
New Mexico Alignment Study of Short-Cycle Diagnostic Assessment Systems

Report to the New Mexico
Legislative Education Study Committee
August 19, 2009

Stanley Rabinowitz, Ph.D.
Senior Program Director
Assessment and Standards Development Services (ASDS)

Agenda

- Introduction
- Purpose
- Methodology
- Findings
- Recommendations

Introduction

- Many state, district, or local assessment systems now use interim or short-cycle assessments to measure progress toward proficiency on state content standards.
- Alignment is critical for these assessments if valid inferences are to be made from them about students' mastery of standards and preparation for accountability assessments.

Introduction

- Short-cycle assessments are considered a class of formative assessments
- Short-cycle assessments should not duplicate state tests; should provide additional timely information
- Professional Development essential for appropriate use of information from short-cycle assessments
- Evidence of technical adequacy (reliability, validity, bias-free) essential for short-cycle assessments—**alignment**
- Balanced Assessment System (state and local assessments)

Introduction

- New Mexico's current assessment system includes:
 - Assessment for accountability at Grades 8 and 11 (HS exit exam)
 - District-selected short-cycle diagnostic assessments at Grades 9 and 10

Purpose

- The Assessment and Standards Development Services (ASDS) program at WestEd was contracted by NMPED to conduct a study to determine the alignment between the Grades 9 and 11 Standards-Based Assessments (SBA) test blueprints in reading and mathematics with six short-cycle diagnostic assessment systems.
- This study (conducted in two phases) will help NMPED determine the validity of the short-cycle diagnostic assessments in assessing student progress toward proficiency on the Grade 11 SBA.

Purpose

- Short-cycle diagnostic assessments included in this study:
 - Phase I
 - Riverside (Assess2Learn)
 - NWEA (MAP)
 - Phase II
 - Pearson (Learnia)
 - Scholastic
 - Discovery
 - ACT (COMPASS)

Purpose

What was aligned?

Short-cycle assessments (Grades 9 and 10)



NM Standards-Based Assessment (SBA)
blueprint for reading and mathematics
(Grades 9 and 11)

Purpose

Study was designed to answer two key questions:

1. To what degree does each test item align for content and depth to the NM SBA test blueprints?
 2. To what degree do the assessments represent the breadth and range of knowledge of the NM SBA test blueprints?
- Purpose NOT to find best test—we are only looking at alignment; quality includes many other features

Methodology

- Alignment approach is unique and highly collaborative
 - Draws upon work of Norm Webb
- Targets two levels of analysis
 - Item-Level Analyses
 - Categorical Concurrence
 - Depth of Knowledge
 - Test-Level Analyses
 - Range of Depth of Knowledge
 - Range of Knowledge (Comprehensiveness/Breadth)
 - Balance of Representation

Methodology

- Analyst Training, Calibration, and Rating Process
 - Review alignment protocol and SBA blueprints
 - Train analysts using a subset of items and benchmarks and performance standards
 - Calibrate analysts to ensure they can apply the criteria accurately and consistently
 - Conduct analysis at item and test level
 - Develop/apply decision rules as appropriate

Findings—Riverside (Assess2Learn)

- Riverside (Assess2Learn) Short-Cycle Assessment Sample
 - English Language Arts
 - Grade 9: 113 items (3 forms)
 - Grade 10: 104 items (3 forms)
 - Mathematics
 - Algebra I: 112 items (3 forms)
 - Geometry: 105 items (3 forms)
 - Algebra II: 102 items (3 forms)

Findings—Riverside (Assess2Learn)

Overall Item Alignment

Reading (Grades 9 and 10)

- 43% (93) of the 217 items were aligned to at least one benchmark.

Mathematics (Algebra I, Geometry, and Algebra II)

- 66% (211) of the 319 items were aligned to at least one benchmark.
 - 81% (91) of the 112 **Algebra I** items were aligned to at least one Grade 9 benchmark.
 - 67% (70) of the 105 **Geometry** items were aligned to at least one Grade 11 benchmark.
 - 49% (50) of the 102 **Algebra II** items were aligned to at least one Grade 11 benchmark.

