

Oklahoma
School Testing Program / College- and Career-
Readiness Assessment
Grades 3–8, 11

2024–25 Technical Report

Part II—Appendix M through Appendix V

Prepared by Cognia and
the Oklahoma Department of Education

APPENDIX M
OKLAHOMA COLLEGE AND CAREER READINESS
ASSESSMENT
2024–2025 EQUATING REPORT



Oklahoma College and Career Readiness Assessment

2024–2025: EQUATING REPORT

2024 - 2025 Oklahoma College and Career Readiness Assessment

Equating Report

The purpose of this document is to summarize the equating results obtained from Cognia for CCRA. Presented in this report are various program summary statistics and specific results related to the equating study.

1. Aggregate Results

1. Percentage of Students by Performance Level Categories
2. Theta Cuts and Scaling Constants
3. Calibration Report
4. Equating Item Summary Statistics

2. Grade Subject Results

1. A/A, B/B, Delta, and Cumulative Scale Score Distribution Plots
2. Cumulative Scale Score Distribution Tables
3. Tabled Delta Analysis Results
4. Tabled B/B Analysis Results
5. Final Item Parameters
6. Fit Plots of Watchlist Items

Section 1.1

Percentage of Students by Performance Level Categories

Table 1.1.1
Percentage of Students by Performance Level Categories
Science

Grade	Year	Count	BB	B	P	A	P+A	Delta	Ave. SS	SD SS
11	2025	48353	58	21	16	6	22	-2.6	271.9	32.2
	2024	47712	53	23	17	7	24	2.4	275.3	32.5
	2023	46302	57	21	16	6	22	-3.5	272.0	32.9
	2022	44157	54	21	18	8	25	1.5	272.4	37.9
	2021	42566	52	24	17	6	24	-0.2	271.5	38.4
	2019	43638	57	20	17	7	24		271.2	36.8

Table 1.1.2
Percentage of Students by Performance Level Categories
U.S. History

Grade	Year	Count	BB	B	P	A	P+A	Delta	Ave. SS	SD SS
11	2025	48309	43	14	34	9	43	-1.8	295.3	25.2
	2024	47745	40	15	36	9	45	1.2	296.6	24.3
	2023	46355	42	14	35	9	44	-2.7	295.7	24.7
	2022	44168	40	14	36	10	46		294.6	30.6

Section 1.2

Theta Cuts and Scaling Constants



Table 1.2.1
Theta and Scale Score Cuts

Subject	Grade	Type	Theta 1	Theta 2	Theta 3	Minimum	Scale Score 1	Scale Score 2	Scale Score 3	Maximum
Science	Grade 11	Scaling	0.17	0.80	1.53	200	278	300	327	399
U.S. History	Grade 11	Scaling	-0.26	0.14	1.30	200	290	300	330	399

Table 1.2.2
Scaling Slope and Intercept

Subject	Grade	Slope	Intercept
Science	Grade 11	35.8776457400000	271.2212872000000
U.S. History	Grade 11	25.9553119200000	296.3759796000000

Section 1.3

Calibration Report

Calibration Report—Executive Summary

PARSCALE 4.1 was used for all analyses. All command files were set up in a way that all general settings were identical to last year. For example, the calibration statement reads:

```
CAL GRADED,LOGISTIC,CYCLE=(150,1,1,1,1),TPRIOR,SPRIOR,GPRIOR;
```

Thus, a 3PLM was used for all MC items, and a Graded Response Model was specified for the polytomous items. The logistic version of the IRT models was used, and default priors were used for all parameter estimates. Each item occupied its own unique block in the command file, and for most items, initial guessing parameters were set to 0.22.

The resulting parameters demonstrated excellent model fit. In particular, the largest change in parameter values (from one iteration to the next) was monotonically decreasing and tended to flatten out towards the end of the calibration process. The number of Newton cycles to conversion for each grade/content for the initial calibrations are listed in the following table:

Table 1.3.1
Number of Cycles to Convergence

Subject	Grade	Initial Cycles
Science	Grade 11	86
U.S. History	Grade 11	60

For some items, the guessing parameter was poorly estimated. This is not at all unusual as difficulty in estimating the c-parameter has been well documented in psychometric literature. This often happens when item discrimination is low (e.g., less than 0.50). After carefully studying these items, we found that fixing the lower asymptote (for example to a value of 0.00) resulted in stable and reasonable estimates for both the a and b parameters (relative to CTT statistics). This technique also produced item parameters that resulted in excellent model fit (comparing theoretical ICCs to observed ICCs).

Three methods of evaluating the suitability of the equating items were used: the delta analysis, the b/b analysis, and the rescore analysis. As a result of all three analyses, very few items were removed from the equating analysis. Results such as this are very common, particularly given the number of grade/content combinations and the number and types of items in the program. Results from these analyses are included in Section II of this report.

Items flagged by the delta or b/b analyses, or any item that required intervention during the calibration process, were compiled and placed in our item watch list, which includes the final actions taken on these items. The final watch list is presented in the following table:

Table 1.3.2
Final Items Watch List

Subject	Grade	ItemID	Reason	Action
Science	11	592069	delta analysis	retained for equating
U.S. History	11	658060	c-parameter	set c = 0
U.S. History	11	658072	delta analysis	removed from equating

Stocking and Lord procedure was used to transform parameter estimates onto the operational scale. This procedure results in constants which were applied to the resulting IRT parameters for each grade/content. These transformation constants were found using the STUIRT program which can be found at the CASMA website: <http://www.education.uiowa.edu/casma/>. The Stocking & Lord transformation constants that were used in the equating process are listed in the following table:

Table 1.3.3
Stocking and Lord Constants

Subject	Grade	Slope	Intercept	Num Eq Items	Num Eq Items Rem
Science	11	0.98	0.01	59	0
U.S. History	11	1.03	-0.04	50	1

