

**Kansas Assessment Program
Technical Manual Addendum
2019**

University of Kansas Achievement & Assessment Institute

July 2019

Table of Contents

Table of Contents	2
I. Test Sample	2
Table I-1. Test Sample for 2019 Administration	2
II. Item Analysis	3
II-1. Classical Item Statistics	3
Table II-1. Classical Item Difficulties for ELA	3
Table II-2. Classical Item Difficulties for Mathematics	3
Table II-3. Classical Item Difficulties for Science	3
Table II-4. Classical Item Discrimination for ELA	4
Table II-5. Classical Item Discrimination for Mathematics	4
Table II-6. Classical Item Discrimination for Science.....	4
II-2. Item Response Theory Item Statistics	4
Table II-7. Item Response Theory Item Difficulty for ELA.....	5
Table II-8. Item Response Theory Item Difficulty for Mathematics.....	5
Table II-9. Item Response Theory Item Difficulty for Science	5
Table II-10. Item Response Theory Item Discrimination for ELA	6
Table II-11. Item Response Theory Item Discrimination for Mathematics.....	6
Table II-12. Item Response Theory Item Discrimination for Science.....	6
II-3. Cognitive Complexity	6
Table II-13. Percentage of Items by Depth of Knowledge Level	7
III. Test Analysis	8
III-1. Operational Test Results	8
Table III-1. Scale-Score Descriptive Statistics for ELA.....	8
Table III-2. Scale-Score Descriptive Statistics for Mathematics.....	8
Table III-3. Scale-Score Descriptive Statistics for Science	8
Table III-4. Percentage of Students Achieving at Each Performance Level for ELA and Mathematics	9
Table III-5. Percentage of Students Achieving at Each Performance Level for Science ...	9
Figure III-1. Performance-level results for ELA.	10
Figure III-2. Performance-level results for mathematics.....	10
Figure III-3. Performance-level results for science.	11
III-2. Test Result Trends	11
Table III-6. Longitudinal Scale-Score Trend for ELA	11

Table III-7. Longitudinal Scale-Score Trend for Mathematics	12
Table III-8. Longitudinal Scale-Score Trend for Science.....	12
Figure III-4. Performance-level trend for ELA.....	13
Figure III-5. College- and career-readiness trend for ELA.....	14
Figure III-6. Performance-level trend for mathematics.	15
Figure III-7. College- and career-readiness trend for mathematics.	16
IV. Reliability	21
Table IV-1. Test Reliability by Subject and Grade.....	21
Table IV-2. Classification Consistency and Accuracy	22
Table IV-3. Grade 3 Subgroup Reliability and Performance for ELA	23
Table IV-4. Grade 4 Subgroup Reliability and Performance for ELA	24
Table IV-5. Grade 5 Subgroup Reliability and Performance for ELA	25
Table IV-6. Grade 6 Subgroup Reliability and Performance for ELA	26
Table IV-7. Grade 7 Subgroup Reliability and Performance for ELA	27
Table IV-8. Grade 8 Subgroup Reliability and Performance for ELA.....	28
Table IV-9. Grade 10 Subgroup Reliability and Performance for ELA	29
Table IV-10. Grade 3 Subgroup Reliability and Performance for Mathematics	30
Table IV-11. Grade 4 Subgroup Reliability and Performance for Mathematics	31
Table IV-12. Grade 5 Subgroup Reliability and Performance for Mathematics	32
Table IV-13. Grade 6 Subgroup Reliability and Performance for Mathematics	33
Table IV-14. Grade 7 Subgroup Reliability and Performance for Mathematics	34
Table IV-15. Grade 8 Subgroup Reliability and Performance for Mathematics	35
Table IV-16. Grade 10 Subgroup Reliability and Performance for Mathematics	36
Table IV-17. Grade 5 Subgroup Reliability and Performance for Science	37
Table IV-18. Grade 8 Subgroup Reliability and Performance for Science	38
Table IV-19. Grade 11 Subgroup Reliability and Performance for Science	39
Table IV-20. Path Reliability for ELA by Grade.....	39
Table IV-21. Subscore Reliability, Classification Consistency, and Accuracy.....	40
V. Validity Evidence	41
V-1. Evidence Based on Internal Structure.....	41
Table V-1. DIF for ELA Across Subgroups	41
Table V-2. DIF for Mathematics Across Subgroups	41
Table V-3. DIF for Science Across Subgroups	42
V-2. Evidence Based on Relationships to Other Variables	42

Table V-4. Correlations Between ELA, Mathematics, and Science Scores by Grade	42
Table V-5. Correlations Between KAP Scores (2017) and ACT (2017–2018) Scores	42
References	43

This 2019 Kansas Assessment Program (KAP) technical manual addendum is an update of the 2017 KAP technical manual. It provides statistical information based on the spring 2019 administration of English language arts (ELA), mathematics, and science. For more details, refer to the 2017 technical manual.

A number of changes to test content were made in 2019 and are described below.

- ❖ Some mathematics content standards changed, and test administration changed from adaptive with two blocks in 2018 to fixed forms in 2019.
 - Items that did not align with new standards were removed. In some cases, entire clusters (i.e., targets) were removed.
 - The new, fixed forms were built from the pool of operational items from the 2018 accommodated forms that were also aligned with the new standards.
 - Grade 3 and grade 4 had minor changes to the test forms, including the removal of a cluster in grade 4. Grade 8 and grade 10 had the most changes to the test forms and content coverage.
- ❖ Science had item-pool changes. This included removing items at all three grade levels that had poor statistics (e.g., a negative a value), were simulation based, or did not match the Next Generation Science Standards. Blueprint coverage was maintained in each grade despite removing items.
- ❖ ELA assessments are multistage adaptive forms. The first stage covers the entire range of difficulty and directed students to either the easy or difficult stage 2 test. The forms were consistent across years, with the exception of grade 6. When 2018 performance was reviewed, it was noticed that, for grade 6, students who were routed into the easier form had no opportunity to achieve at performance level 3. In other words, students routed to the easier form could not demonstrate proficiency. Thus, seven writing items on the grade 6, stage 2 easy form were replaced with new items to make it possible for students to demonstrate proficiency.

I. Test Sample

Assessments were administered in ELA, mathematics, and science in grades 3 through 8 and high school. In high school, students complete ELA and mathematics assessments in grade 10 and science assessments in grade 11.

Table I-1. *Test Sample for 2019 Administration by Grade and Subject*

Grade	ELA (<i>n</i>)	Mathematics (<i>n</i>)	Science (<i>n</i>)
3	37,098	37,184	
4	37,698	37,771	
5	38,372	38,413	38,442
6	38,281	38,329	
7	37,424	37,456	
8	36,779	36,785	36,863
HS	36,318	36,287	34,081

Note. HS = high school.

II. Item Analysis

II-1. Classical Item Statistics

Summaries of item difficulties for ELA, mathematics, and science tests are presented in Tables II-1 through II-3. The ELA grade-level average item difficulties range from .51 to .54, the mathematics grade-level average item difficulties range from .43 to .53, and the science grade-level average item difficulties range from .50 to .58, indicating that overall items are moderately challenging.

Table II-1. *Classical Item Difficulties for ELA*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
3	86	.51	.16	.14	.41	.53	.61	.85
4	82	.51	.17	.16	.39	.50	.62	.92
5	81	.53	.13	.15	.46	.53	.62	.77
6	74	.51	.18	.13	.40	.53	.64	.89
7	74	.54	.17	.22	.43	.50	.66	.94
8	80	.54	.16	.11	.47	.57	.66	.81
10	68	.53	.17	.02	.43	.53	.65	.85

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

Table II-2. *Classical Item Difficulties for Mathematics*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
3	55	.53	.22	.05	.38	.56	.69	.89
4	55	.48	.18	.17	.33	.48	.60	.89
5	55	.50	.17	.08	.38	.52	.63	.82
6	55	.48	.17	.11	.39	.48	.59	.87
7	55	.48	.17	.17	.36	.46	.58	.83
8	55	.46	.20	.02	.30	.48	.64	.87
10	55	.43	.17	.05	.31	.45	.56	.86

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

Table II-3. *Classical Item Difficulties for Science*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
5	33	.58	.17	.26	.47	.57	.72	.91
8	41	.51	.12	.25	.46	.51	.59	.78
11	38	.50	.13	.20	.42	.51	.59	.78

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

Tables II-4 through II-6 present item discrimination for ELA, mathematics, and science. The medians of item discrimination for ELA range from .33 to .38, from .40 to .47 for mathematics, and .34 to .37 for science, indicating that items are moderately discriminating overall.

Table II-4. *Classical Item Discrimination for ELA*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
3	86	.37	.08	.10	.31	.36	.41	.57
4	82	.35	.09	.08	.30	.36	.41	.59
5	81	.34	.09	.06	.29	.35	.40	.54
6	74	.38	.09	.21	.31	.38	.44	.55
7	74	.33	.08	.13	.27	.32	.38	.55
8	80	.33	.12	-.06	.24	.33	.42	.53
10	68	.36	.12	.13	.29	.36	.44	.64

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

Table II-5. *Classical Item Discrimination for Mathematics*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
3	55	.44	.11	.05	.38	.47	.51	.61
4	55	.47	.08	.34	.40	.46	.52	.65
5	55	.46	.08	.29	.41	.45	.51	.63
6	55	.45	.09	.23	.38	.45	.52	.64
7	55	.41	.10	.24	.34	.41	.48	.60
8	55	.40	.11	.14	.33	.42	.48	.61
10	55	.41	.12	.07	.34	.40	.51	.60

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

Table II-6. *Classical Item Discrimination for Science*

Grade	Items (<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	Median	P ₇₅	Max.
5	33	.36	.08	.18	.31	.39	.41	.52
8	41	.34	.08	.08	.30	.36	.39	.46
11	38	.37	.10	.17	.31	.40	.44	.53

Note. P₂₅ and P₇₅ = 25th and 75th percentiles, respectively.

