Balancing on Three Legs: The Tension Between Aligning to Standards, Predicting High-Stakes Outcomes, and Being Sensitive to Growth

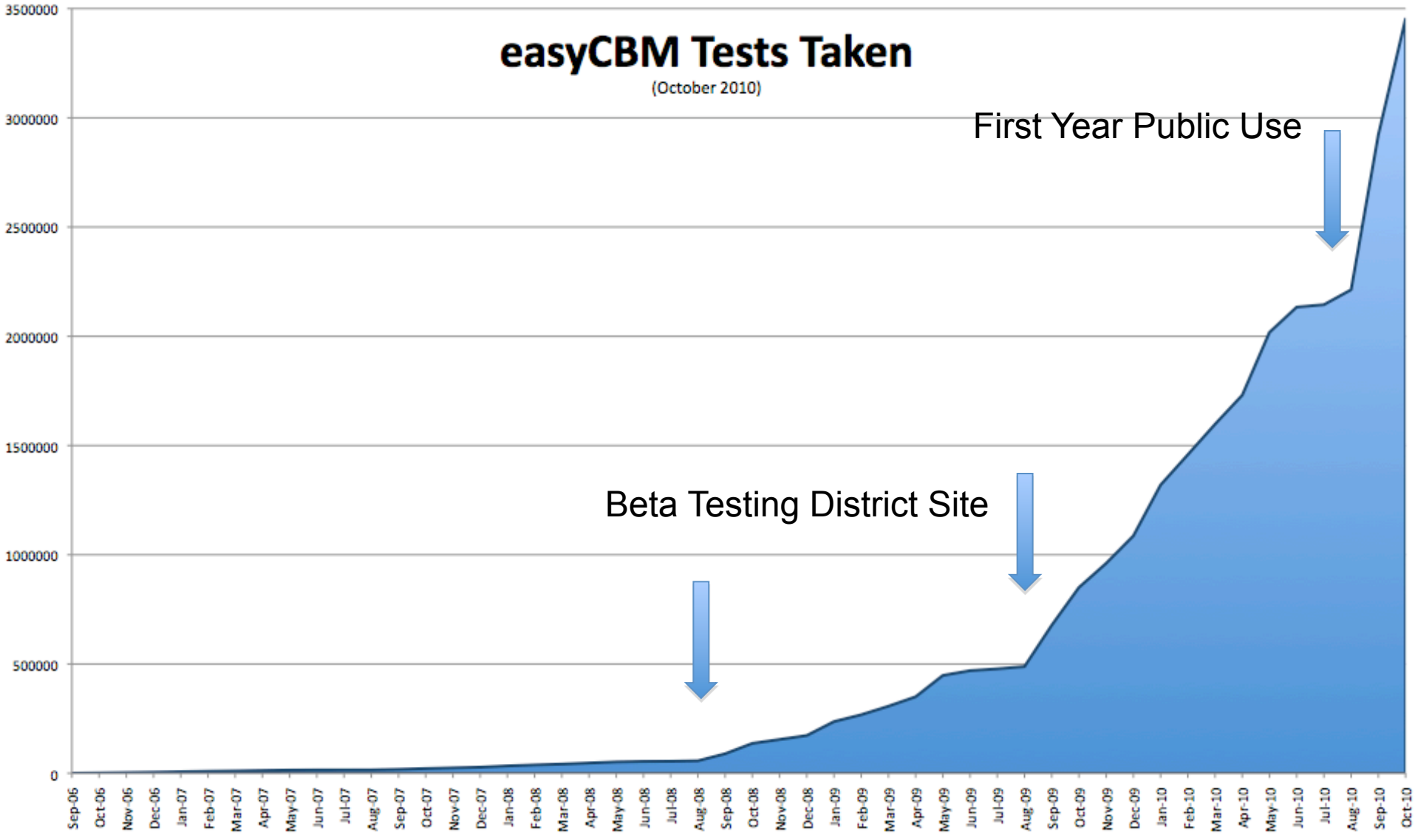
Julie Alonzo, Joe Nese, Leilani Saez & Gerald Tindal
Behavioral Research and Teaching
College of Education – UO

BRT
behavioral research & teaching

Assessment System

- easyCBM (2006)
  - Online benchmarking (45 item) and progress monitoring (16 item) assessment system
  - K-8
  - Mathematics and Reading
- General outcome measures, based on NCTM Focal Point Standards in Mathematics
- Philosophically, we see RTI as focusing on improvement of instruction.

