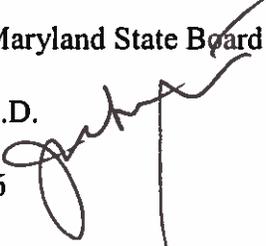




200 West Baltimore Street • Baltimore, MD 21201 • 410-767-0100 • 410-333-6442 TTY/TDD • msde.maryland.gov

TO: Members of the Maryland State Board of Education
FROM: Jack R. Smith, Ph.D. 
DATE: February 23, 2016
SUBJECT: Graduation Assessment Requirements

PURPOSE:

To provide a recommendation as to what PARCC scale score/performance level should be adopted to satisfy the assessment graduation requirement for Algebra I and English 10

BACKGROUND:

As part of the graduation assessment minimum score determination decision by the State Board, the Maryland State Department of Education in collaboration with the Maryland Assessment Research Center conducted a research study to obtain PARCC equivalent scale scores. The High School Assessments in English 10 and Algebra I administered to Maryland students since 2003 required students to meet a minimum scale score in English 10 of 396 and a minimum scale score of 412 in Algebra I.

EXECUTIVE SUMMARY:

Two equating methods (Propensity vs. Correlational Two-Step Linking Study) produced similar scale scores when equating HSA and PARCC test scores. The corresponding PARCC Algebra I score to the HSA Algebra passing score of 412 is 720 and the corresponding PARCC ELA 10 score to the HSA English passing score of 396 is 707. For both contents, the scale score fell within the established confidence interval including one standard deviation above and below the PARCC equivalents scale score.

	HSA Minimum Passing Score	PARCC Equivalent Score	PARCC Confidence Interval	
ALG I	412	720	697	743
ELA 10	396	707	686	728

Given the psychometric analysis, understanding performance trends when introducing new assessments, and practical considerations of aligning to the PARCC performance level descriptors, MSDE recommends a scale score of 725, which corresponds with PARCC Performance Level 3 as the passing high school scale score for both PARCC ALG I and ELA 10.

Members of the State Board of Education
February 23, 2016
Page 2

ACTION:

For information purposes only. No action required at this time.

March: Review proposed language for COMAR.

April: Approve COMAR language for publish for public comment.

Attachment: Investigating the Concordance Relationship between the HSA Cut Scores and the PARCC Cut Scores

Investigating the Concordance Relationship between the HSA Cut Scores and the PARCC Cut Scores

A Research Report Submitted to the Maryland State Department of Education (MSDE)

By

Maryland Assessment Research Center (MARC)

Executive Summary

The purpose of this study is to obtain the PARCC equivalent of the HSA cut score and the HSA equivalent of the PARCC cut score that divides performance level 2 from 3. More specifically, the HSA Algebra cut score needs to be mapped onto the PARCC Algebra I scale and the HSA English cut score to the PARCC ELA10 scale and vice versa. The cut scores for passing HSA English and Algebra are currently 396 and 412 respectively. The cut scores for being in performance level 3 are 725 for both PARCC ELA10 and Algebra I. Based on the discussions at the recent Technical Advisory Committee (TAC) meeting, the MARC team explored the following two options to create the concordance tables:

1. Option I: Using PSAT as an external common test to link HSA and PARCC tests via two-step linking. As item level response data are not available, the equipercentile linking method is used to set up the linkage using a single group design. The exploration was conducted using the first-time test takers' scores.
2. Option II: Using the propensity score matching method to come up with matched equivalent groups so that the equivalent group linking method can be used to map the HSA cut scores onto the PARCC scales directly, and vice versa. The equipercentile linking method is used to set up the linkage using the first-time test takers' scores.

Major Findings

The detailed data cleaning, preparation, and analyses are documented in this report. The following summarizes the major findings based on this current exploration.

1. Using PSAT as an external common test to link HSA and PARCC tests via two-step linking produced PARCC equivalent cut scores of 707 and 720 for PARCC ELA10 and Algebra I respectively. Overall, the PARCC equivalent cut score for ELA10 yields a passing rate falling within the ranges of the HSA historical passing rates for both the May test-takers who resembled the 2015 PARCC test-takers and the yearly passing rates. On the other hand, the PARCC equivalent cut score for Algebra I yields a passing rate slightly lower than the lowest HSA yearly

passing rate and falls outside the range of the historical passing rates of the May test administration.

2. Using the propensity score matching method under different matching conditions produced PARCC equivalent cut scores of 704, 705, and 706 for PARCC ELA10 depending on the matching conditions and 721 and 725 for Algebra I depending on the use of different matching samples. Further when combining Design II and III matched samples, the cut scores were 722, 724, and 727 depending on the matching conditions. The PARCC ELA10 equivalent scores for the HSA English test yield higher passing rates compared with the PSAT linking method. These passing rates fall within the ranges of both the May and the yearly passing rates except for the cut score of 704 leading to a much higher passing rate. Based on this method, the PARCC Algebra I equivalent scores for the HSA Algebra test produced much lower passing rates that fall outside the ranges of the passing rates for both the May tests and the yearly passing rates.
3. 95% confidence intervals and one standard deviation above and below the PARCC equivalents of the HSA cut scores were constructed. For ELA10, the 95% confidence interval around the mapped PARCC equivalent score of the HSA cut score using the mean and the maximum conditional standard error of measurement (CSEM) contained the PARCC cut score of 725 which divides performance level 2 from 3 while the interval one standard deviation above and below the CSEM did not capture the mapped PARCC cut score. Neither does the 95% confidence interval using the minimum CSEM. For Algebra I, all intervals contained the PARCC cut score of 725. The patterns were consistent across linking methods.
4. The HSA equivalents of the PARCC cut score of 725 that divides performance level 2 from 3 are summarized. In general, the HSA equivalents of the PARCC cut score, 725 for both ELA10 and Algebra were higher than the original HSA cut scores.
5. This study provides empirical evidence about the PARCC equivalents of the HSA cut scores and the HSA equivalents of the PARCC cut score of 725 that divides performance level 2 from 3 for ELA10 and Algebra I. The final adoption of cut scores obtained in this study depends on considerations from psychometric, policy, and practical perspectives.

Option I

Using PSAT as External Linking Tests

Data Cleaning and Preparation

The three datasets used in this exploration are from the PARCC, PSAT, and HSA tests. Data cleaning was conducted prior to data analysis for English and Algebra tests respectively. In the HSA layout table, 05 stands for Algebra test. The team used Test Format by Content information in the dataset as supplemental information to find the code for English test (Code 06 for HSA English). For each HSA dataset, the first timer test scores were selected and used in the analyses when multiple attempts were found. Further, only the regular students were selected for the linking study.

For the PARCC test and PSAT tests, the dataset was separated into English and Math test and the first time test scores for each unique student ID were extracted using testing year information. For duplicated cases (the same test year and administration but with different scores), first entry record was used.

The contents areas of the PARCC, PSAT, and HSA tests are summarized in Table 1.1. The subjects used in this study are the PARCC Algebra I, PARCC ELA10, PSAT Math, PSAT Verbal, HSA English, and HSA Algebra. The PARCC test data are from the 2015 administrations. The HSA test data and the PSAT test data are from the administrations during 2008 to 2015. The HSA test was administrated five times a year, and the PSAT test was administrated once a year. The name of subjects such as Algebra and Math, English and Verbal, are used interchangeably in this report.