II-2. Item Response Theory (IRT) Item Statistics

Tables II-7 through II-12 summarize the IRT item statistics, difficulty (i.e., *b* parameter) and discrimination (i.e., *a* parameter) estimates, of operational items in ELA, mathematics, and science tests, respectively. Most items are dichotomous, but some items have more than one score category (thus, more than one *b* parameters yet only one *a* parameter); therefore, the numbers of *b* and *a* parameters are different in these tables. Parameters for all items, irrespective of the number of score categories, are included in the tables. The mean item difficulty fluctuates across grades. The grade-10 mathematics test has the highest *b* value (i.e., .34). A large standard deviation (*SD*) of difficulty parameters indicates a large variability of item difficulties. The mean

item discrimination also fluctuates across grades. For mathematics, it fluctuates but shows the trend of decreasing as the grade increases.

Table II-7. *Item Response Theory Item Difficulty for ELA*

Grade	Number of items	<i>b</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
3	101	-.31	1.29	-4.80	-1.08	-.35	.41	2.50
4	95	-.62	1.27	-3.87	-1.55	-.70	.14	3.15
5	94	-.33	1.29	-4.69	-0.98	-.50	.48	2.58
6	93	-.69	1.55	-4.27	-1.50	-.79	.37	3.48
7	97	-.21	2.03	-5.03	-1.26	-.64	.82	7.94
8	101	-.48	1.47	-5.55	-1.18	-.52	.35	2.92
10	89	-.35	1.54	-3.86	-1.34	-.50	.53	4.26

Note. *b* = difficulty parameter; Q1 = first quartile; Q3 = third quartile.

Table II-8. *Item Response Theory Item Difficulty for Mathematics*

Grade	Number of items	<i>b</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
3	61	-.37	1.24	-3.21	-1.16	-.38	0.51	2.66
4	78	-.14	1.41	-3.45	-1.00	-.07	0.75	4.08
5	72	-.33	1.30	-3.65	-0.88	-.32	0.36	3.38
6	61	-.08	1.10	-3.00	-0.62	-.17	0.49	1.97
7	72	-.18	1.66	-4.67	-0.87	-.10	0.85	4.55
8	78	-.13	2.06	-6.28	-1.09	.00	1.18	4.29
10	70	.34	1.64	-3.03	-0.51	.22	1.26	5.54

Note. *b* = difficulty parameter; Q1 = first quartile; Q3 = third quartile.

Table II-9. *Item Response Theory Item Difficulty for Science*

Grade	Number of items	<i>b</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
5	35	-.68	1.66	-5.10	-1.32	-.56	-.06	2.91
8	46	.40	6.24	-6.09	-0.87	-.15	.35	40.24
11	38	.15	1.39	-1.39	-0.65	-.25	.45	5.93

Note. *b* = difficulty parameter; Q1 = first quartile; Q3 = third quartile.

Table II-10. *Item Response Theory Item Discrimination for ELA*

Grade	Number of items	<i>a</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
3	86	1.04	.36	.26	.77	0.99	1.30	1.92
4	82	0.98	.34	.34	.76	0.91	1.22	1.99
5	81	0.91	.32	.40	.71	0.90	1.06	2.43
6	74	1.02	.36	.36	.78	1.01	1.16	2.01
7	74	0.89	.41	.26	.59	0.80	1.18	2.11
8	80	0.91	.35	.29	.67	0.91	1.08	2.00
10	68	0.94	.32	.35	.72	0.94	1.12	2.21

Note. *a* = discrimination parameter; Q1 = first quartile; Q3 = third quartile.

Table II-11. *Item Response Theory Item Discrimination for Mathematics*

Grade	Number of items	<i>a</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
3	54	1.18	.33	.64	0.96	1.14	1.38	2.11
4	55	1.16	.34	.64	0.86	1.12	1.35	2.02
5	55	1.19	.29	.56	1.02	1.16	1.41	1.97
6	55	1.18	.43	.43	0.89	1.12	1.44	2.36
7	55	1.01	.36	.51	0.74	0.96	1.21	2.11
8	55	0.99	.40	.12	0.71	0.96	1.22	2.04
10	55	1.05	.41	.33	0.80	0.95	1.29	2.17

Note. *a* = discrimination parameter; Q1 = first quartile; Q3 = third quartile.

Table II-12. *Item Response Theory Item Discrimination for Science*

Grade	Number of items	<i>a</i>						
		<i>M</i>	<i>SD</i>	Min.	Q1	Median	Q3	Max.
5	33	.81	.31	.27	.65	.8	.97	1.51
8	41	.68	.23	.02	.54	.71	.86	1.08
11	38	.84	.37	.26	.54	.87	.98	1.72

Note. *a* = discrimination parameter; Q1 = first quartile; Q3 = third quartile.

II-3. Cognitive Complexity

Table II-13 shows the percentage of operational items by depth of knowledge (DOK) level, subject, and grade. This information reveals the proportions of DOK requirements according to content standards. Most ELA and mathematics items are at DOK levels 1 and 2; fewer items are at level 3. For science, most items are at DOK levels 2 and 3, with a few items at level 1. There are no level 4 items.

Table II-13. *Percentage of Items by Depth of Knowledge Level*

G	ELA					Mathematics					Science				
	TI	DOK level (%)				TI	DOK level (%)				TI	DOK level (%)			
		1	2	3	4		1	2	3	4		1	2	3	4
3	86	25	54	7	0	55	24	30	1	0					
4	82	20	52	10	0	55	19	33	2	0					
5	81	23	46	12	0	55	25	30	0	0	33	3	18	12	0
6	74	22	39	13	0	55	22	32	1	0					
7	74	9	52	13	0	55	26	27	2	0					
8	80	20	51	9	0	55	13	39	3	0	41	4	17	20	0
HS	68	16	48	4	0	55	23	28	4	0	38	0	20	16	0

Note. G = grade; TI = total number of items.

III. Test Analysis

III-1. Operational Test Results

Summaries of scale scores, ranging from 220 to 380, are presented in Tables III-1 through III-3 by subject and grade. The differences between (a) P₅₀ and P₂₅ and (b) P₇₅ and P₅₀ are indicators of the shapes of score distributions: the larger of the two differences indicates the direction of any skewness in the distribution (i.e., a negative skew when the first difference is larger and a positive skew when the second difference is larger). If the two differences match, the distribution is symmetric. For example, in ELA, the distribution for grade 5 is symmetric in shape, and distributions for the other grades are slightly positively skewed.

Table III-1. *Scale-Score Descriptive Statistics for ELA*

Grade	<i>M</i>	<i>SD</i>	Min.	P ₁₀	P ₂₅	P ₅₀	P ₇₅	P ₉₀	Max.
3	294.8	29.1	220	257	274	293	316	335	380
4	299.1	28.0	220	262	279	298	318	337	380
5	295.9	29.4	220	257	274	295	316	335	380
6	290.4	28.7	220	252	269	290	312	327	380
7	288.3	31.1	220	250	265	286	310	330	380
8	282.3	28.5	220	246	262	281	301	321	380
10	283.4	29.8	220	245	261	282	304	323	380

Note. P₁₀, P₂₅, P₅₀, P₇₅, and P₉₀ = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

Table III-2. *Scale-Score Descriptive Statistics for Mathematics*

Grade	<i>M</i>	<i>SD</i>	Min.	P ₁₀	P ₂₅	P ₅₀	P ₇₅	P ₉₀	Max.
3	303.2	27.9	220	268	282	301	322	342	380
4	292.9	28.6	220	258	271	291	311	332	380
5	290.9	27.1	220	259	271	287	307	328	380
6	291.6	27.4	220	261	270	287	307	331	380
7	288.2	28.1	220	256	268	283	304	326	380
8	286.2	28.8	220	254	266	283	301	326	380
10	286.1	27.9	220	257	266	280	301	326	380

Note. P₁₀, P₂₅, P₅₀, P₇₅, and P₉₀ = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

Table III-3. *Scale-Score Descriptive Statistics for Science*

Grade	<i>M</i>	<i>SD</i>	Min.	P ₁₀	P ₂₅	P ₅₀	P ₇₅	P ₉₀	Max.
5	298.7	30.8	220	263	276	299	321	335	380
8	286.9	30.3	220	252	265	285	306	325	380
11	289.3	29.8	220	254	266	286	305	332	380

Note. P₁₀, P₂₅, P₅₀, P₇₅, and P₉₀ = 10th, 25th, 50th, 75th, and 90th percentiles, respectively.

The proportion of students achieving at each performance level (level 1 through level 4) and the college- and career-ready rating (combined level 3 and level 4) are provided by subject and grade in Tables III-4 through III-5 and Figures III-1 through III-3. The readiness rates range from 25%

to 52% across subjects and grades. All three subjects tend to have lower readiness rates in higher grade levels.

Table III-4. *Percentage of Students Achieving at Each Performance Level for ELA and Mathematics*

Grade	ELA PL (%)					Mathematics PL (%)				
	1	2	3	4	CCR	1	2	3	4	CCR
3	28	31	28	14	42	17	31	34	17	52
4	16	35	39	10	49	17	46	26	11	37
5	25	30	29	15	45	28	39	22	11	33
6	33	27	35	5	40	29	38	23	10	33
7	35	32	24	9	33	23	48	25	5	29
8	29	45	22	5	26	38	35	20	7	27
10	34	37	24	5	29	40	34	17	8	25

Note. CCR = college and career ready (combination of performance levels 3 and 4); PL = performance level. Column percentages may not total 100% because of rounding.

