Consortium 2017 State-by-State Comparisons

D. J. McRae, Ph.D.

As of 09/27/17

When two consortiums of states were chosen by the US Department of Education in 2010 to develop new statewide assessment systems, one of the purposes was to generate state-by-state comparable achievement data. Roughly 45 states initially signed up for potential use of consortium tests, but by 2015 when the new tests were ready for their first "operational" use, only 18 states administered the Smarter Balanced tests and 11 states (plus the District of Columbia) administered the PARCC tests, representing just under 50 percent of total K-12 enrollments across the country. In 2017, 14 states administered Smarter Balanced tests, and 6 states (plus DC) administered PARCC tests, representing just over 30 percent of K-12 enrollments across the country. In addition, Louisiana and Massachusetts (PARCC) and Michigan (Smarter Balanced) used public domain test questions for their statewide tests but are not using the full consortium protocols and hence are not included in this set of state-by-state comparisons. Plans to use consortium tests in 2018 are noted on the data charts below.

The data charts on pages 2 and 5 provide state-by-state results for the Smarter Balanced and PARCC states, respectively, for spring 2017 testing. The results are expressed as "percents meeting target" grade-by-grade for English Language Arts and Mathematics, along with average percents across grades. On pages 3-4 and 6, the average gain scores for each state for 2015 thru 2017 are provided, respectively, for Smarter Balanced and PARCC states (plus DC). Notes describing the data in the charts are provided at the bottom of pages 3-4 and 6, respectively, for SBAC and PARCC states. Note that not all states have released grade-by-grade results yet; see footnote on page 7 for details. The results represent preliminary data released by states in many cases, with final data to be released later. Results from Smarter Balanced states are not comparable to results from PARCC states, and even within consortia there may be some differences in test administration or reporting practices across states. However, within consortiums, the comparability of scores is sufficient for general comparisons. The gain scores are comparable across consortiums.

Finally, it is fair game to average gain scores for ELA and Math for each state to produce an annual overall gain score for 2016 and 2017 results. And it is fair game to interpret each annual gain score as a letter grade based on a 4.0 grade point average (GPA) metric, with 4.0 being an A, 3.0 being a B, 2.0 being a C, 1.0 being a D, and 0.0 being an F. Annual gain scores for 2016 and 2017 for all consortium states (plus DC) are provided on page 7. See footnote for the 3 states without full data to date. As noted in the Observations on pages 8-9 that follow the data charts, the annual overall gain scores and assigned letter grades are comparable across consortiums.

Smarter Balanced 2017 State-by-State Comparisons [Level 3 & Above Percents]

Compiled by D. J. McRae, Ph.D. [As of 09/27/17]

English/Language Arts

	Grade	3	4	5	6	7	8	Ave
1	California	44	45	47	47	49	49	46.8
2	Connecticut	52	54	56	54	55	54	54.2
3	Delaware	52	54	60	52	53	52	53.8
4	Hawaii	48	48	53	49	50	47	49.2
5	Idaho	47	48	55	51	54	52	51.2
6	Montana	49	50	54	49	53	48	49.8
7	Nevada	45	46	50	43	47	46	46.2
8	New Hampshire	54	56	61	57	63	58	58.2
9	North Dakota	NA	NA	NA	NA	NA	NA	NA
10	Oregon	45	48	53	52	56	55	51.5
11	South Dakota	48	49	51	49	53	49	49.8
12	Vermont	49	49	55	52	55	55	52.5
13	Washington	53	55	59	56	60	59	57.0
14	West Virginia	45	47	49	45	48	45	46.5
	Averages	49	50	54	50	54	51	51.3

Mathematics

1	California	46	40	34	36	37	36	38.2
2	Connecticut	53	50	43	44	43	42	45.8
3	Delaware	53	50	44	41	41	38	44.5
4	Hawaii	53	48	42	41	37	38	43.2
5	Idaho	50	47	42	40	42	39	43.3
6	Montana	48	45	40	38	40	36	41.2
7	Nevada	48	41	34	30	29	18	33.3
8	New Hampshire	55	51	46	47	50	45	49.0
9	North Dakota	NA	NA	NA	NA	NA	NA	NA
10	Oregon	46	43	39	40	42	41	41.8
11	South Dakota	54	50	42	42	45	42	45.8
12	Vermont	52	47	42	39	44	41	44.2
13	Washington	58	54	49	48	50	48	51.2
14	West Virginia	48	43	34	32	31	29	36.2
	Averages	51	47	41	40	41	38	42.8