Growth in Test Use



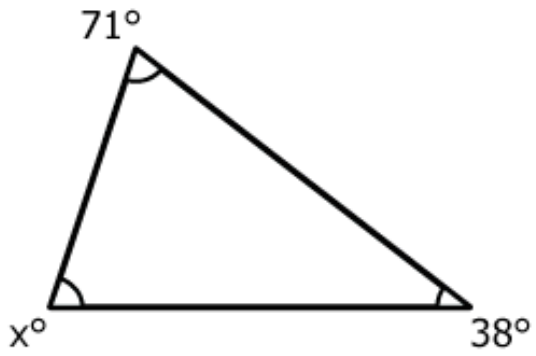
Student – Taking Test

http://www.easycbm.com/cbms/math_take.php?tid=382&gid=&sid=demo&cbmid=396&cbm_type

Google

easy CBM

DEMO STUDENT, #4 of 16



x = ___

- 16°
- 38°
- 71°

Back

Next

Alignment to Standards

National Council of Teachers of Mathematics (NCTM) Three Curriculum Focal Points and Objectives Grade 6

	Number & Operations: Developing an understanding of and fluency with multiplication and division of fractions and decimals.	**Algebra**: Writing, interpreting, and using mathematical expressions and equations.	**Number/Operations/Ratios**: Connecting ratio and rate to multiplication and division.
Objective 1	*Develop and use strategies to estimate the result of decimal and fraction computations & judge the reasonableness of results.* (16)	*Use order of operations to simplify expressions, including exponents and grouping symbols.* (22)	*Determine simple probabilities, both experimental and theoretical.* (35)
Objective 2	*Order, model, and compare fraction and decimals.* (15)	Construct and analyze tables (e.g., to show quantities that are in equivalent ratios), and they use equations to describe simple relationships (such as 3x = y) shown in a table. (15)	Extend whole number multiplication and division to ratios and rates. (49)
Objective 3	[Use the commutative, associative, and distributive properties to show that two expressions are equivalent.] (19)	Identify and represent equivalent expressions. (18)	Solve a wide variety of problems involving ratios and rates. (44)
Objective 4	Multiply and divide fractions and decimals to solve problems, including multistep problems and problems involving measurement. (19)	Know that the solutions of an equation are the values of the variables that make the equation true. (16)	Use simple reasoning about multiplication and division to solve ratio and rate problems. (0)
Objective 5	Use common procedures to multiply and divide fractions and decimals efficiently and accurately. (41)	Solve simple one-step equations by using number sense, properties of operations and the idea of maintaining equality on both sides of an equation. (19)	Expand the repertoire of problems that they can solve by using multiplication and division, and build on understanding of fractions to understand ratios. (0)
Objective 6	Use the meanings of fractions, multiplication and division, and the inverse relationship between multiplication and division to make sense of procedures for multiplying and dividing fractions and explain why they work. (18)	Understand that variables represent numbers whose exact values are not yet specified, and use variables appropriately. (19)	
Objective 7	Use the relationship between decimals and fractions, as well as the relationship between finite decimals and whole numbers (i.e., a finite decimal multiplied by an appropriate power of 10 is a whole number), to understand and explain the procedures for multiplying and dividing decimals. (0)	Write mathematical expressions and equations that correspond to given situations, evaluate expressions, and use expressions and formulas to solve problems. (19)	
Objective 8		Understand that expressions in different forms can be equivalent, and rewrite an expression to represent a quantity in a different way. (0)	

Note. The gray objectives are NCTM Focal Points that did not overlap with any objectives from the Oregon standards, so no items were written to these gray objectives. Those objectives in italics are exclusive to the state of Oregon. Those objectives in brackets are "Connections to Focal Points," as described by NCTM. Numbers inside the parentheses indicates the total number of items written to that focal point objective.

Alignment to Standards

- **Categorical Concurrence**: the degree to which the assessment covers the content of each standard (Webb, 2002).
  - In our analysis, the number of items that aligned to the content for each focal point objective, or "hits."