Table III-5. *Percentage of Students Achieving at Each Performance Level for Science*

Grade	Performance Level (%)				
	1	2	3	4	CCR
5	25	30	30	14	44
8	37	31	22	10	32
11	39	27	23	11	33

Note. CCR = college and career ready (combination of performance levels 3 and 4). Column percentages may not total 100% because of rounding.

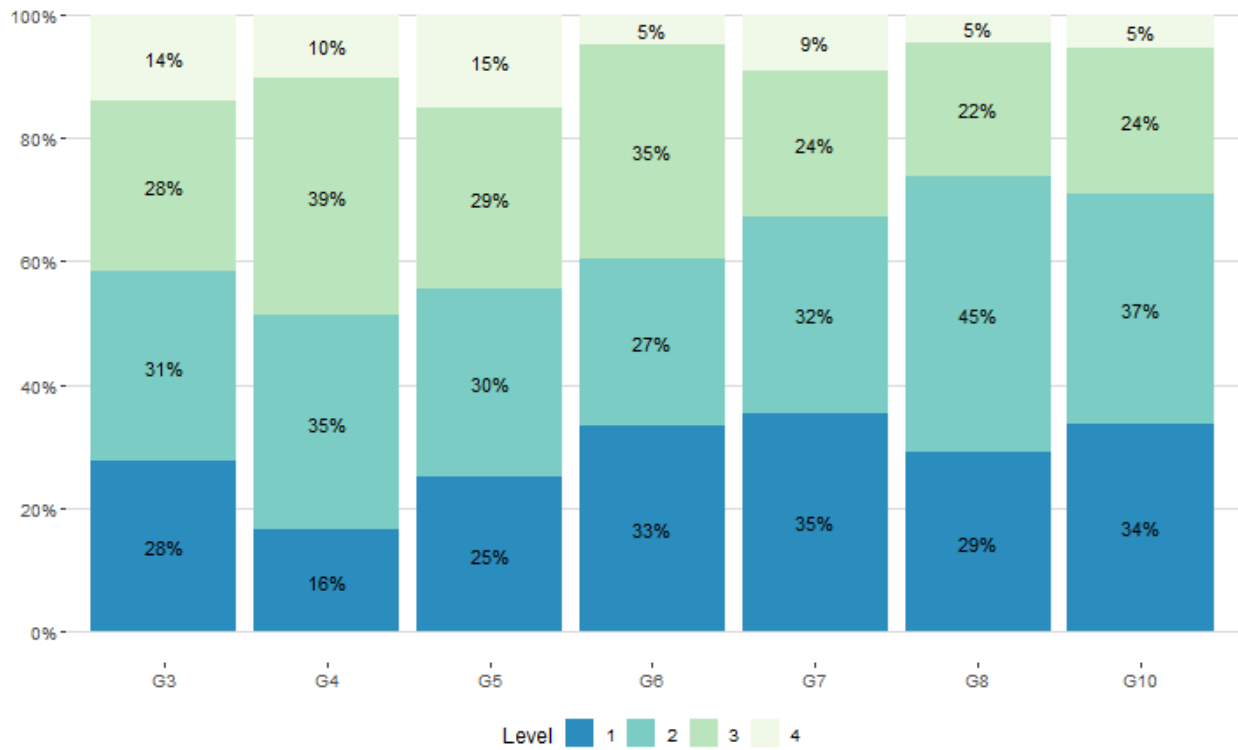


Figure III-1. Performance-level results for ELA.

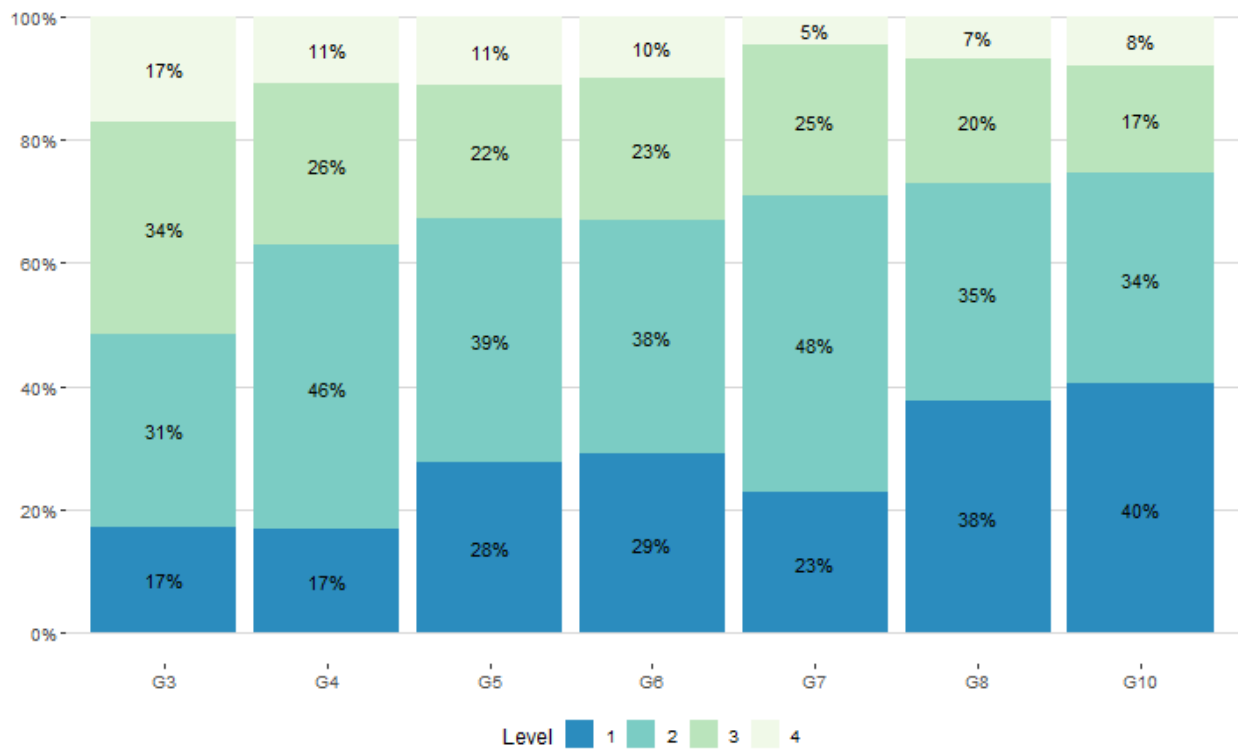


Figure III-2. Performance-level results for mathematics.

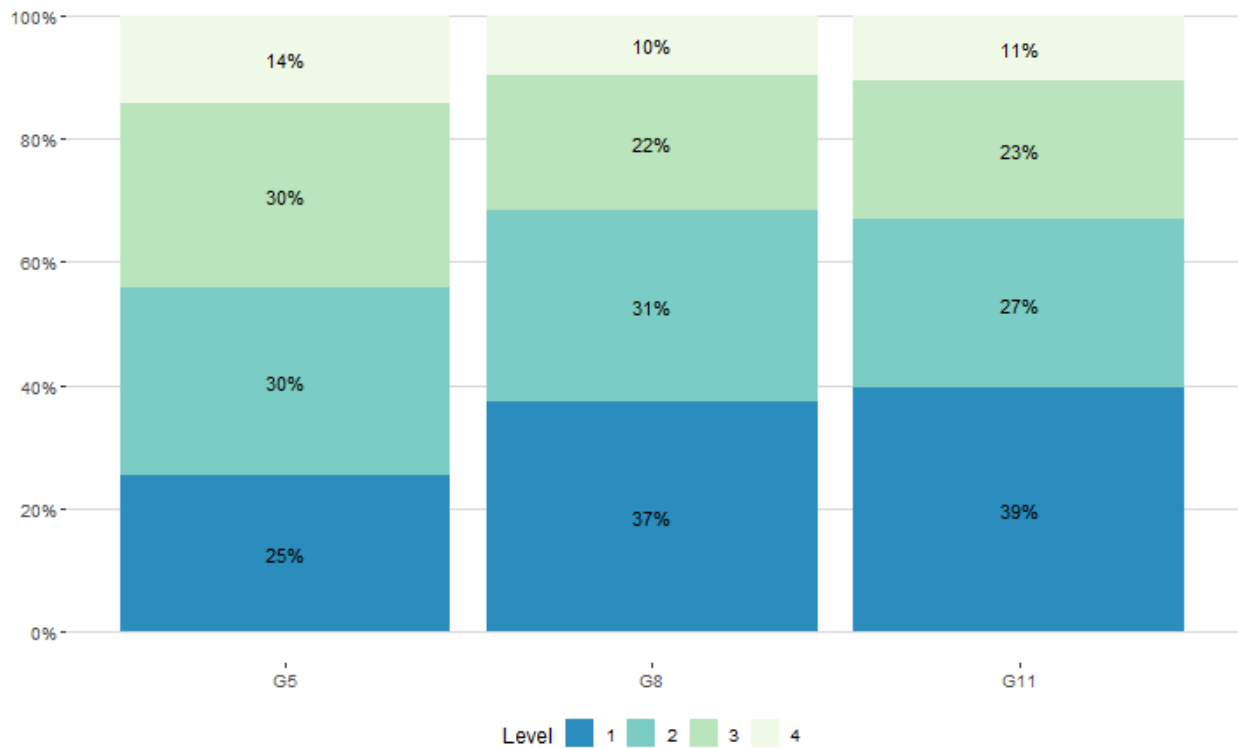


Figure III-3. Performance-level results for science.