Table 1.1
Subjects in Each Test

Test	Subjects
PARCC	Algebra I, Algebra II, ELA10
PSAT	Math, Verbal, Writing
HSA	English, Biology, Government, Algebra/Data Analysis

Table 1.2 provides the summary statistics for the HSA Algebra and English tests after data cleaning. For both the HSA Algebra and English tests, the minimum score is 240 and the maximum score is 650. The average Algebra test score is 424.96. The standard deviation of Algebra test scores is also higher than that of English test scores.

Table 1.2
Summary Statistics for the HSA Test

Test	N	Mean	SD	Min	Max
English	441,957	409.49	33.68	240	650
Algebra	485,673	424.96	41.43	240	650

Table 1.3 provides the summary results for the PARCC Algebra I and ELA10 tests using the first-time test takers' scores. The total number of PARCC Algebra I test takers is 61,760 while that for the PARCC ELA10 test is 55,629. The standard deviation of the PARCC Algebra I test scores is lower than that of the PARCC ELA10 test scores.

Table 1.3

Summary Statistics for the PARCC Test

Test	N	Mean	SD	Min	Max
ELA10	55,629	737.8	44.95	650	850
ALG I	61,760	734.3	32.81	650	850

Table 1.4 provides the summary results for the PSAT test scores. All students are required to take both the PSAT Verbal and Math tests at the same time; therefore, the sample size for the Math and Verbal test is the same. The standard deviations of both tests are similar.

Table 1.4

Summary Statistics for the PSAT Test

	N	Mean	SD	Min	Max
Verbal	515,109	40.37	10.90	20	80
Math	515,109	41.47	11.09	20	80

In order to use the PSAT test as an external linking test, the HSA test was merged with the PSAT test and the PSAT test was merged with the PARCC test using the state issued student ID. Specifically, the PSAT Verbal test was merged with the HSA English test, the PSAT Verbal test was merged with the PARCC ELA10 test using the student ID. The PSAT Math test was merged with the HSA Algebra test, the PSAT Math test was merged with the PARCC Algebra I test. In total, there are four merged datasets and the descriptive statistics for the PSAT test in each merged dataset are summarized in Table 1.5. Descriptive statistics for the HSA test and the PARCC test in the merged datasets are summarized in Table 1.6.

Table 1.5

Summary Statistics for the PSAT Scores after Merging with the HSA and PARCC Tests

Subject	Test	N	Mean	SD	Min	Max	Correlation
English	PSAT Verbal & HSA English	381,599	40.28	10.69	20	80	0.711
	PSAT Verbal & PARCC ELA10	46,680	40.80	10.49	20	80	0.712
Math	PSAT Math & HSA Algebra	366,632	40.10	10.43	20	80	0.711
	PSAT Math & PARCC ALG I	11,018	33.09	7.36	20	69	0.581

Table 1.6

Summary Statistics for the HSA and PARCC Test Scores after Merging with the PSAT

Subject	Test	N	Mean	SD	Min	Max
English	HSA	381,599	413.27	31.42	240	650
	PARCC	46,680	742.11	44.34	650	850
Math	HSA	366,632	428.71	37.67	240	650
	PARCC	11,018	718.87	26.16	650	850

Using the PSAT Tests to Link the HSA and PARCC Tests

After data cleaning and matching samples, the equipercentile linking method was conducted based on the matched samples of HSA and PSAT first and then those of PSAT and PARCC for both Algebra and English tests. The Linking with Equivalent Group or the Single Group Design (LEGS) program developed by Kolen and Brennan was used to link the two matched samples. After specifying the input data format which is the scores and frequencies, subgroup information (no subgroup in this study), smoothing parameters and score truncation in the original scale scores, the LEGS program reported the results for the equipercentile linking based on the single group design for mapping HSA to PSAT, then PSAT to PARCC based on a two-step linking approach. In Appendix A, a screenshot capturing the input window for linking HSA and PSAT English using the first-time test-takers' scores was shown. Two smoothing values were compared in post-linking: 0.3 and 1. The choice of using smoothing parameters is supported by simulation studies that show the smoothed results outperforming the non-smoothed results in reducing linking errors (Cui & Kolen, 2009; Hanson et al., 1994). The results using smoothing value of 1 were reported due to the fact that after rounding there was little difference between the results based on the two smoothing parameters.

The concordance tables were generated using LEGS. Single group design was used in this part. The passing score or proficiency score for the HSA English is 396 and for the HSA Algebra is 412. As was shown in Tables 1.7 to 1.10, the corresponding score for the PARCC ELA10 is 707 and for the PARCC Algebra I test is 720. The direct concordance tables between the HSA and PARCC tests are presented in Tables 1.11 and 1.12 for ELA and Algebra respectively. An imputation equation was developed based on the available HSA and PARCC scores matched via the same PSAT scores. Impact data or the passing rate for different cut score are presented in the concluding part of this report.

In other words, the HSA English cut score of 396 was mapped to a PSAT score of 33. Then the PSAT score of 33 was mapped to a PARCC score of 707. Therefore, a PARCC equivalent score of the HSA English cut score of 396 is 707. Following the same logic, the cut score of 412 for the HSA algebra test was mapped to a PSAT score of 33. Then the PSAT score of 33 was mapped to a PARCC Algebra I score of 720. Therefore, a PARCC Algebra I equivalent score of the HSA Algebra cut score of 412 is 720.

Option II

Using Equivalent Groups Based on Propensity Score Matching to Link HSA and PARCC Tests

Based on the discussions at the recent TAC meeting, the following three designs were suggested to link the HSA and PARCC tests based on equivalent groups from propensity score matching. Six covariates were used in matching; they are gender, race, limited English proficiency (LEP), FARMS, Title I, and MSA test scores in the same content area.

Design I (English)

Group 1: HSA 2014 Grade 10 English + MSA 2012 Grade 8 Reading

Group 2: PARCC 2015 Grade 10 Algebra I + MSA 2013 Grade 8 Reading

Design II (Algebra)

Group 1: HSA 2014 Grade 9 Algebra+ MSA 2013 Grade 8 Math

Group 2: PARCC 2015 Grade 9 Algebra I + MSA 2014 Grade 8 Math

Design III (Algebra)

Group 1: HSA 2014 Grade 8 Algebra + MSA 2013 Grade 7 Math

Group 2: PARCC 2015 Grade 8 Algebra I + MSA 2014 Grade 7 Math

Combined Design II & III (Algebra)

Group 1: HSA 2014 Grade 9 Algebra+ MSA 2013 Grade 8 Math + HSA 2014 Grade 8 Algebra + MSA 2013 Grade 7 Math

Group 2: PARCC 2015 Grade 9 Algebra I + MSA 2014 Grade 8 Math + PARCC 2015 Grade 8 Algebra I + MSA 2014 Grade 7 Math

Prior to data analysis, the HSA test scores were merged with the above matched MSA test scores using testing year, grade, and state issued ID information for the regular first-time test-takers for each of the above mentioned three designs. Further, the PARCC test scores were also merged with the MSA test scores based on the above matched test year, grade, and state issued ID information for each design.