Findings—Riverside (Assess2Learn)

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No.	%	No.	%	No.	%	No.
Language Arts Grade 9	44	5	2	95	42	0	0	0	0
Language Arts Grade 10	52	0	0	65	34	35	18	0	0
Mathematics Algebra I	104	39	41	54	56	7	7	0	0
Mathematics Geometry	86	22	19	73	63	5	4	0	0
Mathematics Algebra II	60	18	11	82	49	0	0	0	0

Findings—Riverside (Assess2Learn)

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Forms I, II & III (n = 113)	
Standard	SBA Item Ct	SBA % of Points	A2L Item Ct	A2L % of Points
I. Reading and Listening for Comprehension	32	65%	23	20%
II. Writing and Speaking for Expression	N/A	N/A	17	15%
III. Literature and Media	17	35%	1	1%
No Relationship	N/A	N/A	72	64%
Reading	Grade 11 SBA Blueprint (n = 51)		Forms I, II & III (n = 104)	
Standard	SBA Item Ct	SBA % of Points	A2L Item Ct	A2L % of Points
I. Reading and Listening for Comprehension	35	65%	5	15%
III. Literature and Media	16	35%	6	18%
No Relationship	N/A	N/A	22	67%

Findings—Riverside (Assess2Learn)

Balance of Representation (by subject and grade with test forms totaled)

Math (Algebra I)	Grade 9 SBA Blueprint (n = 60)		Forms I, II & III (n = 112)	
Strand	SBA Item Ct	SBA % of Points	A2L Item Ct	A2L % of Points
Algebra, Functions, and Graphs	24	40%	80	71%
Geometry and Trigonometry	18	30%	2	2%
Data Analysis and Probability	18	30%	9	8%
No Relationship	N/A	N/A	21	19%
Math (Geometry and Algebra II)	Grade 11 SBA Blueprint (n = 60)		Forms I, II & III (n = 206)	
Strand	SBA Item Ct	SBA % of Points	A2L Item Ct	A2L % of Points
Algebra, Functions, and Graphs	23	36%	49	24%
Geometry and Trigonometry	23	39%	69	33%
Data Analysis and Probability	14	24%	2	1%
No Relationship	N/A	N/A	87	42%

Findings—NWEA (MAP)

- NWEA (MAP) Short-Cycle Assessment Sample
 - Reading
 - Grade 9: 120 items (low, mid, high performing forms)
 - Grade 10: 120 items (low, mid, high performing forms)
 - Mathematics
 - Grade 9: 150 items (low, mid, high performing forms)
 - Grade 10: 150 items (low, mid, high performing forms)

Findings—NWEA (MAP)

Overall Item Alignment

Reading (Grades 9 and 10)

- 50% (121) of the 240 items were aligned to at least one benchmark.

Mathematics (Grades 9 and 10)

- 92% (276) of the 300 items were aligned to at least one benchmark.

Findings—NWEA (MAP)

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No.	%	No.	%	No.	%	No.
Reading Grade 9	60	5	3	58	35	37	22	0	0
Reading Grade 10	72	1	1	44	32	55	39	0	0
Mathematics Grade 9	169	52	88	44	75	4	6	0	0
Mathematics Grade 10	162	52	84	45	73	3	5	0	0

Findings—NWEA (MAP)

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Low, Mid, and High Performing Forms (n = 120)	
	Standard	SBA Item Ct	SBA % of Points	MAP Item Ct
I. Reading and Listening for Comprehension	32	65%	54	45%
III. Literature and Media	17	35%	1	1%
No Relationship	N/A	N/A	65	54%
Reading	Grade 11 SBA Blueprint (n = 51)		Low, Mid, and High Performing Forms (n = 120)	
	Standard	SBA Item Ct	SBA % of Points	MAP Item Ct
I. Reading and Listening for Comprehension	35	65%	50	42%
III. Literature and Media	16	35%	16	13%
No Relationship	N/A	N/A	54	45%

Findings—NWEA (MAP)

Balance of Representation (by subject and grade with test forms totaled)