Categorical Concurrence, Items Aligned with Objectives – Algebra

	Benchmarks % (Frequency)						Progress Monitoring % (Frequency)									
	Fall		Winter		Spring		Form 1		Form 3		Form 5		Form 7		Form 9	
Raters	I & L		I & L		I & L		I & L		I & L		I & L		I & L		I & L	
Not aligned (0)	0	(0)	0	(0)	0	(0)	0	(0)	0	(0)	9	(3)	13	(4)	25	(8)
Vaguely aligned (1)	3	(1)	0	(0)	3	(1)	9	(3)	3	(1)	19	(6)	0	(0)	9	(3)
Somewhat aligned (2)	9	(3)	6	(2)	3	(1)	0	(0)	3	(1)	9	(3)	0	(0)	3	(1)
Directly aligned (3)	88	(28)	94	(30)	94	(30)	91	(29)	94	(30)	63	(20)	88	(28)	63	(20)
Total aligned ratings	97	(31)	100	(32)	97	(31)	91	(29)	97	(31)	72	(23)	88	(28)	66	(21)
Hit items	94	(15)	100	(16)	94	(15)	81	(13)	94	(15)	50	(8)	88	(14)	56	(9)
Group total			96 (46)								74 (59)					

Note. For an item to be considered aligned, both raters had to give the item a rating of 2 or 3; if either rater judged the item to be not aligned (rating of 0 or 1), the item as a whole was deemed unaligned.

BRT
behavioral research & teaching

Alignment to Standards

- **Depth of knowledge (DOK)** consistency indicates the DOK required by the standards and assessments.

Results of Depth of Knowledge Analysis

Form	% of item DOK ratings *below* standard DOK rating	% of item DOK ratings *at* standard DOK rating	% of item DOK ratings *above* standard DOK rating	Hit items (of 48)	Hits across raters (of 144)	% of item DOK rating agreement between raters	Frequency of item DOK ratings		
							1	2	3
Fall	28	48	27	46	92	27	29	49	14
Winter	24	54	21	45	90	22	35	45	10
Spring	20	54	26	45	90	26	23	49	18
PM_1	23	68	10	42	84	23	36	44	4
PM_3	19	62	20	43	86	22	34	47	5
PM_5	26	53	21	38	76	16	25	43	8
PM_7	24	54	23	42	84	15	32	43	9
PM_9	22	60	18	39	78	18	28	44	6

Note. PM = progress monitoring.

Alignment to Standards

- **Range of Knowledge (ROK):** the number of objectives within the focal point for which there is at least one related item.
- **Balance of Representation (BOI):** the degree to which one objective is given more emphasis on the assessment than another.

Range of Knowledge Correspondence and Balance of Representation Index

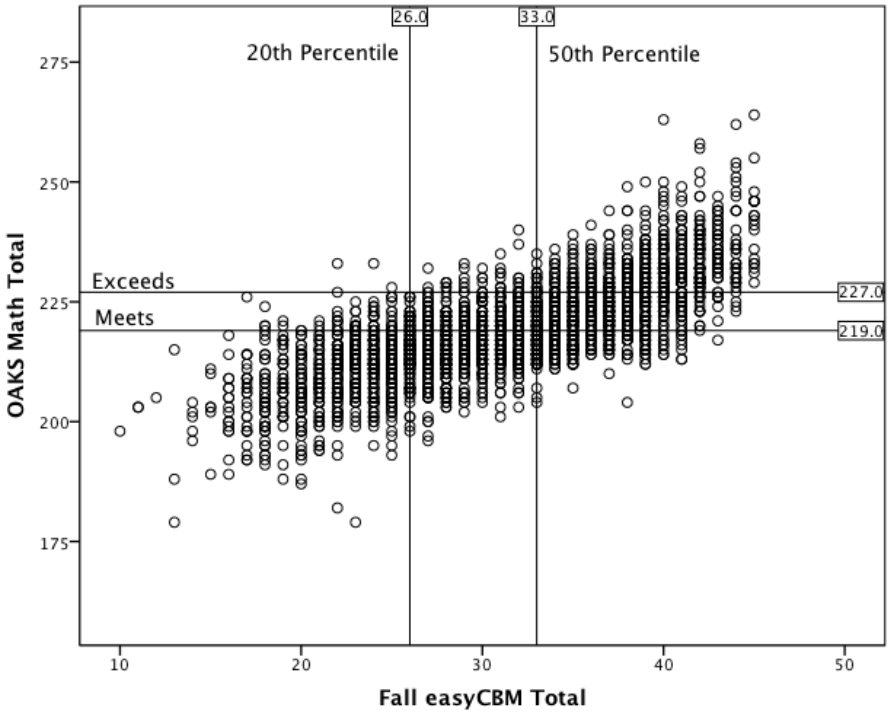
	Number & Operations (6)		Algebra (7)		Numbers/ Operations/Ratios (3)	
Forms	ROK %	BOI	ROK %	BOI	ROK %	BOI
Fall	83[a]	0.69	86[d]	0.76	100	0.85
Winter	100	0.81	100	0.72	100	0.90
Spring	100	0.85	100	0.72	100	0.90
PM_1	100	0.71	100	0.85	100	0.79
PM_3	83[b]	0.79	86[e]	0.74	100	0.85
PM_5	100	0.79	86[f]	0.78	100	0.73
PM_7	83[c]	0.75	100	0.80	100	0.90
PM_9	100	0.79	100	0.79	100	0.83