III-2. Test Result Trends

The scale-score and performance-level trends across years for the three subjects are presented in Tables III-6 through III-8 and in Figures III-4 through III-9. Compared to 2018 results, 2019 ELA level-3 and level-4 percentages increased in grade 3 and grade 6, remained the same in grade 4 and grade 10, and declined in other grades. Mathematics level 3 and level 4 percentages increased in grade 8, remained the same in grade 6 and grade 7 and declined in all other grades; science level 3 and level 4 percentages remained the same in grade 8 but decreased in grade 5 and grade 11.

Table III-6. Longitudinal Scale-Score Trend for ELA

Grade	2017			2018			2019		
	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>
3	295.4	28.8	38,340	294.0	29.1	37,579	294.8	29.1	37,098
4	300.5	28.0	38,424	299.3	27.8	38,440	299.1	28.0	37,698
5	297.0	29.9	37,526	295.8	29.7	38,374	295.9	29.4	38,372
6	291.1	28.9	36,858	290.1	29.2	37,447	290.4	28.7	38,281
7	289.7	30.7	36,863	289.0	31.2	36,754	288.3	31.1	37,424
8	284.0	28.4	36,695	283.0	28.5	36,832	282.3	28.5	36,779
10	284.8	29.8	35,673	284.0	29.8	35,651	283.4	29.8	36,318

Table III-7. *Longitudinal Scale-Score Trend for Mathematics*

Grade	2017			2018			2019		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
3	303.2	27.5	38,438	302.6	28.0	37,641	303.2	27.9	37,184
4	293.9	28.2	38,514	292.7	28.3	38,493	292.9	28.6	37,771
5	291.1	27.6	37,608	290.5	27.5	38,413	290.9	27.1	38,413
6	291.3	26.8	36,923	290.6	27.0	37,487	291.6	27.4	38,329
7	288.4	27.7	36,910	287.5	27.6	36,784	288.2	28.1	37,456
8	284.7	29.1	36,758	283.9	29.4	36,870	286.2	28.8	36,785
10	285.6	28.5	35,653	285.2	28.6	35,658	286.1	27.9	36,287

Table III-8. *Longitudinal Scale-Score Trend for Science*

Grade	2017			2018			2019		
	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>
5	298.5	30.2	37,619	299.0	30.6	38,458	298.7	30.8	38,442
8	288.4	29.6	36,774	287.5	29.8	36,934	286.9	30.3	36,863
11	291.7	29.1	34,161	291.2	29.4	34,314	289.3	29.8	34,081

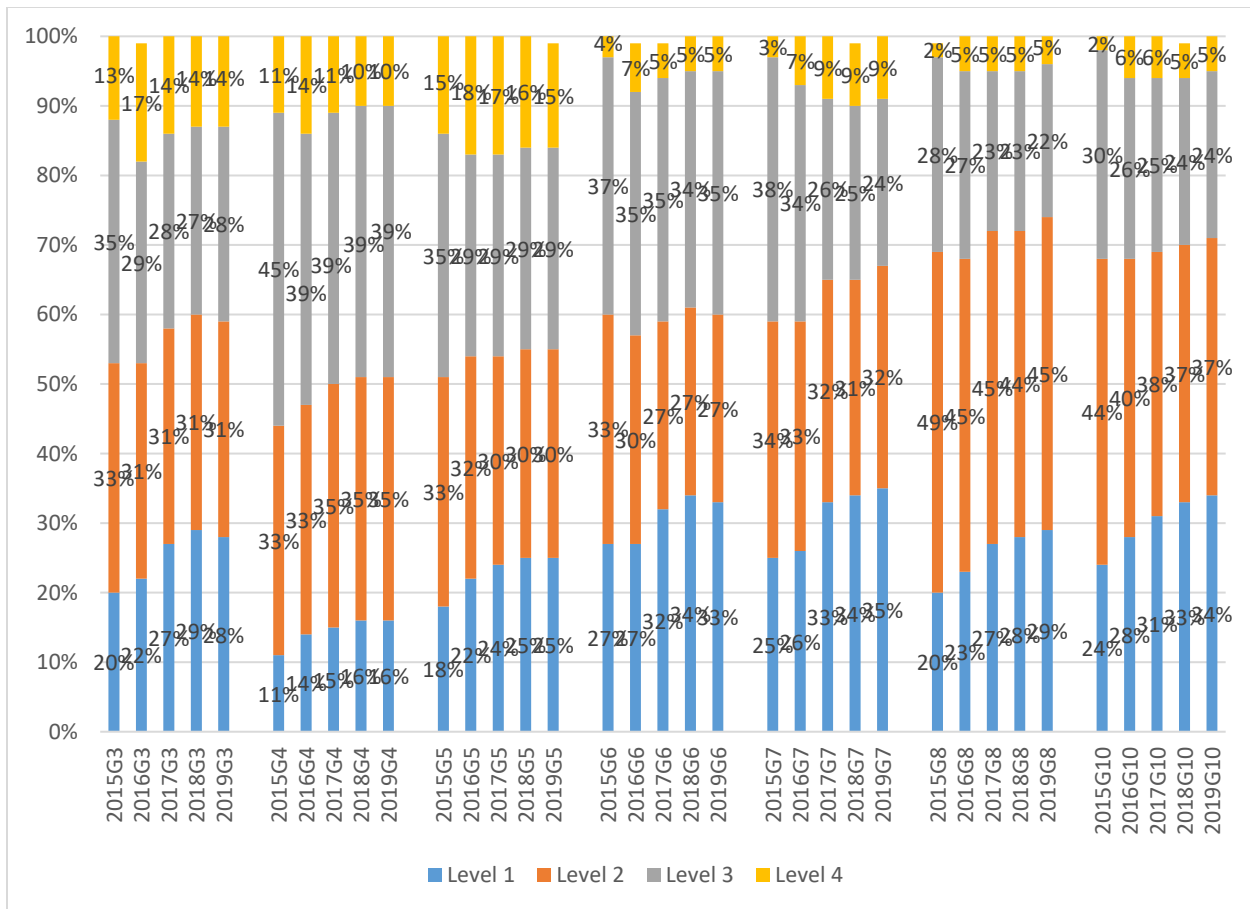


Figure III-4. Performance-level trend for ELA. Labels of column percentages may not total 100% because of rounding. G = grade.

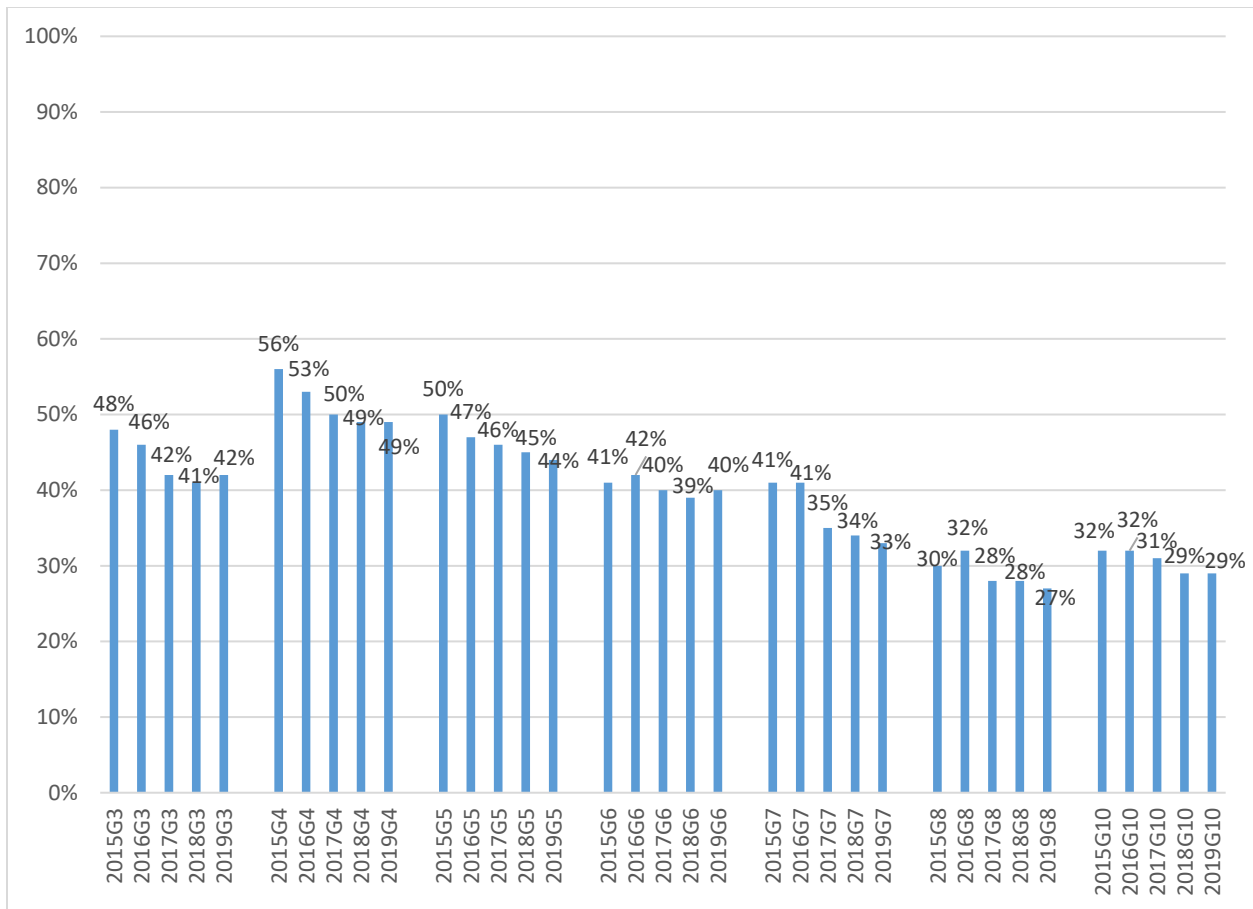


Figure III-5. College- and career-readiness trend for ELA. Students who score at or above the level-3 cut score are considered to be on target for postsecondary success. G = grade.

