluency</i>	3			
<i>Oral Reading Fluency</i>	4			
<i>Oral Reading Fluency</i>	5			
<i>Oral Reading Fluency</i>	6			
<i>Oral Reading Fluency</i>	7			
<i>Oral Reading Fluency</i>	8			
<i>Vocabulary</i>	3			
<i>Vocabulary</i>	4			

READING MEASURES	Grade	Fall	Winter	Spring
<i>Vocabulary</i>	5			
<i>Vocabulary</i>	6			
<i>Vocabulary</i>	7			
<i>Vocabulary</i>	8			

COMPOSITE READING MEASURES	Fall	Winter	Spring
Proficient • Vocabulary • Fluency – Grade 3			
Proficient • Vocabulary • Fluency – Grade 4			
Proficient • Vocabulary • Fluency – Grade 5			
Proficient • Vocabulary • Fluency – Grade 6			
Proficient • Vocabulary • Fluency – Grade 7			
Proficient • Vocabulary • Fluency – Grade 8			

MATH MEASURES	Grade	Fall	Winter	Spring
<i>Proficient Math</i>	3			
<i>Proficient Math</i>	4			
<i>Proficient Math</i>	5			
<i>Proficient Math</i>	6			
<i>Proficient Math</i>	7			
<i>Proficient Math</i>	8			

COMPOSITE MATH	Fall	Winter	Spring
Grade 3			
Grade 4			
Grade 5			
Grade 6			
Grade 7			
Grade 8			

References

- Alonzo, J., & Anderson, D. (2018). *Classification accuracy of the easyCBM Kindergarten – Grade 2 reading measures (Technical Report # 1801)*. Retrieved from Eugene, OR: Behavioral Research and Teaching, University of Oregon.
- Alonzo, J. A., D. (2018). *Supplementary report on easyCBM Grade K – 2 literacy measures: Classification accuracy and predictive validity, a follow-up to Technical Report 1403 (Technical Report # 1805)*. Retrieved from Eugene, OR: Behavioral Research and Teaching, University of Oregon.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*.
- Anderson, D., Alonzo, J., Tindal, G., Farley, D., Irvin, P. S., Lai, C. F., . . . Wray, K. A. (2014). *Technical manual: easyCBM (Technical Report # 1408)*. Eugene, OR: Behavioral Research and Teaching, University of Oregon
- Clarke, B., Nese, J. F. T., Alonzo, J., Mercier-Smith, J. L. M., Tindal, G., Kame' enui, E. J., Baker, S. K. (2011). Classification accuracy of easyCBM first grade mathematics measures: Findings and implications for the field. *Assessment for Effective Intervention*, 36(4), 243-255. doi:10.1177/1534508411414153