Note. PM = progress monitoring. ROK = Range of knowledge. BI = Balance of representation index. Each superscript represents an objective that was not measured. The numbers inside the parentheses indicates the number of objectives for each focal point.

[a] Objective 3. [b] Objective 4. [c] Objective 6. [d] Objective 2. [e] Objective 4. [f] Objective 3.

Ability to Predict Statewide Assessments (Math)

Oregon (grade 4 fall BM)

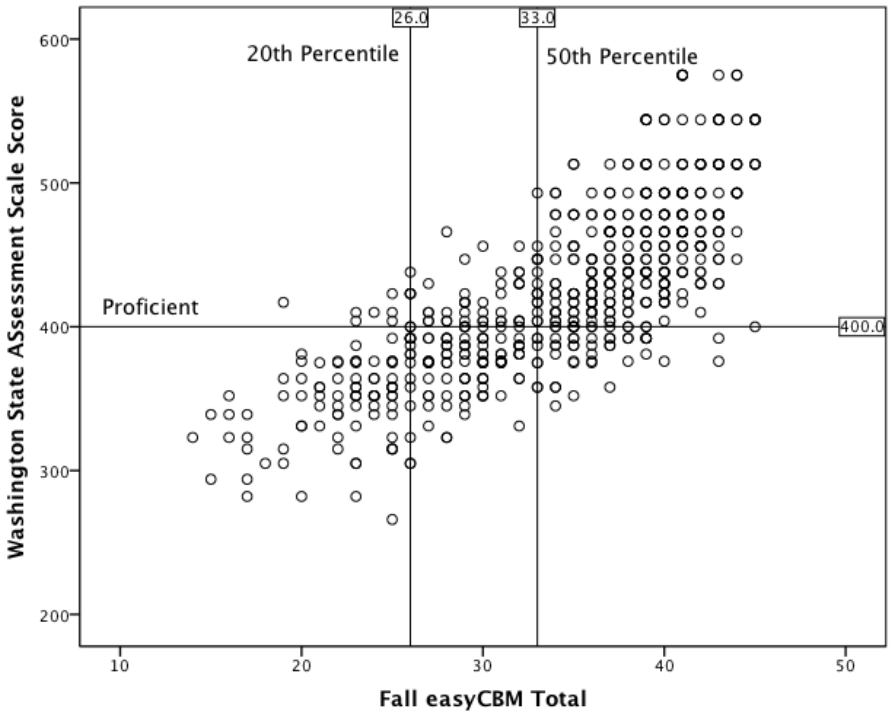


Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.750[a]	.562	.562	6.538

a. Predictors: (Constant), fall_tot

Washington (grade 4 fall BM)



Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.780[a]	.609	.608	35.991

a. Predictors: (Constant), Fall09TotMath

Ability to Predict Statewide Assessments (Math)