Math	Grade 9 SBA Blueprint (n = 60)		Low, Mid, and High Performing Forms (n = 150)	
Strand	SBA Item Ct	SBA % of Points	MAP Item Ct	MAP % of Points
Algebra, Functions, and Graphs	24	40%	60	40%
Geometry and Trigonometry	18	30%	50	33%
Data Analysis and Probability	18	30%	29	19%
No Relationship	N/A	N/A	11	7%
Math	Grade 11 SBA Blueprint (n = 60)		Low, Mid, and High Performing Forms (n = 150)	
Strand	SBA Item Ct	SBA % of Points	MAP Item Ct	MAP % of Points
Algebra, Functions, and Graphs	23	36%	68	45%
Geometry and Trigonometry	23	39%	43	29%
Data Analysis and Probability	14	24%	26	17%
No Relationship	N/A	N/A	13	9%

Findings—Pearson (Learnia)

- Pearson (Learnia) Short-Cycle Assessment Sample
 - Reading
 - Grade 9: 80 items (2 forms)
 - Grade 10: 80 items (2 forms)
 - Mathematics
 - Algebra I: 80 items (2 forms)
 - Geometry: 78 items (2 forms)
 - Algebra II: 79 items (2 forms)

Findings—Pearson (Learnia)

Overall Item Alignment

Reading (Grades 9 and 10)

- 62% (99) of the 160 items were aligned to at least one benchmark.

Mathematics (Algebra I, Geometry, and Algebra II)

- 62% (N) of the 237 items were aligned to at least one benchmark.
 - 79% (63) of the 80 Algebra I items were aligned to at least one Grade 9 benchmark.
 - 72% (56) of the 78 Geometry items were aligned to at least one Grade 11 benchmark.
 - 35% (28) of the 79 Algebra II items were aligned to at least one Grade 11 benchmark.

Findings—Pearson (Learnia)

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No.	%	No.	%	No.	%	No.
Reading Grade 9	40	5	2	75	30	20	8	0	0
Reading Grade 10	59	12	7	73	43	15	9	0	0
Mathematics Algebra I	74	68	50	27	20	5	4	0	0
Mathematics Geometry	65	37	24	57	37	6	4	0	0
Mathematics Algebra II	31	68	21	32	10	0	0	0	0

Findings—Pearson (Learnia)

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Learnia Tests 1 and 2 (n = 80)	
Standard	SBA Item Ct	SBA % of Points	Learnia Item Ct	Learnia % of Points
I. Reading and Listening for Comprehension	32	65%	41	49%
III. Literature and Media	17	35%	0	0%
No Relationship	N/A	N/A	40	51%
Reading	Grade 11 SBA Blueprint (n = 51)		Learnia Tests 1 and 2 (n = 80)	
Standard	SBA Item Ct	SBA % of Points	Learnia Item Ct	Learnia % of Points
I. Reading and Listening for Comprehension	35	65%	41	47%
III. Literature and Media	16	35%	18	29%
No Relationship	N/A	N/A	21	24%

Findings—Pearson (Learnia)

Balance of Representation (by subject and grade with test forms totaled)

Math	Grade 9 SBA Blueprint (n = 60)		Learnia Grade 9 Forms (n = 80)	
Strand	SBA Item Ct	SBA % of Points	Learnia Item Ct	Learnia % of Points
Algebra, Functions and Graphs	24	40%	53	66%
Geometry and Trigonometry	18	30%	4	5%
Data Analysis and Probability	18	30%	6	8%
No Relationship	N/A	N/A	17	21%
Math	Grade 11 SBA Blueprint (n = 60)		Learnia Grade 10 Forms (n = 157)	
Strand	SBA Item Ct	SBA % of Points	Learnia Item Ct	Learnia % of Points
Algebra, Functions and Graphs	23	36%	32	20%
Geometry and Trigonometry	23	39%	52	33%
Data Analysis and Probability	14	24%	0	0%
No Relationship	N/A	N/A	73	46%

Findings—Scholastic

- Scholastic Short-Cycle Assessment Sample
 - Reading
 - Grade 9: 150 items (low, mid, high performing forms)
 - Grade 10: 150 items (low, mid, high performing forms)

Findings—Scholastic

Overall Item Alignment

Reading (Grades 9 and 10)

- 47% (71) of the 150 Grade 9 items were aligned to at least one benchmark.
- 52% (78) of the 150 Grade 10 items were aligned to at least one benchmark.