Smarter Balanced 2015-17 Gain Scores

English/Language Arts

	State	2015 ELA	2016 ELA	2017 ELA	2015-17 Gain	2016-17 Gain
1	California	42.3	46.7	46.8	+4.5	+0.1
2	Connecticut	xx*	55.8	54.2	xx*	-1.6
3	Delaware	51.7	54.8	53.8	+2.1	-1.0
4	Hawaii	47.7	50.5	49.2	+1.5	-1.3
5	Idaho	49.7	51.8	51.2	+1.5	-0.6
6	Montana	xx*	50.0	49.8	xx*	-0.2
7	Nevada	xx*	48.3	46.2	xx*	-2.1
8	New Hampshire	58.7	59.8	58.2	-0.5	-1.6
9	North Dakota	xx*	50.2	NA*	xx*	xx*
10	Oregon	53.8	53.3	51.5	-2.3	-1.8
11	South Dakota	47.5	51.2	49.8	+2.3	-1.4
12	Vermont	53.7	56.5	52.5	-1.2	-4.0
13	Washington	55.5	57.8	57.0	+1.5	-0.8
14	West Virginia	45.5	48.0	46.5	+1.0	-1.5

Notes for Smarter Balanced Data:

All averages and gains are based on Grade 3-8 data only. Less than half Smarter Balanced states uniformly administer Smarter Balanced tests for High Schools, and hence HS results are not included.

*Montana, Nevada, and North Dakota participated in Smarter Balanced testing in 2015, but all three states experienced technology difficulties that prevented generation of representative scores for the entire state. These circumstances prevent calculation of selected gain scores. Connecticut indicated they discontinued the Performance Task portion of ELA tests in 2016, so for comparability reasons the 2015 Connecticut average is not recorded. North Dakota's numerical results for 2017 are not available yet.

New Hampshire, North Dakota and West Virginia have indicated plans to discontinue use of Smarter Balanced tests spring 2018.

Mathematics

	State	2015 Math	2016 Math	2017 Math	2015-17 Gain	2016-17 Gain
1	California	34.2	37.3	38.2	+4.0	+0.9
2	Connecticut	40.3	44.2	45.8	+5.5	+1.6
3	Delaware	40.7	43.7	44.5	+3.8	+0.8
4	Hawaii	42.2	43.0	43.2	+1.0	+0.2
5	Idaho	40.8	43.3	43.3	+2.5	0.0
6	Montana	xx*	41.0	41.2	xx*	+0.2
7	Nevada	xx*	33.8	33.3	xx*	-0.5
8	New Hampshire	47.7	50.3	49.0	+1.3	-1.3
9	North Dakota	xx*	41.7	NA*	xx*	xx*
10	Oregon	43.5	42.8	41.8	-1.7	-1.0
11	South Dakota	41.2	44.5	45.8	+4.6	+1.3
12	Vermont	43.2	46.7	44.2	1.0	-2.5
13	Washington	49.8	51.5	51.2	+1.4	-0.3
14	West Virginia	30.8	34.8	36.2	+5.4	+1.4

Notes for Smarter Balanced Data:

All averages and gains are based on Grade 3-8 data only. Less than half Smarter Balanced states uniformly administer Smarter Balanced tests for High Schools, and hence HS results are not included.

*Montana, Nevada, and North Dakota participated in Smarter Balanced testing in 2015, but all three states experienced technology difficulties that prevented generation of representative scores for the entire state. These circumstances prevent calculation of selected gain scores. North Dakota's numerical results for 2017 are not available yet.

New Hampshire, North Dakota and West Virginia have indicated plans to discontinue use of Smarter Balanced tests for spring 2018.

PARCC 2017 State-by-State Comparisons [Level 4 & Above Percents]

Compiled by D. J. McRae, Ph.D. [As of 09/27/17]

English Language Arts

	Grade	3	4	5	6	7	8	Ave
1.	Colorado	40	44	46	41	44	43	43.0
2.	Dist Columbia	29	34	35	33	34	30	32.5
3.	Illinois	36	37	37	35	40	37	37.0
4.	Maryland	40	42	41	38	43	39	40.5
5.	New Jersey	50	56	59	53	59	59	56.0
6.	New Mexico	26	26	29	25	26	28	26.7
7.	Rhode Island	40	41	42	38	40	37	39.7
	Averages	37	40	41	38	41	39	39.3