Table 1

Demographics

Grade	*n*	% ELL	% FRL	% SPED	% Female	% Ethnicity						
						Amer Ind	Asian/Pac Islander	Black	Hispanic	White	Multi	Decline/ Missing
District 1												
3	1023	3.1	45.2	12.7	48.5	2.8	10.9	5.2	8.7	57.9	11.9	2.5
4	993	2.9	43.1	11.7	48.8	2.1	9.4	5.5	9.4	57.5	13.9	2.2
5	1000	2.9	39.7	15.1	42.6	1.9	10.8	5.3	7.8	57.3	14.7	2.2
6	940	2.1	40.1	11.6	49.1	3.2	10.0	5.5	8.9	59.0	10.9	2.4
7	982	2.0	38.9	13.1	48.8	2.3	10.3	9.0	9.6	58.5	6.2	4.2
8	1107	2.3	34.3	10.3	41.9	3.0	13.6	9.8	11.1	60.7	1.0	0.8
District 2												
3	271	12.2	-	13.7	47.2	5.5	4.1	1.1	24.0	61.3	2.6	1.5
4	262	8.4	-	18.7	48.5	4.2	2.7	0.4	22.9	67.6	2.3	-
5	258	6.2	-	21.3	57.8	7.8	3.5	1.2	20.9	65.5	0.4	0.8
6	245	4.9	-	7.8	49.0	5.3	1.6	1.6	18.4	70.2	2.4	0.4
7	225	4.4	-	4.9	49.3	6.7	1.8	1.3	17.3	70.2	0.9	1.8
8	592	3.4	-	12.5	47.6	7.4	2.0	1.7	14.9	71.6	1.0	1.4
District 3												
3	638	6.1	29.5	15.5	49.2	0.9	16.8	6.7	7.2	56.4	11.9	-
4	673	5.6	27.0	15.5	44.9	1.0	18.1	6.7	4.5	59.0	10.7	-
5	638	5.2	27.9	14.6	45.5	1.4	15.7	7.8	7.4	64.1	3.6	-
6	667	4.5	27.0	13.0	50.5	1.6	17.1	9.0	8.4	61.2	2.5	0.1
7	623	5.3	28.4	10.4	48.8	0.3	19.4	8.2	7.5	60.7	3.7	0.2
8	661	4.8	25.9	10.7	49.6	1.4	18.8	7.9	7.7	62.0	2.1	0.2

Ability to Predict Statewide Assessments (Math)

Table 4

Resulting Statistics for Each Chosen Cut Score

Measure	Meeting score	Sensitivity	Specificity	Positive Predictive Power	Negative Predictive Power	Area Under the Curve	Overall Correct Classification
				Grade 3			
Fall	31	.79	.75	.67	.85	.84	.77
Winter	35	.82	.77	.64	.89	.87	79
Spring							
				Grade 4			
Fall	33	.83	.84	.75	.89	.90	.84
Winter	36	.84	.80	.71	.89	.90	.81
Spring	39	.88	.75	.69	.91	.93	.80
Fall	33	.84	.81	.71	.91	.91	.82
Winter	37	.87	.84	.76	.91	.93	.85
Spring	42	.89	.73	.68	.91	.93	.79
				Grade 6			
Fall	31	.85	.78	.69	.90	.90	.81
Winter	33	.86	.82	.72	.91	.92	.83
Spring	37	.85	.85	.78	.90	.94	.85
				Grade 7			
Fall	29	.80	.82	.72	.88	.90	.82
Winter	29	.80	.83	.71	.89	.91	.82
Spring	34	.89	.78	.74	.91	.93	.82
				Grade 8			
Fall	31	.84	.82	.65	.93	.92	.82
Winter	34	.87	.79	.68	.92	.92	.82
Spring	34	.76	.81	.73	.84	.91	.79

Table 8

Grade 4 Fall Benchmark

Cut Score	Sensitivity	Specificity
13	1	0
14.5	1	0.005
15.5	1	0.015
16.5	1	0.03
17.5	1	0.044
18.5	1	0.049
19.5	0.997	0.064
20.5	0.997	0.099
21.5	0.997	0.128
22.5	0.997	0.167
23.5	0.989	0.251
24.5	0.986	0.3
25.5	0.978	0.394
26.5	0.962	0.468
27.5	0.948	0.517
28.5	0.937	0.576
29.5	0.907	0.64
30.5	0.891	0.724
31.5	0.866	0.768
32.5	**0.837**	**0.818**
33.5	0.79	0.847
34.5	0.738	0.882
35.5	0.676	0.916
36.5	0.61	0.936
37.5	0.529	0.961
38.5	0.452	0.97
39.5	0.362	0.985
40.5	0.264	0.99
41.5	0.177	0.99

Test forms are designed to target middle of year difficulty, with a # of items intentionally 'easy' to ensure access to the scale for students with low math ability.