For Design I, after extracting first-time test takers' scores and removing students taking the Modified MSA tests, the matched sample size for HSA and MSA for Group 1 is 47,656. For Group 2, the matched sample size for PARCC and MSA is 46,692. For Design II, the matched sample size between HSA and MSA for Group 1 is 23,738; for Group 2, the matched sample size for PARCC and MSA is 26,704. For Design III, Group 1 matched sample size between HSA and MSA is 24,420 while the matched sample size

between PARCC and MSA for Group 2 is 16,525. Table 2.1 summarizes the matched sample sizes for each pair.

Table 2.1
Sample Sizes for Matched Cases in Each Group under Each Design

Design	Matched Pair	Sample Size
Design I	Group1 HSA English with MSA	47,656
	Group2 PARCC ELA10 with MSA	46,692
Design II	Group1 HSA Algebra with MSA	23,738
	Group2 PARCC ALG I with MSA	26,704
Design III	Group1 HSA Algebra with MSA	24,420
	Group2 PARCC ALG I with MSA	16,525

In the merged dataset, six covariates were utilized for propensity score matching. As stated above, the six covariates are Gender, Race, LEP, Farms, Title I and MSA scores in the same content area. Gender, Race, LEP, Farms and Title I are variables from the HSA test dataset in Group 1 and the PARCC test dataset in Group 2 in all three designs. For the Gender variable, males are coded as 1 and females are coded as 0. For the Race variable, White is coded as 1 and all others are coded as 0. We also explored to code White/Asian as 1 and 0 for others. However, the standardized mean difference for each covariate is larger than the current method. LEP is coded as 1 for students with limited English proficiency and 0 for others. The Farms variable is coded as 1 for students who take free and reduced priced meals and 0 for students who do not. The Title I variable is coded as 1 for students who belong to this category and 0 for students who do not belong to this category. The MSA scale score was used as a covariate directly with no recoding needed. No missing data were detected for the six covariates in the three designs.

R studio was used for propensity score matching. The package “MatchIt” developed by Ho, Imai, K. and Imai, M. (2013) was used to match cases in the control group to those in the treatment group. Usually the group with a smaller sample size is treated as the treatment group, and this was done in matching HSA and PARCC tests. The Match.Matrix function in the package was called to export one-to-one matched case IDs. For better matching, this study explored four conditions for each design by using different caliper values and the use of replacement of cases in matching. Caliper, which is the maximum degree of difference to be considered as a match, was set at two levels: caliper of 0.1 and caliper of 0.25. Replacement was set at two levels: with and without replacement of cases. Replacement means that the cases in the control group can be used multiple times to match those in the treatment group.

To compare the similarity of the treated and control subjects in the matched sample, the standardized mean difference is commonly used as an indicator for what is called a balance check. It can be used to compare the mean of continuous and binary variables between the treatment and control groups. For a continuous covariate, the standardized mean difference is defined as

$$d = \frac{\bar{x}_{treatment} - \bar{x}_{control}}{\sqrt{\frac{s_{treatment}^2 + s_{control}^2}{2}}}$$

where $\bar{x}_{treatment}$ and $\bar{x}_{control}$ denote the sample mean of the covariate in treated and control subjects, respectively, whereas $s_{treatment}^2$ and $s_{control}^2$ denote the sample variance of the covariate in the treated and control groups, respectively.

The standardized mean difference compares the difference in means in units of the pooled standard deviation. Furthermore, it is not influenced by sample size and allows for the comparison of the relative balance of variables measured in different units. Although there is no universally agreed upon criterion as to what threshold of the standardized difference can be used to indicate important imbalance, an absolute value of standardized mean difference that is less than 0.25 has been suggested to indicate a negligible difference in the mean of a covariate between the treatment group and control group (Stuart, 2010).

Table 2.2
Propensity Score Matching Results for Design I

Condition No.	1.1	1.2	1.3	1.4
K	1	1	1	1
Caliper	0.1	0.1	0.25	0.25
Replacement	NO	YES	NO	YES
Gender	0.0010	0.0005	0.0003	0.0068
Race	0.0038	0.0114	0.0047	0.0020
LEP	0.0062	0.0095	0.0151	0.0132
FARMS	0.0017	0.0027	0.0037	0.0118
Title1	0.0006	0.0004	<0.0001	0.0006
MSA	0.0030	0.0027	0.0040	0.0014
HSA English (Control)	46,228	29,663	46,311	29,421
PARCC ELA10 (Treatment)	46,228	46,691	46,311	46,692

Table 2.3

Propensity Score Matching Results for Design II

Condition No.	2.1	2.2	2.3	2.4
K	1	1	1	1
Caliper	0.1	0.1	0.25	0.25
Replacement	NO	YES	NO	YES
Gender	0.0009	0.0013	0.0011	0.0141
Race	0.0091	0.0102	0.0082	0.0033
LEP	0.0021	0.0053	0.0014	0.0014
FARMS	0.0017	0.0050	0.0004	0.0065
Title1	0.0071	0.0070	0.0084	0.0111
MSA	0.0139	0.0145	0.0043	0.0016
HSA Algebra (Treatment)	23,316	23,733	23,522	23,736
PARCC ALG I (Control)	23,316	15,627	23,522	15,590

Table 2.4

Propensity Score Matching Results for Design III

Condition No.	3.1	3.2	3.3	3.4
K	1	1	1	1
Caliper	0.1	0.1	0.25	0.25
Replacement	NO	YES	NO	YES
Gender	0.0009	0.0021	0.0025	0.0064
Race	0.0144	0.0143	0.0184	0.0139
LEP	0.0219	0.0173	0.0141	0.0179
FARMS	0.0204	0.0140	0.0186	0.0160
Title1	0.0170	0.0116	0.0123	0.0048
MSA	0.0067	0.0035	0.0049	0.0017
HSA Math (Control)	16,118	11,754	16,268	11,820
PARCC ALG I (Treatment)	16,118	16,504	16,268	16,522

In Tables 2.2 to 2.4, the 12 conditions are labeled from 1.1 to 3.4 for convenience. The first number represents each of the three designs and the second number represents the matching condition based on the combination of different caliper values and matching with or without replacement. For example, Condition 3.1 represents one-to-one matching with a caliper value of 0.1 and no replacement. Each of these tables presents the absolute standardized mean difference values for each covariate. The bottom part in each of the three tables contains the number of matched cases in the treatment group and the control group. In this study, the group with fewer cases (the sample size is indicated in Table 2.1) was chosen as the treatment group and the other group was chosen as the control group in order to maximize the sample size of the matched cases in both the treatment and control groups. Therefore, in each design, either Group 1 or Group 2 was chosen as a treatment group based on the sample size of the matched cases in Table 2.1. The values of the absolute standardized mean differences in Tables 2.2 to 2.4 were checked. The results

indicated that for these three designs, the covariates were balanced after matching. The descriptive statistics for the HSA and PARCC test scores for the matched groups for each design and each matching condition are summarized in Tables 2.5 to 2.7.