Mathematics

	Grade	3	4	5	6	7	8	Ave
1.	Colorado	40	34	34	31	26	xx	33.0
2.	Dist Columbia	38	33	31	26	23	xx	30.2
3.	Illinois	39	31	35	28	27	32	31.2
4.	Maryland	43	37	36	32	xx	xx	37.0
5.	New Jersey	52	47	46	44	40	xx	45.8
6.	New Mexico	30	23	23	20	16	xx	22.4
7.	Rhode Island	44	35	35	29	27	xx	34.0
	Averages	41	34	34	30	27	xx	33.4

PARCC 2015-17 Gain Scores

English/Language Arts

State	2015 Ave	2016 Ave	2017 Ave	2015-17 Gain	2016-17 Gain
1. Colorado	40.3	40.5	43.0	+2.7	+2.5
2. Dist Columbia	26.2	26.5	32.5	+6.3	+6.0
3. Illinois	36.0	36.5	37.0	+1.0	+0.5
4. Maryland	38.8	38.7	40.5	+1.7	+1.8
5. New Jersey	50.0	52.8	56.0	+6.0	+3.2
6. New Mexico	23.2	24.5	26.7	+3.5	+2.2
7. Rhode Island	36.8	39.8	39.7	+2.9	-0.1

Mathematics

1. Colorado	31.2	32.6	33.0	+1.8	+0.4
2. Dist Columbia	25.3	29.5	30.2	+4.9	+0.7
3. Illinois	28.8	31.3	31.2	+2.4	-0.1
4. Maryland	31.8	37.3	37.0	+5.2	-0.3
5. New Jersey	41.0	45.6	45.8	+4.8	+0.2
6. New Mexico	21.4	23.0	22.4	+1.0	-0.6
7. Rhode Island	28.2	34.0	34.0	+5.8	0.0

Notes for PARCC Data:

Most PARCC states utilized PARCC End-of-Course Math tests for the High School level rather than PARCC grade level tests. Since course taking patterns differ from state-to-state, HS results are not included.

All averages and gains reflect grades 3-8 only, and the xx under a grade means the state does not uniformly administer PARCC grade level tests to all students at that grade level. All xx's reflect administrations of Algebra and/or Geometry End-of-Course tests to grade 7 and 8 students taking these courses, rather than the regular grade level tests. The pattern of xx's in the 2017 chart are identical to the patterns for 2015 and 2016. Averages are based on grade levels with comparable results.

Colorado and Rhode Island have indicated plans to discontinue use of PARCC tests for spring 2018.

ELA / Math Gain Score Averages by Year with GPA's and Letter Grades [As of 9/27/17]

	Ave Gain	GPA /	Ave Gain	GPA /
SBAC	15-16	Letter	16-17	Letter
1. California	3.75	A	0.50	D-
2. Connecticut	xx	--	0.00	F
3. Delaware	3.05	B	-0.10	F
4. Hawaii	1.80	C	-0.55	F
5. Idaho	2.30	C+	-0.30	F
6. Montana	xx	--	0.00	F
7. Nevada	xx	--	-1.30	F
8. New Hampshire	1.85	C	-1.45	F
9. North Dakota	xx	--	PC	Inc
10. Oregon	-0.60	F	-1.40	F
11. South Dakota	3.50	A-	-0.05	F
12. Vermont	3.65	A-	-3.25	F
13. Washington	2.00	C	-0.55	F
14. West Virginia	3.25	B+	-0.05	F
Averages	2.46	C+	-0.65	F
PARCC				
1. Colorado	0.80	D	1.45	D+
2. Dist Columbia	2.25	C+	3.35	B+
3. Illinois	1.50	C-	0.20	F
4. Maryland	2.70	B-	0.75	D
5. New Jersey	3.70	A-	1.70	C-
6. New Mexico	1.45	D+	0.80	D
7. Rhode Island	4.40	A+	-0.05	F
Averages	2.40	C+	1.17	D

GPA to Letter Conversions

A = 3.50 to 4.49, B = 2.50 to 3.49, C = 1.50 to 2.49, D = 0.50 to 1.49, F = 0.00 to +0.49; Within each range, the higher range of 0.25 to 0.49 merits a + sign, the lower range of 0.50 to 0.74 merits a – sign; Equal to or greater than 4.50 merits an A++, less than 0.00 merits a F.

Inc: Incomplete. North Dakota has not released numerical results for 2017 yet, but personal communication with ND Dept Educ officials indicated overall gain scores for 2016-17 were less than zero.