Grade 4 Fall easyCBM Norms

25th percentile =

40th percentile =

50th percentile =

n =

Table 9

Grade 4 Winter Benchmark

Cut Score	Sensitivity	Specificity
16	1	0
17.5	1	0.01
18.5	1	0.03
19.5	1	0.039
20.5	1	0.049
21.5	1	0.054
22.5	1	0.079
23.5	1	0.103
24.5	0.997	0.143
25.5	0.995	0.177
26.5	0.986	0.217
27.5	0.984	0.256
28.5	0.984	0.32
29.5	0.975	0.399
30.5	0.948	0.463
31.5	0.94	0.562
32.5	0.924	0.611
33.5	0.888	0.695
34.5	0.856	0.764
35.5	**0.801**	**0.837**
36.5	0.747	0.897
37.5	0.673	0.911
38.5	0.597	0.956
39.5	0.48	0.97
40.5	0.392	0.985
41.5	0.289	0.995
42.5	0.183	0.995
43.5	0.112	1
44.5	0.041	1
46	0	1

Grade 4 Winter easyCBM Norms

25th percentile =

40th percentile =

50th percentile =

n =

Table 10

Grade 4 Spring Benchmark

Cut Score	Sensitivity	Specificity
9	1	0
11.5	1	0.005
14	1	0.01
15.5	1	0.015
16.5	1	0.02
18	1	0.025
19.5	1	0.03
20.5	1	0.034
21.5	1	0.059
22.5	1	0.074
23.5	1	0.084
24.5	1	0.094
25.5	1	0.128
26.5	0.997	0.158
27.5	0.997	0.197
28.5	0.997	0.241
29.5	0.992	0.291
30.5	0.989	0.335
31.5	0.981	0.399
32.5	0.967	0.458
33.5	0.967	0.552
34.5	0.954	0.626
35.5	0.932	0.7
36.5	0.918	0.749
37.5	0.88	0.823
38.5	**0.82**	**0.867**
39.5	0.76	0.931
40.5	0.689	0.97
41.5	0.594	0.99
42.5	0.471	0.995
43.5	0.283	1
44.5	0.128	1
46	0	1

By the spring, higher-performing students have reached a ceiling.