Table 2.5

Descriptive Statistics for HSA and PARCC in the Matched Data in Design I (English)

	Test Name	N	Mean	SD	Min	Max
Condition 1	HSA	46,228	413.48	28.97	240	650
	PARCC	46,228	741.33	43.96	650	850
Condition 2	HSA	29,663	413.15	29.05	240	650
	PARCC	46,691	740.97	44.02	650	850
Condition 3	HSA	46,311	413.46	28.76	240	650
	PARCC	46,311	741.26	43.97	650	850
Condition 4	HSA	29,421	413.46	28.02	240	650
	PARCC	46,692	740.97	44.02	650	850

Table 2.6

Descriptive Statistics for HSA and PARCC in the Matched Data in Design II (Algebra)

	Test Name	N	Mean	SD	Min	Max
Condition 1	HSA	23,316	404.20	40.63	240	650
	PARCC	23,316	721.67	26.11	650	850
Condition 2	HSA	23,733	404.44	40.51	240	650
	PARCC	15,627	721.58	26.22	650	850
Condition 3	HSA	23,522	404.41	40.53	240	650
	PARCC	23,522	721.64	26.15	650	850
Condition 4	HSA	23,736	404.44	40.44	240	650
	PARCC	15,590	721.51	26.09	650	850

Table 2.7

Descriptive Statistics for HSA and PARCC in the Matched Data in Design III (Algebra)

	Test Name	N	Mean	SD	Min	Max
Condition 1	HSA	16,118	433.48	33.41	240	650
	PARCC	16,118	749.56	29.05	650	850
Condition 2	HSA	11,754	432.1	34.4	240	650
	PARCC	16,504	750.5	29.58	650	850
Condition 3	HSA	16,268	433.56	33.46	240	650
	PARCC	16,268	749.88	29.24	650	850
Condition 4	HSA	11,820	432.42	34	240	650
	PARCC	16,522	750.59	29.69	650	850

After propensity score matching, the matched data were exported from all conditions in the three designs. LEGS program was again used for equipercentile linking using the equivalent group design using frequency data. The propensity score matching with replacement weighs different cases differently. Weights for cases in the control group (with a larger sample size) may be a value larger or smaller than 1 while the weights for cases in the treatment group (with a smaller sample size) are still 1. Thus, in computing the frequency for the control group in the matched sample, weights assigned to each case were summed up and used as the frequency for each case. The sum of the weights is rounded up if larger than 0.5.

In total, there are 16 concordance tables created based on propensity score matching. The PARCC equivalents of the HSA cut scores for each matching condition are summarized in Table 2.8. The 16 HSA and PARCC concordance tables are presented in Tables 2.9 to 2.24.

Table 2.8

PARCC Equivalent Scores of the HSA Cut Scores Using Propensity Score Matching

Sub-Condition	1	2	3	4
Caliper	0.1	0.1	0.25	0.25
Replacement	NO	YES	NO	YES
Design I (ELA10)	706	706	705	704
Design II (ALG I)	721	721	721	721
Design III (ALG I)	725	725	725	725
Combined Design II & III (ALG I)	722	727	722	724

Impact

To evaluate the impact of the cut scores obtained using different methods to link HSA and PARCC tests, the percentage of passing for each cut score is summarized in Tables 2.25 and 2.26 for ELA10 and Algebra respectively. The red color indicates the cut scores obtained using PSAT as an external linking test while the green color indicates the cut scores obtained using the propensity score matching method. For Algebra I, the blue color indicates the cut scores using the combined matched samples from Design II and III using propensity score matching. The black color indicates the passing rates for other PARCC scores adjacent to the cut scores obtained in this study.

Table 2.25

Passing Rates for the PARCC ELA10 Test

Cut score	700	701	702	703	704	705	706	707	708	709	710	711	712
Passing rate	78.65%	77.98%	77.29%	76.72%	76.07%	75.30%	74.67%	73.96%	73.23%	72.60%	71.88%	71.23%	70.32%
Count	43,750	43,378	42,997	42,676	42,314	41,889	41,540	41,144	40,737	40,386	39,984	39,625	39,118
Cut score	713	714	715	716	717	718	719	720	721	722	723	724	725
Passing rate	69.91%	69.02%	68.40%	67.76%	67.89%	66.22%	65.43%	64.62%	63.86%	63.02%	62.32%	61.35%	60.60%
Count	38,889	38,395	38,049	37,695	37,768	36,839	36,398	35,945	35,526	35,058	34,670	34,127	33,713

Table 2.26

Passing Rates for the PARCC Algebra I Test

Cut score	700	701	702	703	704	705	706	707	708	709	710	711	712	713
Passing rate	86.99%	85.54%	84.36%	83.47%	83.07%	82.51%	81.44%	79.18%	78.58%	78.04%	77.20%	75.53%	73.83%	73.26%
Count	53,722	52,832	52,100	51,550	51,305	50,958	50,298	48,904	48,533	48,199	47,679	46,649	45,597	45,246
Cut score	714	715	716	717	718	719	720	721	722	723	724	725	726	727
Passing rate	72.36%	70.68%	69.15%	68.29%	63.58%	65.83%	65.06%	63.58%	62.08%	61.00%	60.34%	58.37%	57.170%	56.346%
Count	44,690	43,649	42,709	42,177	39,264	40,659	40,182	39,264	38,342	37,675	37,267	36,047	35,308	34,799

These passing rates are also compared with the HSA historical passing rates as shown in Tables 2.27 and 2.28 for English and Algebra respectively. Figures 1 and 2 present the trend of the passing rate for HSA tests across years. In general, students taking HSA in different months differed in their test scores for both English and Algebra. Within each year, a majority of the students took the May HSA tests. Students who took the 2015 PARCC would be expected to resemble the May test takers of HSA better than other months' test-takers. The passing rates for the May HSA English tests ranged from 68.78 % to 76.74% while those for Algebra ranged from 67.70% to 75.23%. The yearly passing rates from 2008 to 2014 go from 64.32 % to 75.62% for English and from 65.51% to 73.77% for Algebra.

Overall, the PARCC ELA10 equivalent cut scores based on both methods produced the passing rates falling within the range of the HSA historical May and yearly passing rates except the cut score of 704 yielding a higher passing rate. Compared with the propensity score matching method, the PSAT linking produced a slightly higher PARCC equivalent cut score which leads to slightly lower passing rate for ELA10.