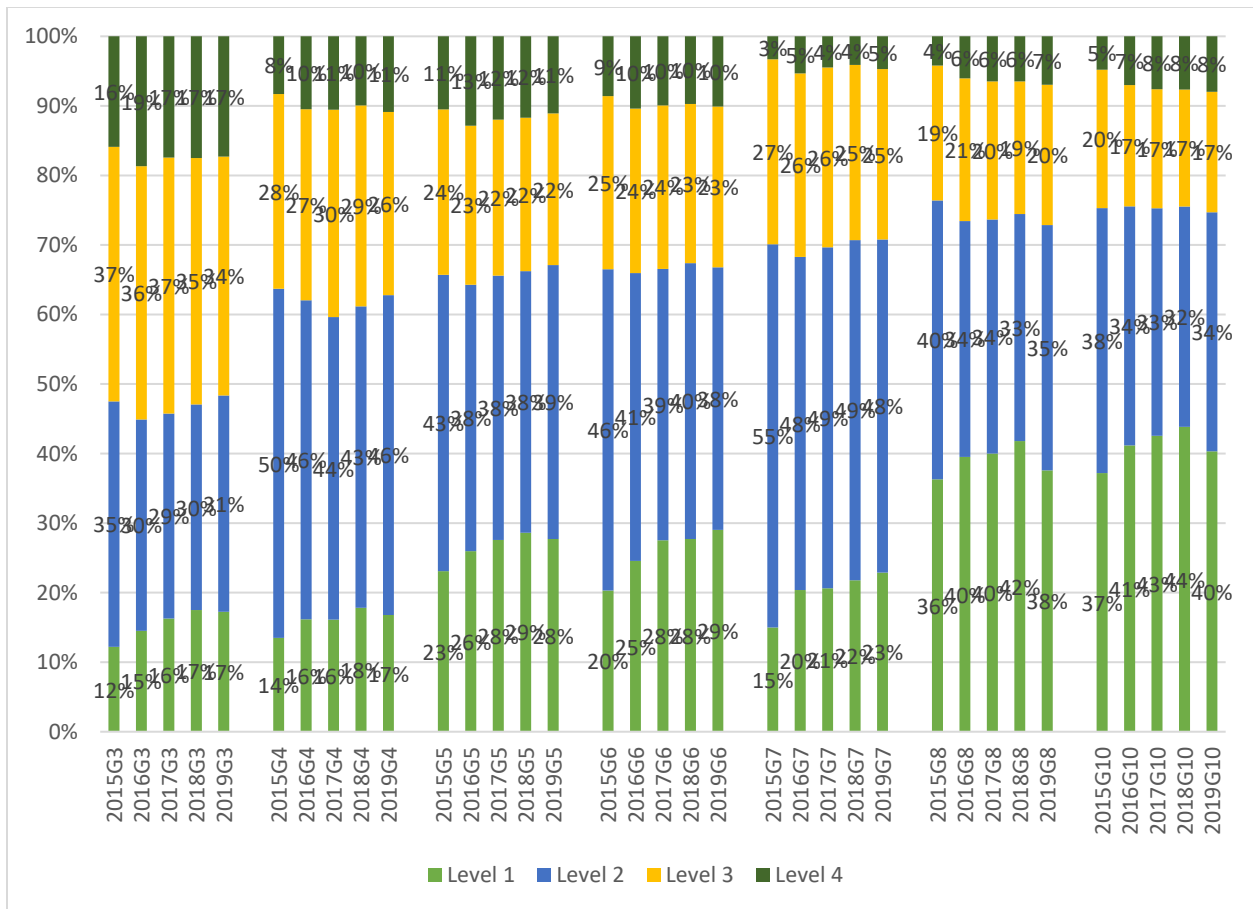


Figure III-6. Performance-level trend for mathematics. Labels of column percentages may not total 100% because of rounding. G = grade.

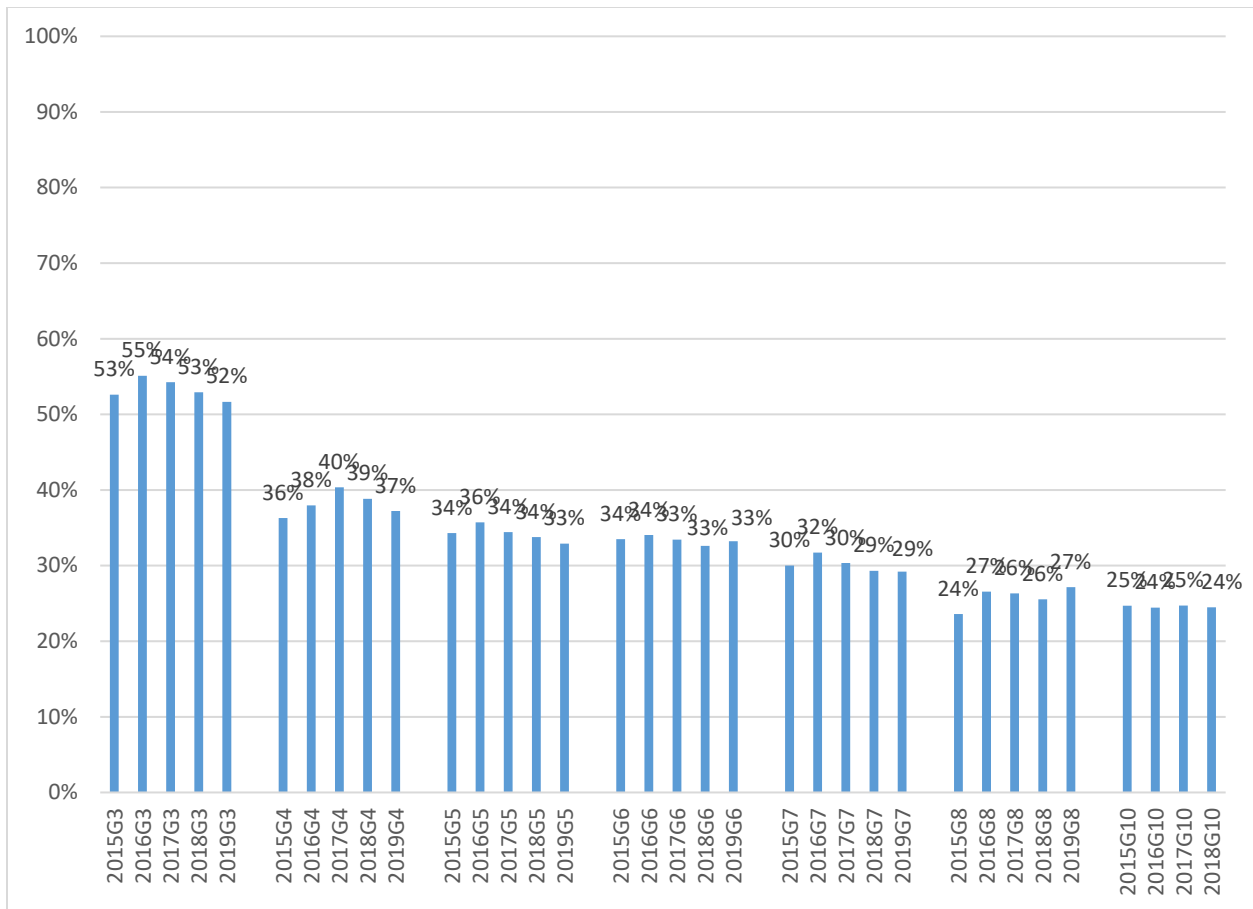


Figure III-7. College- and career-readiness trend for mathematics. Students who score at or above the level-3 cut score are considered to be on target for postsecondary success. G = grade.