Observations for Smarter Balanced and PARCC 2016 State-by-State Comparison Scores

Smarter Balanced vs PARCC Results, ELA vs Math Results, Trends Across Grades, and Gain Results

It is clear students score better on Smarter Balanced tests than on PARCC tests. Smarter Balanced states averaged 51 percent meeting targets for ELA and 43 percent for Math, while PARCC states averaged 39 percent meeting targets for ELA and 33 percent for Math. In 2004, widely respected educational measurement expert Bob Linn noted differences of 3 to 4 percent are clearly meaningful. Differences of 10 to 12 percent are large meaningful differences.

In addition, consortium ELA tests averaged 45 percent meeting target, while consortium Math tests averaged 38 percent meeting target, another meaningful difference.

While there may be demographic differences between the two cohorts of states, or there may be differences for implementation of common core instruction, it is unlikely either of these reasons would cause the large differences in Smarter Balanced scores vs PARCC scores. Rather, it is likely that the differences between Smarter Balanced and PARCC results are due to the tests themselves, either in the difficulty of the items or in the setting of threshold scores for the respective targets upon which the data in the charts are based. Perhaps the best way to describe the differences between Smarter Balanced and PARCC results is simply that PARCC has the more difficult set of tests.

A look at trends across grades shows no obvious trends for ELA results for both consortiums, but do show declining results for Math as the grades increase for both consortiums. These trends across grades are very similar to the trends across grades found for both Smarter Balanced and PARCC 2015 and 2016 results

The gain scores on pages 3-4 and 6 as well as the gain scores on page 7 are comparable across Smarter Balanced and PARCC. For Smarter Balanced states, it appears the losses from 2016 to 2017 are substantially greater for ELA than for Math. For PARCC, the gains from 2016 to 2017 are greater for ELA than for Math. These results suggest that 2017 Smarter Balanced results are more problematic for ELA than for Math. On page 7, the overall Average Gains (losses for most Smarter Balanced states) are starkly different than the modest gains for PARCC, with all but one of the Smarter Balanced states registering "gains" of zero or less while all but one of the PARCC states registering gains above zero. The GPA-like metrics for overall gains and losses translate into Letter grades that communicate these differences very nicely.

Other Considerations

One factor that may affect results from state-to-state is differential initiation of common core instruction, affecting student Opportunity-to-Learn (OTL) in each state. OTL is the notion that students should be taught material on tests before large scale tests are initiated or administered. For both Smarter Balanced and PARCC states, the timing for initiation of tests was largely set by federal funded consortium agreements, not by state-by-state analysis of OTL status. If a state had not sufficiently implemented common core instruction before administering Smarter Balanced or PARCC tests, results must be interpreted with great caution – for these states, the results most likely do not meet professional technical standards for validity, reliability, and fairness. Analysis of OTL needs to be conducted on a state-by-state basis. This factor was likely notable for many states for the 2015 results, less so for 2016 and 2017.

Finally, it should be noted that changes (or lack of needed changes) in the tests between 2015 and 2017 may affect the gain scores displayed in the charts above. For example, the Smarter Balanced submission for federal peer review covering spring 2015 tests "revealed some gaps in item coverage at the low end of the performance spectrum." How this deficiency for the 2015 Smarter Balanced tests affected Smarter Balanced scores in 2016 and/or 2017 is unknown, but addition of test questions that measure the low end of the performance spectrum to the Smarter Balanced computer-adaptive item bank would likely have a substantial effect. Whether Smarter Balanced released items from their computer-adaptive item bank and/or added new verified items in 2016 or 2017 is unknown. However, it is noteworthy that the Vermont press release for their 2017 results included a sentence to the effect that Vermont's achievement gaps decreased in 2017, but the decreases were not due to increases in scores for low achieving students but rather due to decreases in scores for higher achieving students; this pattern of results may suggest changes in the Smarter Balanced computer-adaptive item bank for 2017 that affected scores in the higher end of the achievement spectrum.

Author Tagline

Doug McRae is a retired educational measurement specialist living in Monterey, California. In his more than 45 years in the K-12 testing field, he has served as an educational testing company executive in charge of the design and development of K-12 tests widely used across the country, as well as an advisor for the design and development of California's STAR system which was used from 1998 through 2013. He has a Ph.D. in Quantitative Psychology from the L. L. Thurstone Psychometric Laboratory at the University of North Carolina, Chapel Hill.