Table 2.27

Passing Rates for the HSA English Test

Month	Year	Min	Max	Mean	SD	N	%pass	year %pass
Jan	2008	240	650	391.88	36.99	11125	44.41%	64.32%
Jan	2009	240	650	402.72	34.97	7492	60.76%	71.27%
Jan	2010	240	650	408.03	33.40	6883	68.63%	73.62%
Jan	2011	240	650	405.02	34.49	7497	67.77%	73.68%
Jan	2012	240	650	407.42	33.23	6765	68.38%	75.62%
Jan	2013	240	522	403.82	36.00	5568	68.12%	73.05%
Jan	2014	240	650	402.70	36.73	4911	67.28%	74.04%
April	2009	240	455	382.67	38.00	307	41.37%	71.27%
April	2010	240	650	387.94	43.51	129	41.09%	73.62%
April	2011	240	450	382.96	34.24	144	37.50%	73.68%
April	2012	240	448	378.82	37.04	101	34.65%	75.62%
April	2013	240	475	385.79	34.69	140	35.00%	73.05%
April	2014	240	447	376.51	46.04	122	37.70%	74.04%
May	2008	240	650	409.82	34.70	58173	68.78%	64.32%
May	2009	240	650	411.41	33.43	55007	73.35%	71.27%
May	2010	240	650	411.39	32.37	54679	74.58%	73.62%
May	2011	240	650	411.34	33.03	53671	74.75%	73.68%
May	2012	240	650	413.29	30.09	52767	76.74%	75.62%
May	2013	240	650	409.94	34.19	52480	73.68%	73.05%
May	2014	240	650	410.76	32.07	52961	74.96%	74.04%
July	2008	240	462	385.70	38.60	310	47.42%	64.32%
July	2009	240	469	391.81	40.94	160	55.00%	71.27%
July	2010	240	484	393.81	39.99	126	57.14%	73.62%
July	2011	240	463	388.22	44.73	103	58.25%	73.68%
July	2012	240	447	394.38	31.11	125	56.80%	75.62%
July	2013	240	449	389.37	40.40	104	49.04%	73.05%
July	2014	240	471	381.80	46.44	154	46.75%	74.04%
Oct	2008	240	538	392.54	30.94	1154	54.59%	64.32%
Oct	2009	240	468	392.63	34.52	700	58.14%	71.27%
Oct	2010	240	500	398.16	30.95	715	65.87%	73.62%
Oct	2011	240	482	399.78	33.35	567	68.08%	73.68%
Oct	2012	240	507	402.92	35.30	587	75.98%	75.62%
Oct	2013	240	510	395.25	36.60	717	62.20%	73.05%
Oct	2014	240	479	392.30	38.55	847	56.67%	74.04%

The PARCC Algebra I equivalent cut score based on PSAT linking produced the lowest cut score which leads to a passing rate slightly lower than the lower bound of the yearly passing rate but below the range of the May passing rates. On the other hand, the PARCC cut scores obtained based on propensity score matching produced even higher cut scores yielding even lower passing rates when compared with both the May and yearly HSA passing rates for Algebra. Compared with the propensity score matching

method, the PSAT linking produced a lower PARCC equivalent cut score which leads to a higher passing rate for Algebra.

Table 2.28
Passing Rates for the HSA Algebra Test

Month	Year	Min	Max	Mean	SD	N	%pass	year %pass
Jan.	2008	240	650	391.65	37.65	11210	26.39%	65.51%
Jan.	2009	240	538	400.24	40.63	6272	41.65%	67.03%
Jan	2010	240	540	401.42	41.02	5057	44.24%	66.98%
Jan	2011	240	650	408.50	46.16	3245	56.12%	72.88%
Jan	2012	240	522	401.74	47.18	3318	49.46%	73.77%
Jan	2013	240	650	410.06	42.49	2852	57.43%	71.59%
Jan	2014	240	502	402.30	48.21	2789	52.35%	66.88%
April	2009	240	460	376.12	45.84	195	17.95%	67.03%
April	2010	240	478	386.84	43.06	164	26.22%	66.98%
April	2011	240	488	394.48	44.13	88	34.09%	72.88%
April	2012	256	499	408.42	48.58	59	55.93%	73.77%
April	2013	240	509	415.43	40.58	79	53.16%	71.59%
April	2014	295	471	394.77	39.59	48	31.25%	66.88%
May	2008	240	650	428.63	37.25	69227	72.59%	65.51%
May	2009	240	650	427.26	41.93	73165	69.88%	67.03%
May	2010	240	650	426.13	40.48	64195	69.11%	66.98%
May	2011	240	650	431.90	39.55	57107	74.08%	72.88%
May	2012	240	650	428.90	39.98	58817	75.23%	73.77%
May	2013	240	650	428.38	38.93	62026	72.33%	71.59%
May	2014	240	650	421.99	43.32	55817	67.70%	66.88%
July	2008	240	500	401.25	50.38	321	48.91%	65.51%
July	2009	240	486	412.33	43.05	161	55.28%	67.03%
July	2010	240	501	407.16	50.64	114	56.14%	66.98%
July	2011	240	500	417.00	48.15	85	64.71%	72.88%
July	2012	240	489	415.85	50.72	96	63.54%	73.77%
July	2013	240	489	417.21	46.94	70	65.71%	71.59%
July	2014	240	469	391.84	51.64	80	42.50%	66.88%
Oct.	2008	240	516	396.61	41.56	1355	41.92%	65.51%
Oct	2009	240	650	401.75	47.04	698	47.99%	67.03%
Oct	2010	240	498	405.89	42.46	513	55.36%	66.98%
Oct	2011	240	506	413.09	47.31	388	65.21%	72.88%
Oct	2012	240	540	409.65	50.07	325	61.85%	73.77%
Oct	2013	240	524	398.66	57.04	359	55.99%	71.59%
Oct	2014	240	519	394.84	52.11	773	45.15%	66.88%