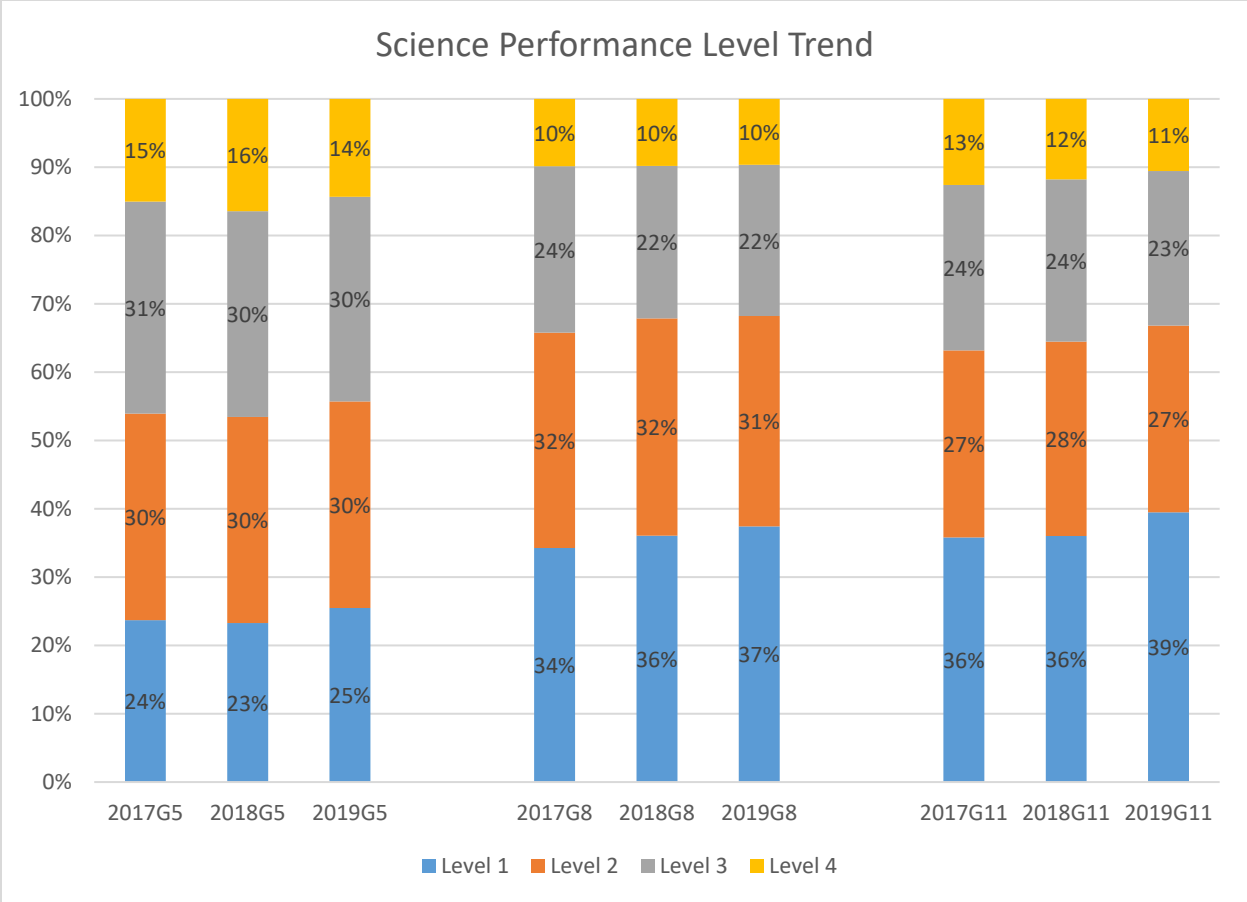


Figure III-8. Performance-level trend for science.

Labels of column percentage may not total to 100% due to rounding. G = grade.

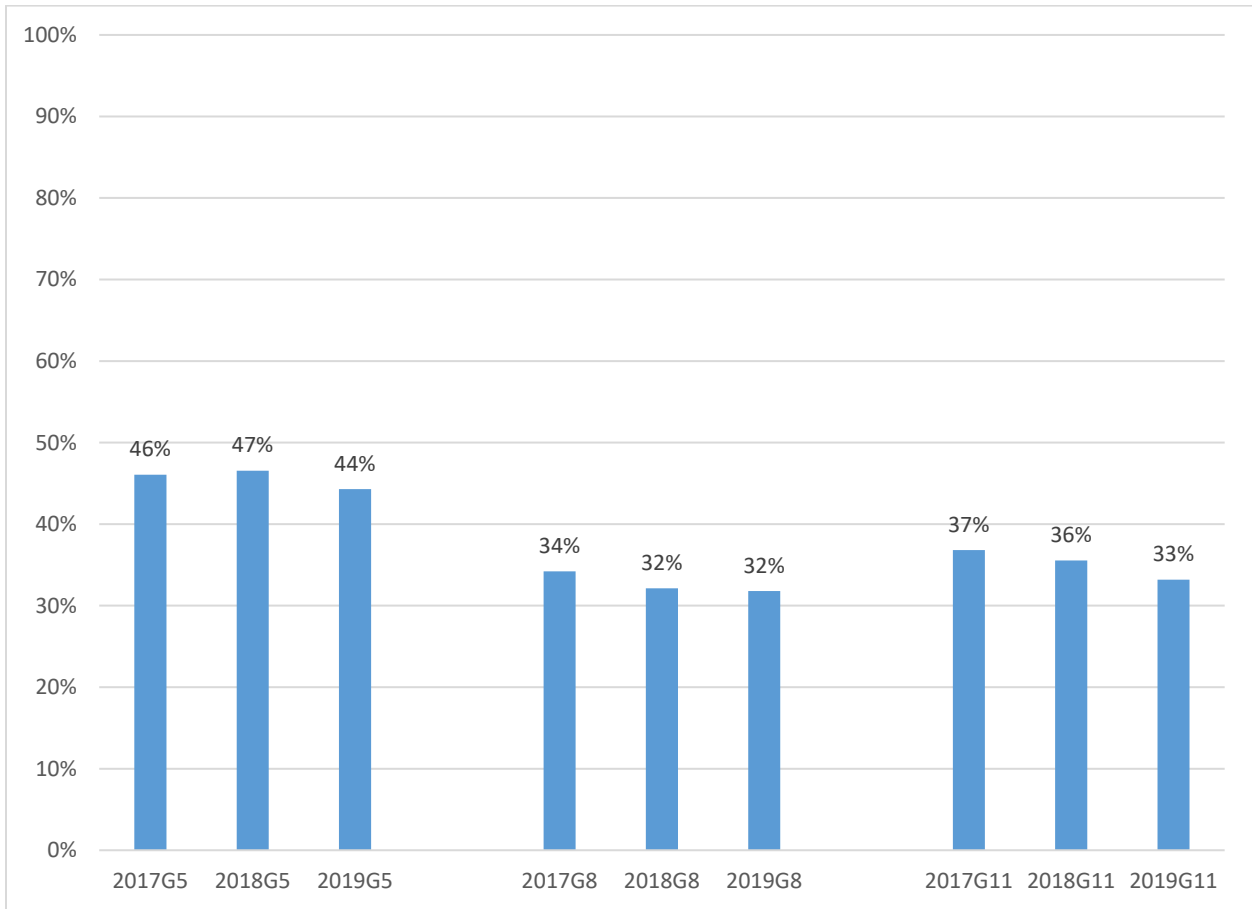


Figure III-9. College- and career-readiness trend for science. Students who score at or above the level-3 cut score are considered to be on target for postsecondary success. G = grade.

IV. Reliability

Marginal reliability was used for all analyses. As shown in Table IV-1, marginal reliabilities of ELA and mathematics are above .90; science has relatively lower reliabilities because there are fewer test items compared to ELA and mathematics, but values are still greater than or equal to .80.

Table IV-1. *Test Reliability by Subject and Grade*

Grade	ELA	Mathematics	Science
3	.93	.92	
4	.90	.94	
5	.91	.93	.80
6	.91	.93	
7	.91	.92	
8	.90	.92	.83
High school	.92	.92	.85

Classification consistency and accuracy indicate how accurately students are classified into performance categories. *Classification consistency* refers to the agreement between two parallel forms, and *classification accuracy* refers to the agreement between true scores and observed scores (Livingston & Lewis, 1995). Table IV-2 presents results for overall consistency across all four performance levels as well as for the dichotomies created by the three cut scores. Science has relatively lower consistency and accuracy because it has fewer test items compared to ELA and mathematics.

Table IV-2. *Classification Consistency and Accuracy*

Subject and grade	Cut-score category							
	Overall		1 vs. 2, 3, 4		1, 2 vs. 3, 4		1, 2, 3 vs. 4	
	C	A	C	A	C	A	C	A
ELA								
3	.62	.80	.74	.93	.78	.92	.75	.95
4	.57	.79	.62	.93	.73	.9	.69	.95
5	.57	.77	.70	.92	.75	.91	.72	.94
6	.59	.79	.73	.92	.73	.91	.62	.97
7	.59	.79	.72	.91	.75	.92	.70	.96
8	.59	.81	.70	.91	.71	.92	.61	.98
10	.62	.81	.73	.92	.75	.93	.66	.97
Mathematics								
3	.58	.78	.56	.92	.76	.92	.77	.95
4	.64	.82	.61	.92	.81	.94	.78	.97
5	.61	.80	.65	.90	.80	.94	.79	.97
6	.58	.79	.58	.89	.81	.94	.80	.97
7	.58	.81	.53	.89	.79	.94	.73	.98
8	.61	.81	.70	.90	.79	.94	.75	.98
10	.57	.79	.63	.88	.81	.95	.80	.98
Science								
5	.37	.65	.48	.87	.62	.86	.59	.93
8	.44	.70	.60	.87	.65	.89	.59	.95
11	.46	.72	.60	.86	.70	.91	.67	.96

Note. C = consistency; A = accuracy.

Tables IV-3 through IV-19 present subgroup marginal reliabilities for each subject by grade. The race analysis has a smaller sample size than the other subgroups because students whose demographic information about race was not provided were excluded from the analysis. For ELA and mathematics, the subgroup marginal reliabilities for each group are close to or above .90 across grades, ranging from .88 to .93 for ELA and from .89 to .94 for mathematics. Science has relatively lower subgroup reliabilities because it has fewer test items compared to ELA and mathematics. Science subgroup marginal reliabilities ranged from .76 to .88 across grades.