Findings—Scholastic

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No.	%	No.	%	No.	%	No.
Reading Grade 9	72	3%	2	94%	68	3%	2	0%	0
Reading Grade 10	88	0%	0	100%	88	0%	0	0%	0

Findings—Scholastic

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Scholastic 25th, 62nd and 95th Percentile Samples (n = 150)	
Standard	SBA Item Ct	SBA % of Points	Scholastic Item Ct	Scholastic % of Points
I. Reading and Listening for Comprehension	32	65%	71	47%
III. Literature and Media	17	35%	0	0%
No Relationship	N/A	N/A	79	53%
Reading	Grade 11 SBA Blueprint (n = 51)		Scholastic 25th, 62nd and 95th Percentile Samples (n = 150)	
Standard	SBA Item Ct	SBA % of Points	Scholastic Item Ct	Scholastic % of Points
I. Reading and Listening for Comprehension	35	65%	80	53%
III. Literature and Media	16	35%	8	5%
No Relationship	N/A	N/A	62	41%

Findings—Discovery

- Discovery Short-Cycle Assessment Sample
 - Reading
 - Grade 9: 60 items (2 forms)
 - Grade 10: 60 items (2 forms)
 - Mathematics
 - Grade 9: 60 items (2 forms)
 - Grade 10: 60 items (2 forms)

Findings—Discovery

Overall Item Alignment

Reading (Grades 9 and 10)

- 80% (96) of the 120 items were aligned to at least one benchmark.

Mathematics (Grades 9 and 10)

- 80% (96) of the 120 items were aligned to at least one benchmark.

Findings—Discovery

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No	%	No.	%	No.	%	No.
Reading Grade 9	45	2	1	71	32	27	12	0	0
Reading Grade 10	55	6	3	58	32	36	20	0	0
Mathematics Grade 9	62	34	21	63	39	3	2	0	0
Mathematics Grade 10	54	52	28	42	23	6	3	0	0

Findings—Discovery

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Discovery Tests A and B (n = 60)	
Standard	SBA Item Ct	SBA % of Points	Discovery Item Ct	Discovery % of Points
I. Reading and Listening for Comprehension	32	65%	34	57%
III. Literature and Media	17	35%	9	15%
No Relationship	N/A	N/A	17	28%
Reading	Grade 11 SBA Blueprint (n = 51)		Discovery Tests A and B (n = 60)	
Standard	SBA Item Ct	SBA % of Points	Discovery Item Ct	Discovery % of Points
I. Reading and Listening for Comprehension	35	65%	32	53%
III. Literature and Media	16	35%	21	35%
No Relationship	N/A	N/A	7	12%

Findings—Discovery

Balance of Representation (by subject and grade with test forms totaled)

Math	Grade 9 SBA Blueprint (n = 60)		Discovery Tests A and B (n = 60)	
Strand	SBA Item Ct	SBA % of Points	Discovery Item Ct	Discovery % of Points
Algebra, Functions, and Graphs	24	40%	22	37%
Geometry and Trigonometry	18	30%	16	26%
Data Analysis and Probability	18	30%	10	17%
No Relationship	N/A	N/A	12	20%
Math	Grade 11 SBA Blueprint (n = 60)		Discovery Tests A and B (n = 60)	
Strand	SBA Item Ct	SBA % of Points	Discovery Item Ct	Discovery % of Points
Algebra, Functions, and Graphs	23	36%	20	33%
Geometry and Trigonometry	23	39%	15	25%
Data Analysis and Probability	14	24%	13	22%
No Relationship	N/A	N/A	12	20%

Findings—ACT (COMPASS)

- ACT (COMPASS) Short-Cycle Assessment Sample
 - Reading
 - Grade 9: 109 items (Vocabulary/Reading Diagnostic)
 - Grade 10: 114 items (Reading Placement)
 - Mathematics
 - Grade 9: 200 items (Pre-Algebra and Algebra)
 - Grade 10: 100 items (Geometry and Trigonometry)

Findings—ACT (COMPASS)

Overall Item Alignment

Reading (Grades 9 and 10)

- **70%** (157) of the 223 items were aligned to at least one benchmark.
 - **0%** (0) of the 49 Grade 9 Vocabulary Diagnostic sample were aligned to at least one benchmark.
 - **100%** (60) of the 60 Grade 9 Reading Diagnostic sample items were aligned to at least one benchmark.
 - **85%** (97) of the 114 Grade 10 Reading Placement sample items were aligned to at least one benchmark.