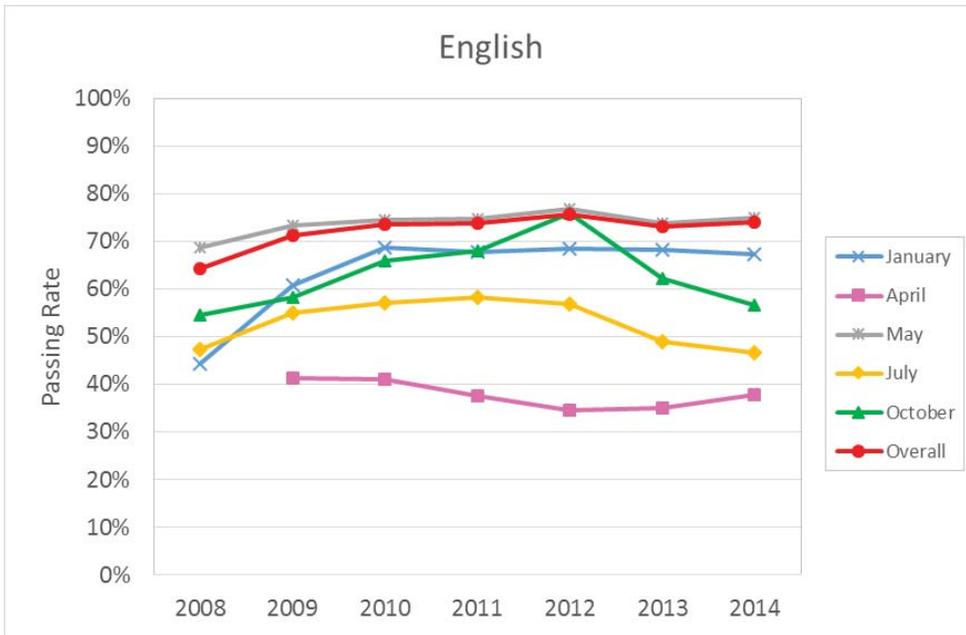


Figure 1. *Passing Rates for the HSA English Test*

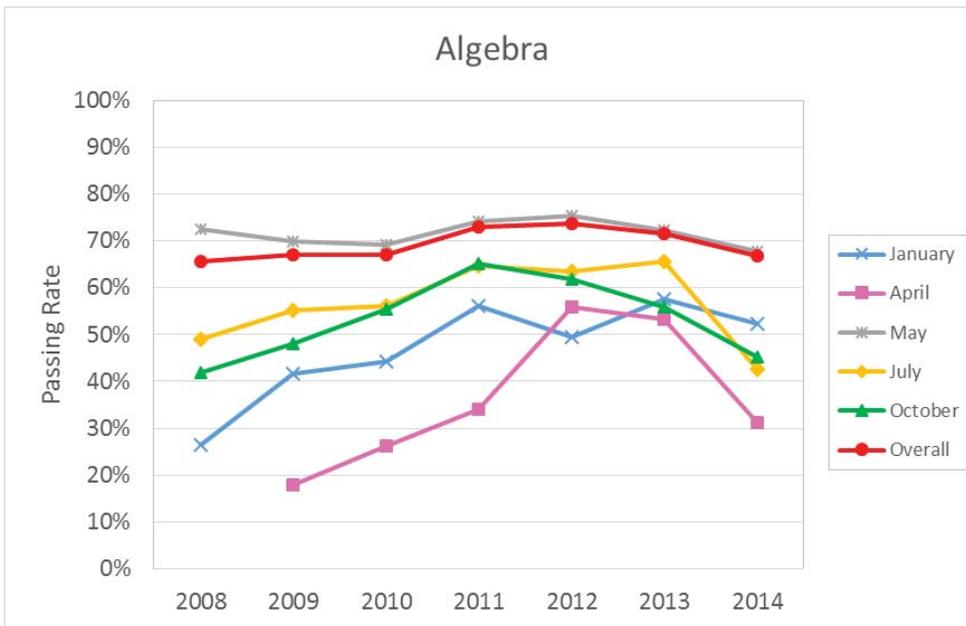


Figure 2. *Passing Rates for the HSA Algebra Test*

To further investigate the relationship between the mapped PARCC equivalents of HSA cut scores and the PARCC cut scores, especially the cut score that divides performance level 2 from 3 (a PARCC score of 725 for both ELA10 and Algebra I), the conditional standard error of measurement (CSEM) for the mapped PARCC cut score is utilized to construct a 95% confidence interval and 1 standard deviation above and below

the mapped cut scores using different methods. As multiple forms were constructed for the PARCC tests, the CSEM for the same PARCC score could be different for different forms. Thus, the mean, minimum, and maximum CSEM are used to construct the intervals respectively. The two intervals around the PARCC equivalent cut scores obtained using PSAT for linking are summarized in Tables 2.29. For ELA10, the 95% confidence interval around the mapped PARCC equivalent score of the HSA cut score using the mean and the maximum CSEM contained the PARCC cut score of 725 dividing level 2 and 3 while the interval one standard deviation above and below the CSEM did not contain the mapped PARCC cut score. Neither does the 95% confidence interval using the minimum CSEM. For Algebra I, all intervals contained the PARCC cut score of 725 as seen in Table 2.29. Similar patterns were found for the cut scores obtained using the propensity score matching method as shown in Tables 2.30 and 2.31.

Table 2.29
95% Confidence Intervals and One Standard Deviation above and below the Mapped PARCC Equivalent Cut Scores for Option I Using PSAT for Linking

Subject	Cut Score	Mean CSEM	Minimum CSEM	Maximum CSEM	95% CI	1 SD	95% CI	1 SD	95% CI	1 SD
					Mean CSEM	Mean CSEM	Minimum CSEM	Minimum CSEM	Maximum CSEM	Maximum CSEM
ELA10	707	9.80	8	10.7	(688, 726)	(697, 717)	(691, 723)	(699, 715)	(686, 728)	(696, 718)
Algebra I	720	10.73	9	11.8	(699, 741)	(709, 731)	(702, 738)	(711, 729)	(697, 743)	(708, 732)

Table 2.30
95% Confidence Intervals and One Standard Deviation above and below the Mapped PARCC Equivalent Cut Scores for Option II Using Propensity Score Matching

Subject	Cut Score	Mean CSEM	Minimum CSEM	Maximum CSEM	95% CI	1 SD	95% CI	1 SD	95% CI	1 SD
					Mean CSEM	Mean CSEM	Minimum CSEM	Minimum CSEM	Maximum CSEM	Maximum CSEM
ELA10	704	10.11	8.1	10.9	(684, 724)	(694, 714)	(688, 720)	(696, 712)	(683, 725)	(693, 715)
	705	9.99	9.2	10.8	(685, 725)	(695, 715)	(687, 723)	(696, 714)	(684, 726)	(694, 716)
	706	9.96	8.0	10.6	(686, 726)	(696, 716)	(690, 722)	(698, 714)	(685, 727)	(695, 717)
Algebra I	721	10.37	8.9	13.4	(701, 741)	(711, 731)	(704, 738)	(712, 730)	(695, 747)	(708, 734)
	725	9.81	8.5	10.9	(706, 744)	(715, 735)	(708, 742)	(717, 734)	(704, 746)	(714, 736)

Table 2.31
95% Confidence Intervals and One Standard Deviation above and below the Mapped PARCC Equivalent Cut Scores for Option II (Combining Design II and III) Using Propensity Score Matching

Subject	Cut Score	Mean CSEM	Minimum CSEM	Maximum CSEM	95% CI	1 SD	95% CI	1 SD	95% CI	1 SD
					Mean CSEM	Mean CSEM	Minimum CSEM	Minimum CSEM	Maximum CSEM	Maximum CSEM
Algebra I	722	10.15	8.8	12.5	(702, 742)	(712, 732)	(705, 739)	(713, 731)	(698, 747)	(710, 735)
	724	10.14	8.7	11.3	(704, 744)	(714, 734)	(707, 741)	(715, 733)	(702, 746)	(713, 735)
	727	9.70	8.4	10.7	(708, 746)	(717, 737)	(711, 743)	(719, 735)	(706, 748)	(716, 738)

In addition, the HSA equivalents of the PARCC cut score of 725 dividing performance level 2 from 3 are summarized in Table 2.32 when using PSAT for linking and in Table 2.33 for propensity score matching. When propensity score matching was used, Design I condition 4 for ELA10 did not have a PARCC score of 725 corresponding to a HSA test score. A reversed mapping was implemented to find a HSA equivalent of a PARCC score of 725. For the condition with Design II and III combined for Algebra I, two HSA scores were equivalent to a PARCC score of 725. Thus a reversed mapping was also implemented to find a single HSA equivalent score of a 725 PARCC cut score. In general, the HSA equivalents of the PARCC cut score, 725 for both ELA10 and Algebra I were higher than the original HSA cut scores.

Table 2.32
HSA Equivalent Scores of the PARCC Cut Score of 725 for Dividing Performance Level 2 from 3 (Option I Using PSAT for Linking)

Subject	HSA
English	408
Algebra	419

Table 2.33
HSA Equivalent Scores of the PARCC Cut Score of 725 for Dividing Performance Level 2 from 3 (Option II Based on Propensity Score Matching)

Condition	1	2	3	4
Design I (English)	407	407	407	407
Design II (Math)	416	417	417	417
Design III (Math)	412	412	412	412
Design II & III Combined	415	410	415	413

Summary

This study explored two methods of obtaining the PARCC equivalent scores of the HSA cut scores for PARCC ELA10 and Algebra I, and vice versa. One method used PSAT as an external linking test to link HSA and PARCC based on a two-step single group linking design. Specifically, the HSA English and Algebra tests were linked to the PSAT Verbal and Math tests respectively and then the PSAT tests were linked to the corresponding PARCC tests. Based on the first-time test-takers' scores, the corresponding PARCC Algebra I score to the HSA Algebra passing score of 412 is **720** and the corresponding PARCC ELA10 score to the HSA English passing score of 396 is **707**.