Table IV-3. *Grade 3 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	36,756					
Black		2,644	7.2	.93	280.2	26.2
American Indian		756	2.1	.93	284.7	25.0
Asian		1,090	3.0	.92	302.3	30.7
NHPI		107	0.3	.93	288.3	28.1
White		30,016	81.7	.93	296.4	28.9
Hispanic	37,067					
Yes		7,420	20.0	.93	283.5	25.8
No		29,647	80.0	.92	297.7	29.1
SWD	37,067					
Yes		5,287	14.3	.93	274.9	26.1
No		31,780	85.7	.93	298.1	28.2
EL	37,067					
Yes		4,764	12.9	.93	279.2	24.6
No		32,303	87.1	.93	297.1	28.9
Gender	37,067					
Boys		18,867	50.9	.93	292.4	29.1
Girls		18,200	49.1	.93	297.3	28.8

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-4. *Grade 4 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	37,382					
Black		2,809	7.5	.91	284.5	25.5
American Indian		848	2.3	.91	287.9	23.9
Asian		1,061	2.8	.89	309.5	31.0
NHPI		112	0.3	.91	288.7	28.8
White		30,378	81.3	.90	300.8	27.6
Hispanic	37,667					
Yes		7,757	20.6	.91	288.7	25.1
No		29,910	79.4	.90	301.9	28.0
SWD	37,667					
Yes		5,236	13.9	.91	278.3	25.9
No		32,431	86.1	.90	302.6	26.8
EL	37,667					
Yes		4,916	13.1	.92	284.1	23.7
No		32,751	86.9	.90	301.5	27.9
Gender	37,667					
Boys		19,181	50.9	.91	296.8	28.0
Girls		18,486	49.1	.90	301.6	27.7

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-5. *Grade 5 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	38,026					
Black		2,851	7.5	.92	281.1	27.5
American Indian		963	2.5	.92	284.6	25.9
Asian		1,072	2.8	.89	307.5	32.0
NHPI		91	0.2	.91	289.0	28.6
White		30,855	81.1	.91	297.5	29.0
Hispanic	38,344					
Yes		7,922	20.7	.92	284.8	25.8
No		30,422	79.3	.91	298.8	29.5
SWD	38,344					
Yes		5,014	13.1	.92	271.8	25.4
No		33,330	86.9	.91	299.5	28.1
EL	38,344					
Yes		4,743	12.4	.92	279.0	23.8
No		33,601	87.6	.91	298.3	29.2
Gender	38,344					
Boys		19,498	50.9	.91	292.8	28.9
Girls		18,846	49.1	.91	299.2	29.4

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-6. *Grade 6 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	37,941					
Black		2,697	7.1	.92	271.9	26.1
American Indian		954	2.5	.92	277.0	24.9
Asian		1,128	3.0	.90	302.1	30.2
NHPI		86	0.2	.92	281.3	27.2
White		30,820	81.2	.91	292.4	28.2
Hispanic	38,233					
Yes		7,692	20.1	.92	278.7	26.1
No		30,541	79.9	.91	293.5	28.6
SWD	38,233					
Yes		4,709	12.3	.92	265.3	25.0
No		33,524	87.7	.91	294.0	27.4
EL	38,233					
Yes		4,536	11.9	.92	272.9	24.2
No		33,697	88.1	.91	292.9	28.4
Gender	38,233					
Boys		19,552	51.1	.91	287.9	28.9
Girls		18,681	48.9	.91	293.2	28.2

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-7. *Grade 7 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	37,135					
Black		2,689	7.2	.92	270.7	26.0
American Indian		1,087	2.9	.92	275.6	27.3
Asian		1,098	3.0	.89	298.5	32.5
NHPI		93	0.3	.92	280.5	27.3
White		29,965	80.7	.90	290.4	30.9
Hispanic	37,363					
Yes		7,530	20.2	.92	277.0	27.7
No		29,833	79.8	.90	291.3	31.1
SWD	37,363					
Yes		4,534	12.1	.93	261.4	24.4
No		32,829	87.9	.90	292.1	30.0
EL	37,363					
Yes		4,274	11.4	.93	269.6	24.5
No		33,089	88.6	.90	290.8	30.9
Gender	37,363					
Boys		18,948	50.7	.91	284.7	31.2
Girls		18,415	49.3	.90	292.2	30.4

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-8. *Grade 8 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	36,482					
Black		2,552	7.0	.91	267.5	24.8
American Indian		1,276	3.5	.91	270.5	25.7
Asian		1,046	2.9	.88	294.5	31.4
NHPI		88	0.2	.91	273.0	26.4
White		29,374	80.5	.90	284.0	28.1
Hispanic	36,707					
Yes		7,297	19.9	.91	272.1	25.5
No		29,410	80.1	.90	284.9	28.5
SWD	36,707					
Yes		4,280	11.7	.92	255.7	22.6
No		32,427	88.3	.90	285.9	27.2
EL	36,707					
Yes		4,130	11.3	.91	265.9	23.2
No		32,577	88.7	.90	284.5	28.3
Gender	36,707					
Boys		18,707	51.0	.91	278.3	28.2
Girls		18,000	49.0	.90	286.6	28.0

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-9. *Grade 10 Subgroup Reliability and Performance for ELA*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	35,901					
Black		2,494	6.9	.93	267.4	25.4
American Indian		1,353	3.8	.93	270.4	24.5
Asian		1,197	3.3	.91	291.2	31.8
NHPI		97	0.3	.92	278.4	28.6
White		28,853	80.4	.91	285.6	29.5
Hispanic	36,109					
Yes		6,859	19.0	.93	272.2	26.7
No		29,250	81.0	.91	286.4	29.6
SWD	36,109					
Yes		3,852	10.7	.93	255.8	22.3
No		32,257	89.3	.91	287.0	28.6
EL	36,109					
Yes		3,822	10.6	.93	263.5	23.4
No		32,287	89.4	.91	286.1	29.3
Gender	36,109					
Boys		18,445	51.1	.92	279.0	29.9
Girls		17,664	48.9	.91	288.5	28.5

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-10. *Grade 3 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	36,823					
Black		2,646	7.2	.94	288.3	24.3
American Indian		766	2.1	.93	295.0	24.2
Asian		1,098	3.0	.90	315.6	31.0
NHPI		108	0.3	.93	299.5	26.1
White		30,070	81.7	.92	304.7	27.7
Hispanic	37,135					
Yes		7,495	20.2	.94	292.9	24.2
No		29,640	79.8	.92	305.9	28.1
SWD	37,135					
Yes		5,278	14.2	.93	285.2	25.9
No		31,857	85.8	.92	306.2	27.1
EL	37,135					
Yes		4,883	13.1	.93	290.7	24.4
No		32,252	86.9	.92	305.1	27.9
Gender	37,135					
Boys		18,895	50.9	.92	304.5	28.9
Girls		18,240	49.1	.93	302.0	26.7

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-11. *Grade 4 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (N)	Group (n)	%	Reliability	Scale score	
					M	SD
Race	37,440					
Black		2,808	7.5	.94	275.9	22.6
American Indian		858	2.3	.94	282.0	23.9
Asian		1,076	2.9	.92	308.6	34.7
NHPI		114	0.3	.94	284.8	26.3
White		30,410	81.2	.94	294.7	28.3
Hispanic	37,725					
Yes		7,822	20.7	.94	281.9	24.1
No		29,903	79.3	.94	295.8	29.0
SWD	37,725					
Yes		5,226	13.9	.94	274.2	25.5
No		32,499	86.1	.94	295.9	27.9
EL	37,725					
Yes		5,027	13.3	.94	279.0	23.7
No		32,698	86.7	.94	295.1	28.7
Gender	37,725					
Boys		19,193	50.9	.93	295.0	30.0
Girls		18,532	49.1	.94	290.7	26.9

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-12. *Grade 5 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	38,047					
Black		2,843	7.5	.94	276.9	22.3
American Indian		971	2.6	.94	282.1	23.3
Asian		1,080	2.8	.91	307.4	33.1
NHPI		96	0.3	.93	282.9	25.3
White		30,875	81.1	.93	292.3	26.9
Hispanic	38,366					
Yes		7,977	20.8	.94	281.6	22.6
No		30,389	79.2	.93	293.5	27.6
SWD	38,366					
Yes		4,995	13.0	.93	271.9	22.0
No		33,371	87.0	.93	293.9	26.6
EL	38,366					
Yes		4,834	12.6	.94	278.7	21.8
No		33,532	87.4	.93	292.8	27.3
Gender	38,366					
Boys		19,497	50.8	.93	292.7	28.5
Girls		18,869	49.2	.93	289.2	25.4

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-13. *Grade 6 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	37,987					
Black		2,710	7.1	.94	275.8	20.6
American Indian		957	2.5	.94	280.3	22.1
Asian		1,134	3.0	.89	308.9	33.6
NHPI		86	0.2	.93	284.5	28.6
White		30,855	81.2	.93	293.1	27.2
Hispanic	38,277					
Yes		7,761	20.3	.94	281.6	22.6
No		30,516	79.7	.93	294.2	27.9
SWD	38,277					
Yes		4,708	12.3	.93	271.4	20.5
No		33,569	87.7	.93	294.5	27.1
EL	38,277					
Yes		4,642	12.1	.94	278.4	21.3
No		33,635	87.9	.93	293.5	27.7
Gender	38,277					
Boys		19,575	51.1	.92	292.4	28.5
Girls		18,702	48.9	.93	290.9	26.2

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-14. *Grade 7 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	37,159					
Black		2,684	7.2	.92	271.5	21.3
American Indian		1,098	3.0	.92	277.1	21.8
Asian		1,102	3.0	.90	304.6	34.4
NHPI		92	0.2	.92	281.4	24.7
White		29,985	80.7	.92	290.0	27.8
Hispanic	37,387					
Yes		7,587	20.3	.92	277.9	23.0
No		29,800	79.7	.92	291.0	28.5
SWD	37,387					
Yes		4,522	12.1	.92	265.6	20.3
No		32,865	87.9	.92	291.4	27.5
EL	37,387					
Yes		4,364	11.7	.92	273.6	21.1
No		33,023	88.3	.92	290.2	28.2
Gender	37,387					
Boys		18,963	50.7	.92	289.1	29.0
Girls		18,424	49.3	.92	287.5	26.8

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-15. *Grade 8 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	36,492					
Black		2,554	7.0	.92	270.8	22.0
American Indian		1,279	3.5	.92	275.0	23.4
Asian		1,056	2.9	.90	306.4	36.3
NHPI		88	0.2	.92	276.3	25.2
White		29,379	80.5	.92	287.7	28.5
Hispanic	36,718					
Yes		7,342	20.0	.92	276.1	23.7
No		29,376	80.0	.92	288.8	29.4
SWD	36,718					
Yes		4,265	11.6	.92	263.2	20.0
No		32,453	88.4	.92	289.3	28.4
EL	36,718					
Yes		4,216	11.5	.92	272.5	22.0
No		32,502	88.5	.92	288.0	29.1
Gender	36,718					
Boys		18,703	50.9	.92	286.2	30.2
Girls		18,015	49.1	.92	286.3	27.2