Math (Grades 9 and 10)

- **62%** (187) of the 300 items were aligned to at least one benchmark.
 - **77%** (153) of the 200 Pre-Algebra and Algebra items were aligned to at least one Grade 9 benchmark.
 - **68%** (34) of the 50 Geometry items were aligned to at least one Grade 10 benchmark.
 - **0%** (0) of the 50 Trigonometry items were aligned to at least one Grade 10 benchmark.

Findings—ACT (COMPASS)

Depth of Knowledge of Aligned Items

Subject	Item Alignments	Recall		Basic Application		Strategic Thinking		Extended Thinking	
		%	No.	%	No.	%	No.	%	No.
Reading Grade 9	60	22	13	78	47	0	0	0	0
Reading Grade 10	97	21	20	78	76	1	1	0	0
Mathematics Pre-Algebra	85	55	47	45	38	0	0	0	0
Mathematics Algebra	94	56	53	42	39	2	2	0	0
Mathematics Geometry	37	11	4	86	32	3	1	0	0
Mathematics Trigonometry	0	0	0	0	0	0	0	0	0

Findings—ACT (COMPASS)

Balance of Representation (by subject and grade with test forms totaled)

Reading	Grade 9 SBA Blueprint (n = 49)		Vocabulary and Reading Diagnostic Samples (n = 109)	
Standard	SBA Item Ct	SBA % of Points	ACT Item Ct	ACT % of Points
I. Reading and Listening for Comprehension	32	65%	60	55%
III. Literature and Media	17	35%	0	0%
No Relationship	N/A	N/A	49	45%
Reading	Grade 11 SBA Blueprint (n = 51)		Reading Placement (n = 114)	
Standard	SBA Item Ct	SBA % of Points	ACT Item Ct	ACT % of Points
I. Reading and Listening for Comprehension	35	65%	84	74%
III. Literature and Media	16	35%	13	11%
No Relationship	N/A	N/A	17	15%

Findings—ACT (COMPASS)

Balance of Representation (by subject and grade with test forms totaled)

Math	Grade 9 SBA Blueprint (n = 60)		Pre-Algebra and Algebra Placement Samples (n = 200)	
Strand	SBA Item Ct	SBA % of Points	ACT Item Ct	ACT % of Points
Algebra, Functions and Graphs	24	40%	138	69%
Geometry and Trigonometry	18	30%	13	7%
Data Analysis and Probability	18	30%	2	1%
No Relationship	N/A	N/A	47	24%
Math	Grade 11 SBA Blueprint (n = 60)		Geometry and Trigonometry Domain Forms (n = 100)	
Strand	SBA Item Ct	SBA % of Points	ACT Item Ct	ACT % of Points
Algebra, Functions and Graphs	23	36%	0	0%
Geometry and Trigonometry	23	39%	34	34%
Data Analysis and Probability	14	24%	0	0%
No Relationship	N/A	N/A	66	66%

Recommendations

- **Review the intended purpose** of the assessments to determine whether they adequately meet New Mexico's needs as short-cycle diagnostic assessments
- **Review alignment of the SBA system** for purposes of ensuring progress toward proficiency on the grade 11 SBA (e.g., consider the degree to which the grade 9 and 11 blueprints are aligned and sequenced and the degree to which the grade 11 SBA is aligned to the grade 11 blueprint)

Recommendations

- **Analyze academic performance data** at the school, district, and state level vis-à-vis the alignment study findings to determine which short-cycle diagnostic assessments are most valid (assessments vary in terms of standards and benchmark coverage and overall purpose)
- **Encourage professional development** at the teacher, school, and district level in the implementation of the short-cycle assessments to maximize benefit and impact
- **Seek “impact” data** from test publishers and NM districts to continually evaluate and monitor whether short-cycle assessments improve student achievement and ensure progress toward proficiency on the grade 11 standards