The other method uses propensity score matching to come up with equivalent groups between students taking HSA and PARCC. Four matching conditions were explored based on the use of different caliper values and the use of replacement of cases for each design. The absolute standardized mean difference values for each covariate indicate the matched samples were relatively equivalent. Among the 16 designs, the

corresponding PARCC ELA10 equivalent scores of the HSA English passing score are 704, 705, and 706 while the corresponding PARCC Algebra I scores equivalent to the HSA Algebra passing scores are 721 and 725 for Design II and III respectively, and 722, 724, and 727 for the combined Design II and III samples (refers to Table 2.5).

Two intervals, 95% confidence intervals and one standard deviation above and below the PARCC equivalents of the HSA cut scores were also constructed. For ELA10, the 95% confidence interval around the mapped PARCC equivalent score of the HSA cut score using the mean and the maximum CSEM captured the PARCC cut score of 725 between performance level 2 vs. 3 while the interval one standard deviation above and below the CSEM did not capture the mapped PARCC cut score. Neither does the 95% confidence interval using the minimum CSEM. For Algebra I, all intervals captured the PARCC cut score of 725. The patterns were consistent across linking methods.

The HSA equivalents of the PARCC cut score of 725 dividing performance level 2 from 3 are summarized. In general, the HSA equivalents of the PARCC cut score, 725 for both ELA10 and Algebra were higher than the original HSA cut scores.

This study provides empirical evidence about the PARCC equivalents of the HSA cut scores and the HSA equivalents of the PARCC cut score of 725 between performance level 2 vs. 3 for ELA10 and Algebra I. The final adoption of cut scores obtained in this study depends on considerations from psychometric, policy, and practical perspectives.

References

- Cui, Z., & Kolen, M. J. (2009). Evaluation of Two New Smoothing Methods in Equating: The Cubic B-Spline Presmoothing Method and the Direct Presmoothing Method. *Journal of Educational Measurement, 46*(2), 135-158.
- Hanson, B. A. (1994). A Comparison of Presmoothing and Postsmoothing Methods in Equipercentile Equating. ACT Research Report Series 94-4.
- Ho, D., Imai, K., & Imai, M. K. (2013). Package 'MatchIt'. Retrieved June, 23, 2014.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking* (pp. 201-205). New York: Springer.
- Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics, 25*(1), 1.

Graduation Assessment Requirements

Maryland State Board of Education
Assessment Update
February 23, 2016

Background

- The High School Assessments (HSAs) have been part of the Maryland graduation requirements since 2003.
- During the 2014-2015 school year, the PARCC Algebra I and English 10 tests replaced the HSA Algebra/Data Analysis and English2 tests
- The HSAs are on a scale ranging from 240 score points to 650 score points
 - 412 - Algebra/Data Analysis passing score
 - 396 - English2 passing score
- The PARCC assessments are on a scale from 650-850 score points

Background con't

PARCC Scale Scores (SSs)	PARCC Performance Levels (PLs)	PL Descriptors (PLDs)
800	Level 5*	Exceeded expectations
750	Level 4*	Met expectations
725	Level 3	Approached expectations
700	Level 2	Partially met expectations
650	Level 1	Did not yet meet expectations

* PL 4 and PL 5 denote on track for College and Career Readiness (CCR)

Research Questions?

- What PARCC Algebra I scale score correlates to the passing Algebra/Data Analysis HSA 412 scale score?
- What PARCC English 10 scale score correlates to the passing English2 HSA 396 scale score?

Purpose

- To share the results of a research study conducted by the Maryland Assessment Research Center (MARC) equating the PARCC Algebra I (ALG I) and PARCC English 10 (ELA 10) scale scores onto the HSA scales.
- Provide a recommendation as to what PARCC scale score/performance level should be used to satisfy the assessment graduation requirement for Algebra I and English 10

Study Findings

- Two equating methods produced similar equivalent scale scores when linking HSA and PARCC tests.
 - The PARCC ALG I scale score of **720** corresponds with the HSA Algebra passing score of 412.
 - The PARCC ELA 10 scale score of **707** corresponds with the HSA English passing score of 396.

Summary of Psychometric Analysis

Content	HSA Passing Score	PARCC Equivalent Score	PARCC Confidence Interval	
ALG I	412	720	697	743
ELA 10	396	707	686	728

PARCC Performance Levels

PARCC Scale Scores (SSs)	PARCC Performance Levels (PLs)	PL Descriptors (<u>PLDs</u>)
800	Level 5 [^]	Exceeded expectations
750	Level 4 [^]	Met expectations
725	Level 3	Approached expectations
700	Level 2	Partially met expectations
650	Level 1	Did not yet meet expectations

[^] PL 4 and PL 5 denote College and Career Readiness (CCR) or on track for CCR

Passing Rates by PARCC Algebra I Performance Level and Scale Score

2014-2015 Maryland PARCC Assessment Data

Scale Score	650 PL-1	700 PL-2	720* PL-3	725 PL-3	750^ PL-4	800^ PL-5
Passing Rate	100%	87%	65%	59%	30%	10%

* Equivalent score to HSA minimum passing 396 scale score

^ PARCC Performance Level (PL) 4 and 5 denote on track for College and Career Readiness