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-16. *Grade 10 Subgroup Reliability and Performance for Mathematics*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	35,929					
Black		2,503	7.0	.93	271.4	18.7
American Indian		1,355	3.8	.93	273.9	19.0
Asian		1,204	3.4	.90	306.1	35.9
NHPI		95	0.3	.93	284.2	25.0
White		28,868	80.3	.92	287.7	27.7
Hispanic	36,140					
Yes		6,875	19.0	.93	275.9	22.4
No		29,265	81.0	.92	288.8	28.4
SWD	36,140					
Yes		3,851	10.7	.92	265.5	16.4
No		32,289	89.3	.92	288.8	27.8
EL	36,140					
Yes		3,869	10.7	.93	271.8	19.5
No		32,271	89.3	.92	288.1	28.1
Gender	36,140					
Boys		18,453	51.1	.92	286.4	29.0
Girls		17,687	48.9	.93	286.2	26.5

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-17. *Grade 5 Subgroup Reliability and Performance for Science*

Subgroup	Grade (N)	Group (n)	%	Reliability	Scale score	
					M	SD
Race	37,999					
Black		2,857	7.5	.84	282.4	27.2
American Indian		974	2.6	.83	287.1	28.2
Asian		1,079	2.8	.76	309.1	32.9
NHPI		95	0.3	.84	290.2	28.1
White		30,804	81.1	.80	300.8	30.5
Hispanic	38,318					
Yes		7,974	20.8	.84	287.2	27.5
No		30,344	79.2	.80	302.0	30.8
SWD	38,318					
Yes		4,997	13.0	.84	279.0	29.7
No		33,321	87.0	.80	301.9	29.7
EL	38,318					
Yes		4,832	12.6	.85	282.5	26.1
No		33,486	87.4	.80	301.3	30.6
Gender	38,318					
Boys		19,493	50.9	.80	300.1	32.1
Girls		18,825	49.1	.81	297.6	29.2

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-18. *Grade 8 Subgroup Reliability and Performance for Science*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	36,579					
Black		2,570	7.0	.84	269.3	24.4
American Indian		1,283	3.5	.84	273.7	26.0
Asian		1,054	2.9	.81	297.0	32.6
NHPI		89	0.2	.84	275.5	26.7
White		29,436	80.5	.83	289.1	30.1
Hispanic	36,804					
Yes		7,373	20.0	.84	274.9	26.1
No		29,431	80.0	.83	290.0	30.5
SWD	36,804					
Yes		4,275	11.6	.83	265.2	25.5
No		32,529	88.4	.83	289.8	29.6
EL	36,804					
Yes		4,229	11.5	.84	269.2	23.3
No		32,575	88.5	.83	289.3	30.3
Gender	36,804					
Boys		18,743	50.9	.82	288.9	32.3
Girls		18,061	49.1	.84	285.0	27.8

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-19. *Grade 11 Subgroup Reliability and Performance for Science*

Subgroup	Grade (<i>N</i>)	Group (<i>n</i>)	%	Reliability	Scale score	
					<i>M</i>	<i>SD</i>
Race	33,700					
Black		2,313	6.9	.88	270.1	22.1
American Indian		1,310	3.9	.88	276.3	23.8
Asian		1,146	3.4	.83	293.4	33.0
NHPI		77	0.2	.87	280.4	26.9
White		27,107	80.4	.85	291.8	29.8
Hispanic	34,013					
Yes		6,163	18.1	.87	277.4	25.0
No		27,850	81.9	.85	292.0	30.1
SWD	34,013					
Yes		3,441	10.1	.88	267.8	23.0
No		30,572	89.9	.85	291.8	29.5
EL	34,013					
Yes		3,294	9.7	.88	269.5	20.3
No		30,719	90.3	.85	291.5	29.8
Gender	34,013					
Boys		17,185	50.5	.84	290.7	31.9
Girls		16,828	49.5	.86	288.1	27.4

Note. NHPI = Native Hawaiian and Pacific Islander; SWD = students with disabilities; EL = English learners.

Table IV-20 shows path reliabilities for ELA by grade. Path reliability is equivalent to the reliability of different test forms. All path reliabilities are about or above .90.

Table IV-20. *Path Reliability for ELA by Grade*

Grade	Path	Stage 1	Stage 2	<i>N</i>	%	Reliability
3	1	Medium	Easy	19,943	53.8	.93
3	2	Medium	Hard	17,124	46.2	.92
4	1	Medium	Easy	7,817	20.8	.92
4	2	Medium	Hard	29,850	79.3	.90
5	1	Medium	Easy	19,291	50.3	.92
5	2	Medium	Hard	19,053	49.7	.89
6	1	Medium	Easy	7,750	20.3	.92
6	2	Medium	Hard	30,483	79.7	.91
7	1	Medium	Easy	20,717	55.5	.93
7	2	Medium	Hard	16,646	44.6	.88
8	1	Medium	Easy	19,332	52.7	.92
8	2	Medium	Hard	17,375	47.3	.89
10	1	Medium	Easy	17,680	49.0	.94
10	2	Medium	Hard	18,429	51.0	.90

Table IV-21 presents marginal reliability, classification consistency, and accuracy for subscores by subject. Science has the highest subscore reliability.

Table IV-21. *Subscore Reliability, Classification Consistency, and Accuracy by Subject*

Subject	Subscores							
	(<i>n</i>)	<i>M</i>	<i>SD</i>	Min.	P ₂₅	P ₅₀	P ₇₅	Max.
Reliability								
ELA	63	.62	.06	.50	.58	.62	.67	.75
Mathematics	44	.61	.09	.43	.54	.60	.66	.76
Science	9	.69	.04	.61	.68	.70	.72	.74
Consistency								
ELA	63	.35	.06	.22	.3	.34	.38	.47
Mathematics	44	.34	.10	.12	.28	.34	.40	.52
Science	9	.31	.03	.26	.31	.32	.32	.37
Accuracy								
ELA	63	.71	.06	.61	.68	.70	.76	.82
Mathematics	44	.72	.09	.50	.65	.72	.79	.87
Science	9	.69	.04	.61	.68	.70	.72	.74

Note. P₂₅ = 25th percentile; P₅₀ = 50th percentile; P₇₅ = 75th percentile.

V. Validity Evidence

V-1. Evidence Based on Internal Structure

Differential item functioning (DIF) analyses evaluate items for potential bias. Logistic regression was used to detect items with DIF. The Jodoin and Gierl (2001) DIF classification criteria were used to indicate the degree of DIF (i.e., negligible, moderate, large). When the DIF test is significant, large DIF is identified by a Nagelkerke R^2 change greater than .070, and moderate DIF has a Nagelkerke R^2 change between .035 and .070. Numbers of items flagged for moderate or large DIF for each subject are listed by grade in Tables V-1 through V-3. Only two grade 5 ELA items were identified with moderate DIF favoring male students. No other items were identified with either moderate or large DIF, suggesting that item-development process and procedures effectively addressed potential bias and sensitivity issues during the development phase. When an item is flagged, test development teams review the item for potential sources of bias against subgroups of the population.

Table V-1. *DIF for ELA Across Subgroups by Grade*

Grade	Items (<i>n</i>)	Gender		Race	
		Moderate	Large	Moderate	Large
3	86	0	0	0	0
4	82	0	0	0	0
5	81	2	0	0	0
6	74	0	0	0	0
7	74	0	0	0	0
8	80	0	0	0	0
10	68	0	0	0	0

Table V-2. *DIF for Mathematics Across Subgroups By Grade*

Grade	Items (<i>n</i>)	Gender		Race	
		Moderate	Large	Moderate	Large
3	55	0	0	0	0
4	55	0	0	0	0
5	55	0	0	0	0
6	55	0	0	0	0
7	55	0	0	0	0
8	55	0	0	0	0
10	55	0	0	0	0

Table V-3. *DIF for Science Across Subgroups by Grade*

Grade	Items (<i>n</i>)	Gender		Race	
		Moderate	Large	Moderate	Large
5	33	0	0	0	0
8	41	0	0	0	0
11	38	0	0	0	0

V-2. Evidence Based on Relationships to Other Variables

The correlations presented in Table V-4 are between subjects of the same grade, and the values range from .68 to .76. The correlations between KAP and ACT scores are presented in Table V-5, and the values range from .67 to .86. The highest correlation is between grade 10 KAP mathematics scores and ACT mathematics scores: $r = .86$.

Table V-4. *Correlations Between ELA, Mathematics, and Science Scores by Grade*

Grade	ELA vs. mathematics	ELA vs. science	Mathematics vs. science
3	.76		
4	.75		
5	.73	.73	.68
6	.75		
7	.73		
8	.73	.73	.70
10	.69		

Table V-5. *Correlations Between KAP Scores (2017) and ACT (2017–2018) Scores*

KAP		ACT				
Subject	<i>n</i>	Composite ^a	English	Reading	Mathematics	Science
Grade 10 ELA	2,015	.79*	.75*	.75*		
Grade 10 mathematics	2,015	.80*			.86*	
Grade 11 science	4,283	.71*				.67*

^aThe ACT composite score is the average of scores on the four subjects (English, reading, mathematics, and science).

* $p < .001$.

References

- Jodoin, M. G., & Gierl, M. J. (2001). Evaluating Type I error and power rate using an effect size measure with the logistic regression procedure for DIF detection. *Applied Measurement in Education, 14*, 329–349. https://doi.org/10.1207/S15324818AME1404_2
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32*, 179–197. <https://doi.org/10.1111/j.1745-3984.1995.tb00462.x>