Grade 4 Fall easyCBM Norms

25th percentile =

40th percentile =

50th percentile =

n =

Ability to Predict TerraNova Math, Grade 1

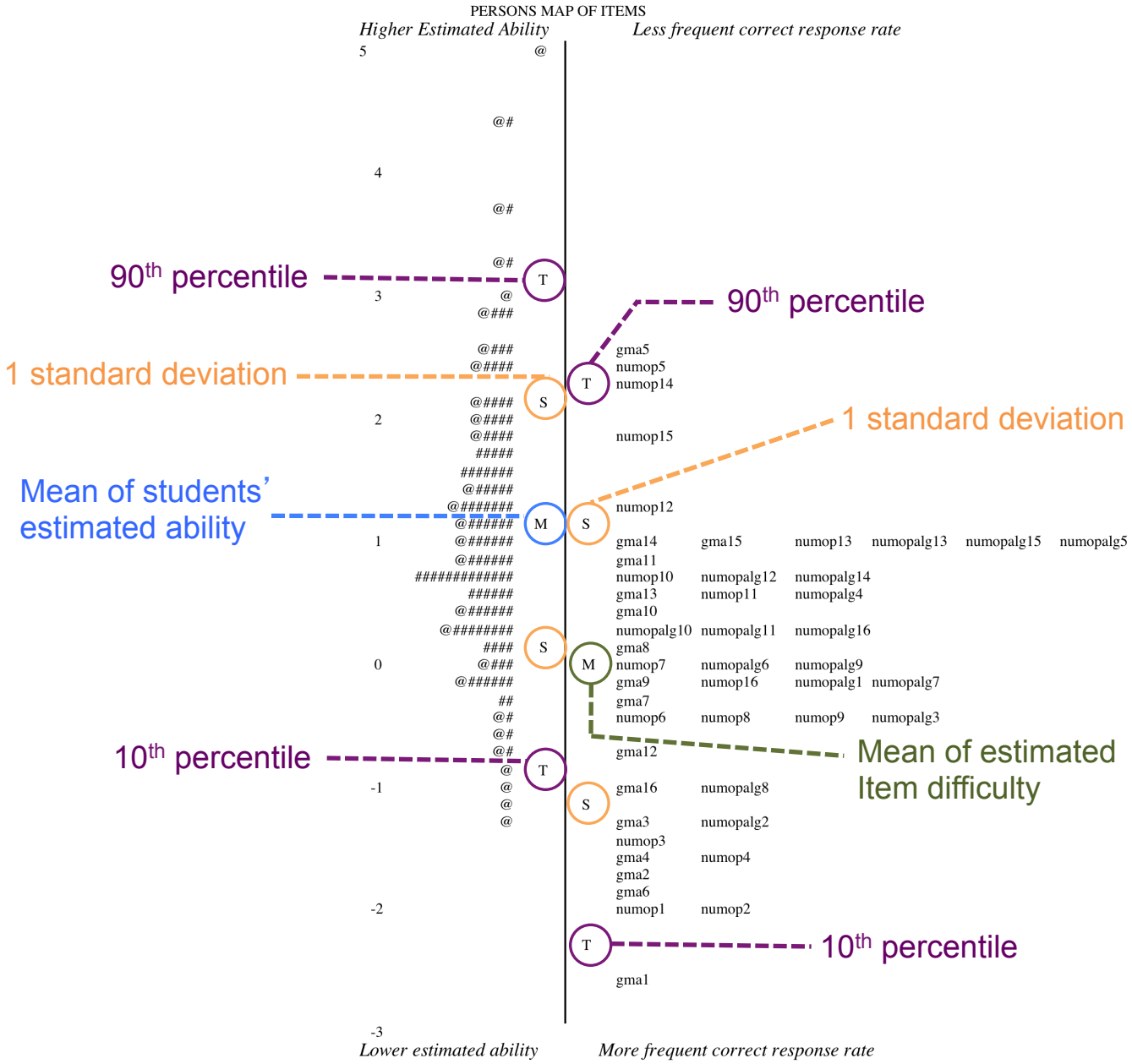
		TeraNova Math, 25th%ile Cut		
Season	easyCBM Cut Score	Sensitivity	Specificity	Overall Correct Classification
Fall	23	.70	.80	.78
	24[a]	.83	.74	.75
	25[bc]	.91	.69	.72
Spring	34	.78	.83	.82
	35[a]	.87	.77	.79
	36[bc]	.91	.74	.77

		TeraNova Math, 40th%ile Cut		
Season	easyCBM Cut Score	Sensitivity	Specificity	Overall Correct Classification
Fall	25[ac]	.73	.73	.73
	26	.77	.65	.69
	27	.82	.63	.69
	28	.89	.52	.63
	29[b]	.93	.47	.61
Spring	35	.73	.84	.81
	36	.77	.81	.80
	37[a]	.82	.75	.77
	38	.86	.68	.74
	39	.89	.61	.70
	40[b]	.93	.49	.62

[a] Cut score if following criteria proposed by Silberglitt and Hintze (2005).
[b] Cut score if the selection criteria was to sensitivity > .90.
[c] Highest sum, sensitivity and specificity.