Passing Rates by PARCC English 10 Performance Level and Scale Score

2014-2015 Maryland PARCC Assessment Data

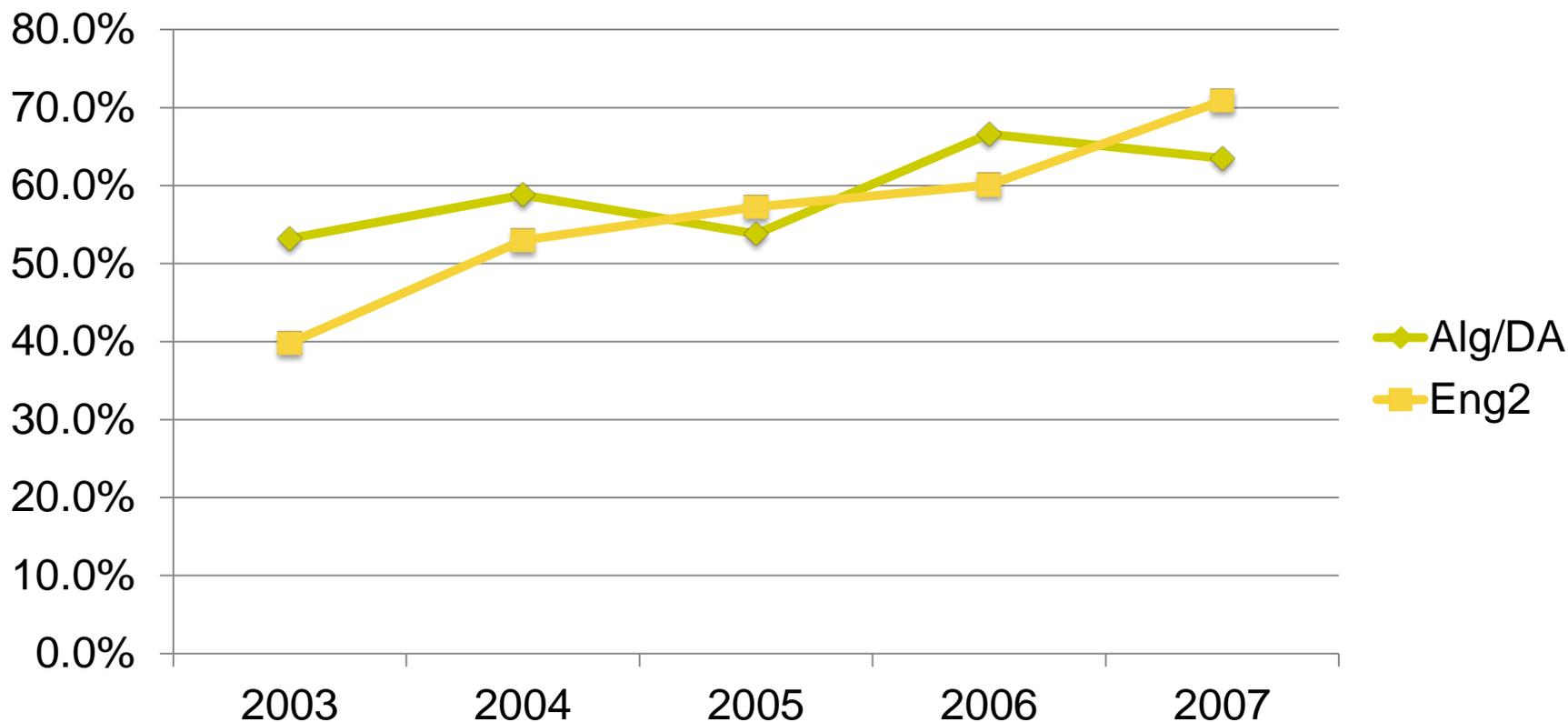
Scale Score	650 PL-1	700 PL-2	707* PL-3	725 PL-3	750^ PL-4	800^ PL-5
Passing Rate	100%	79%	74%	60%	35%	10%

* Equivalent score to HSA minimum passing 396 scale score

^ PARCC Performance Level (PL) 4 and 5 denote College and Career Readiness



Historical Trends for Initial HSAs



MSDE's Recommendations

- Given psychometric analysis, understanding performance trends when introducing new assessments, and practical considerations of aligning to the PARCC performance level descriptors, MSDE recommends:
 - PARCC **725** as the passing score for ALG I
 - PARCC **725** as the passing score for ELA 10
- 725 corresponds with PARCC Performance Level 3

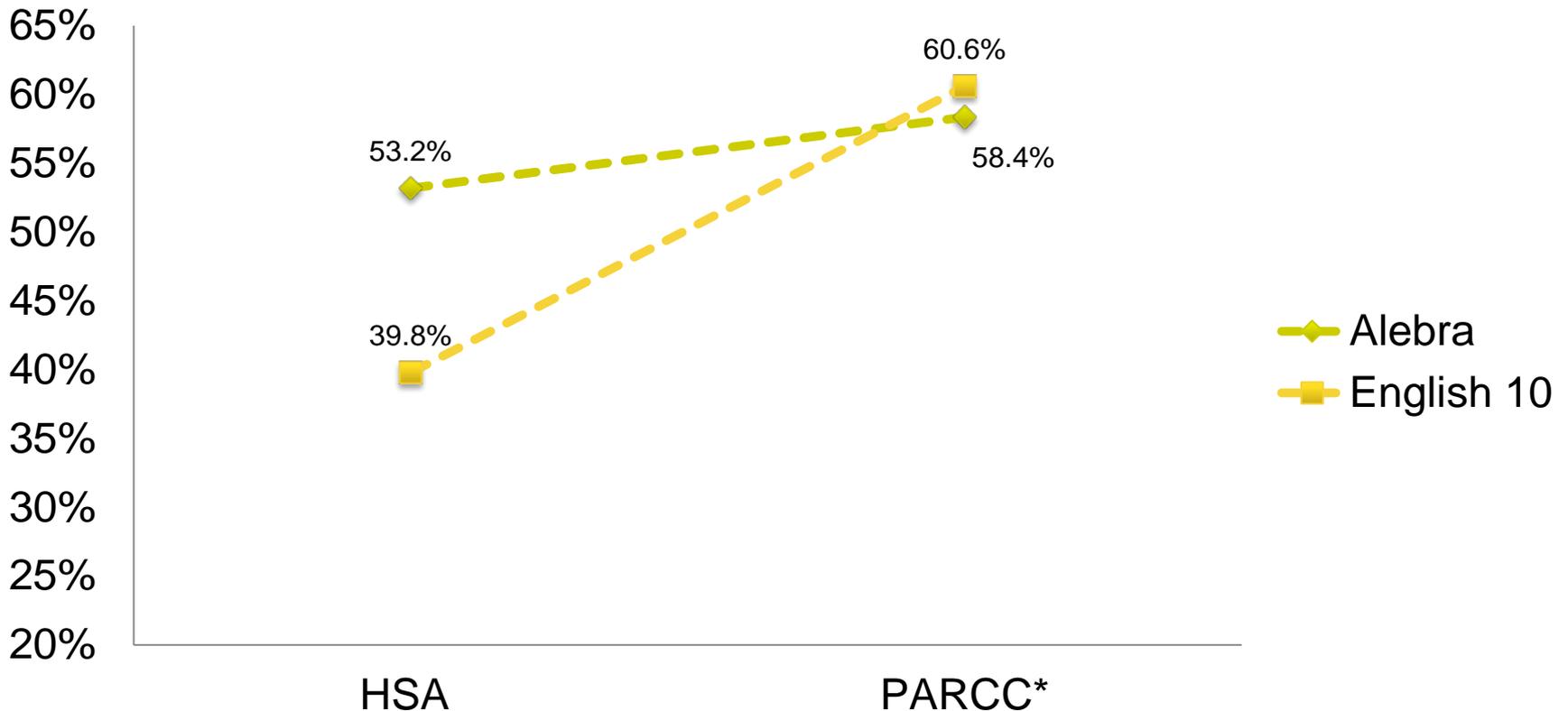
Recommendation

PARCC Scale Scores (SSs)	PARCC Performance Levels (PLs)	PL Descriptors (PLDs)
800	Level 5*	Exceeded expectations
750	Level 4*	Met expectations
725	Level 3 ⁺	Approached expectations
700	Level 2	Partially met expectations
650	Level 1	Did not yet meet expectations

* PL 4 and PL 5 denote College and Career Readiness (CCR) or on track for CCR

+ Proposed PL to satisfy graduation requirement for [PARCC](#) Algebra I and English 10

First Administration Pass Rates



*PARCC Pass Rate based on a PARCC score of 725

Next Steps...

MSDE BOE Meetings	Actions
March 2016	<ul style="list-style-type: none">• Address additional questions and comments from MSDE BOE members• Introduce new updated language for COMAR
April 2016	Vote to approve COMAR for publication and public comment