Sensitivity to Measuring Growth for Low-Performing Students

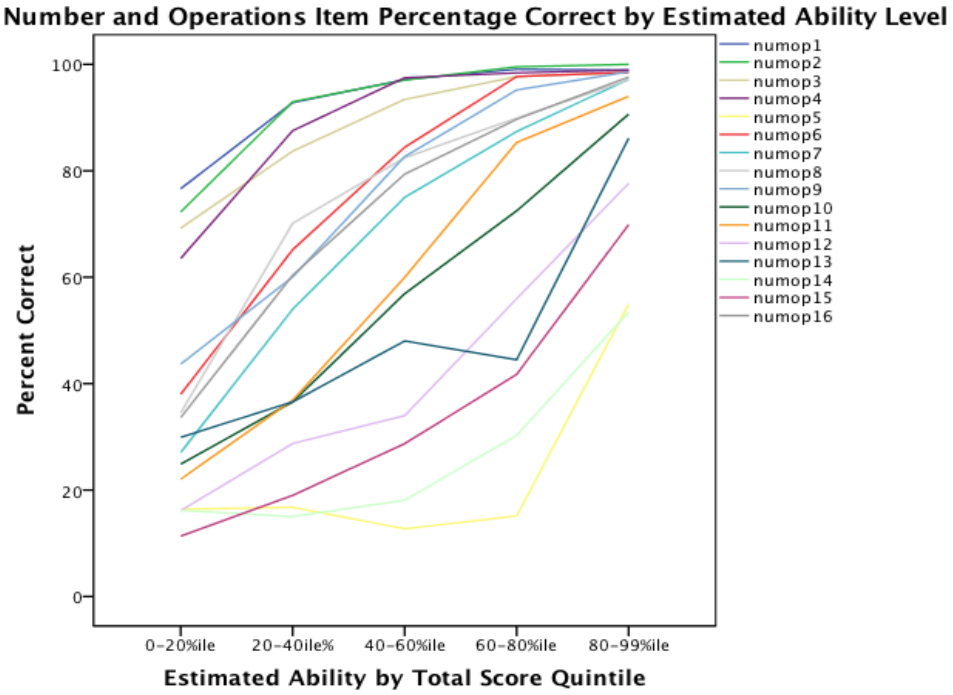
BRT
behavioral research & teaching



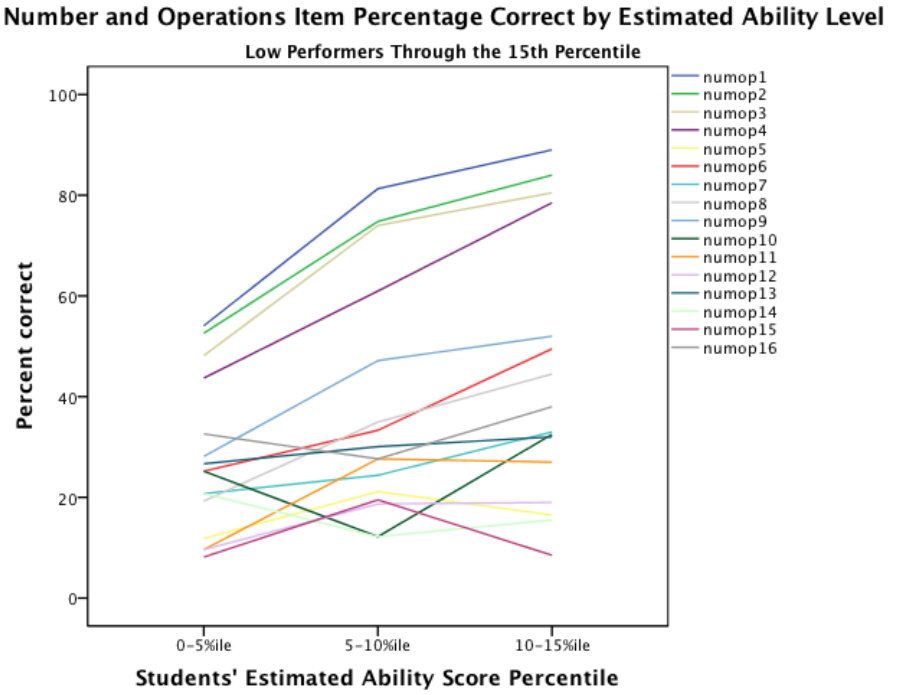
Note. Each '#' represents 16 students, each '@' represents 1 student.

Sensitivity to Measuring Growth for Low-Performing Students (5th grade)

All Students



Low Performers through the 15th Percentile



BRT
behavioral research & teaching

Key Take-Aways

- As student performance on CBMs becomes increasingly high-stakes, our attention to technical adequacy needs to increase commensurately
- Too great a focus on any one of the legs (alignment, ability to predict state test performance, sensitivity to measuring growth for low-performing students) puts the whole at risk.

For More Information

http://www.brtprojects.org

http://easyCBM.com

BRT
behavioral research & teaching