Section 1.4

Equating Item Summary Statistics

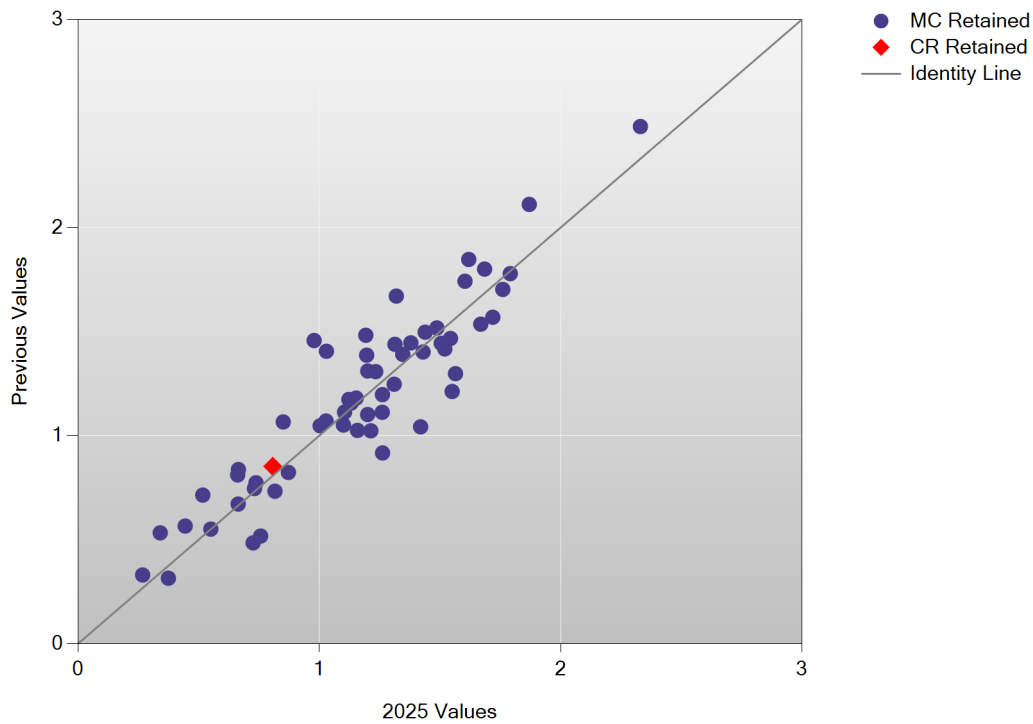
Table 1.4.1
Equating Item Summary Statistics

Subject	Grade	Year	P-Value Mean	P-Value Std Dev	Point Biserial Mean	Point Biserial Std Dev	<i>a</i> Mean	<i>a</i> Std Dev	<i>b</i> Mean	<i>b</i> Std Dev
Science	11	2025	0.44	0.11	0.38	0.11	1.17	0.43	0.87	0.71
		Previous	0.45	0.11	0.38	0.11	1.18	0.44	0.86	0.71
U.S. History	11	2025	0.53	0.12	0.45	0.09	1.12	0.34	0.39	0.65
		Previous	0.54	0.11	0.44	0.08	1.10	0.35	0.35	0.62

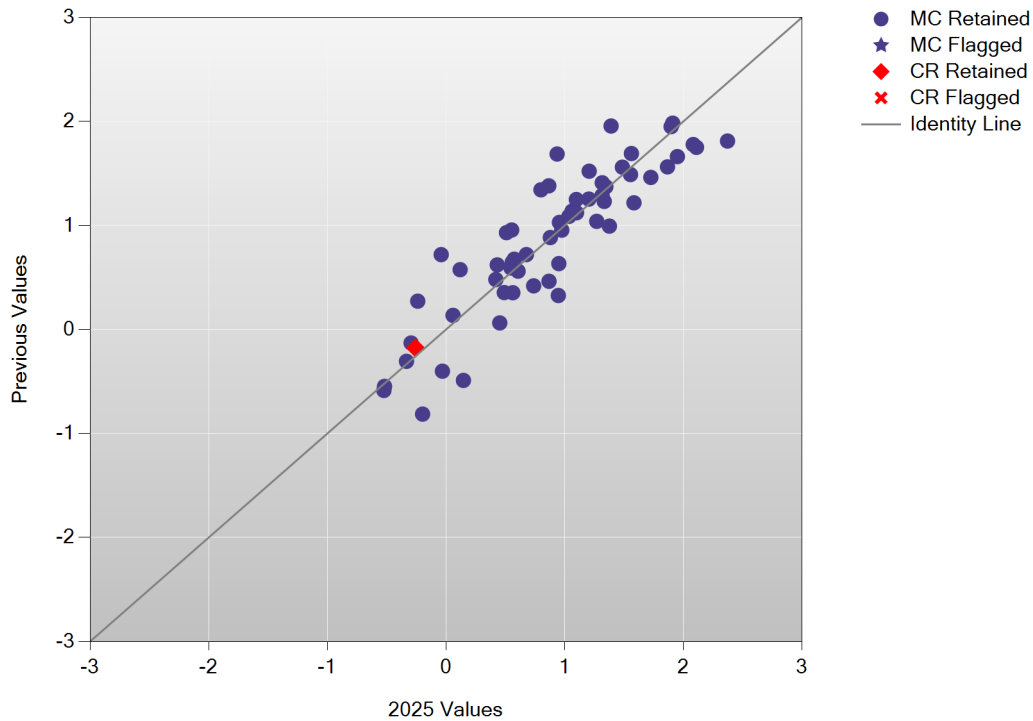
Section 2.1

A/A, B/B, Delta, and Cumulative
Scale Score Distribution Plots

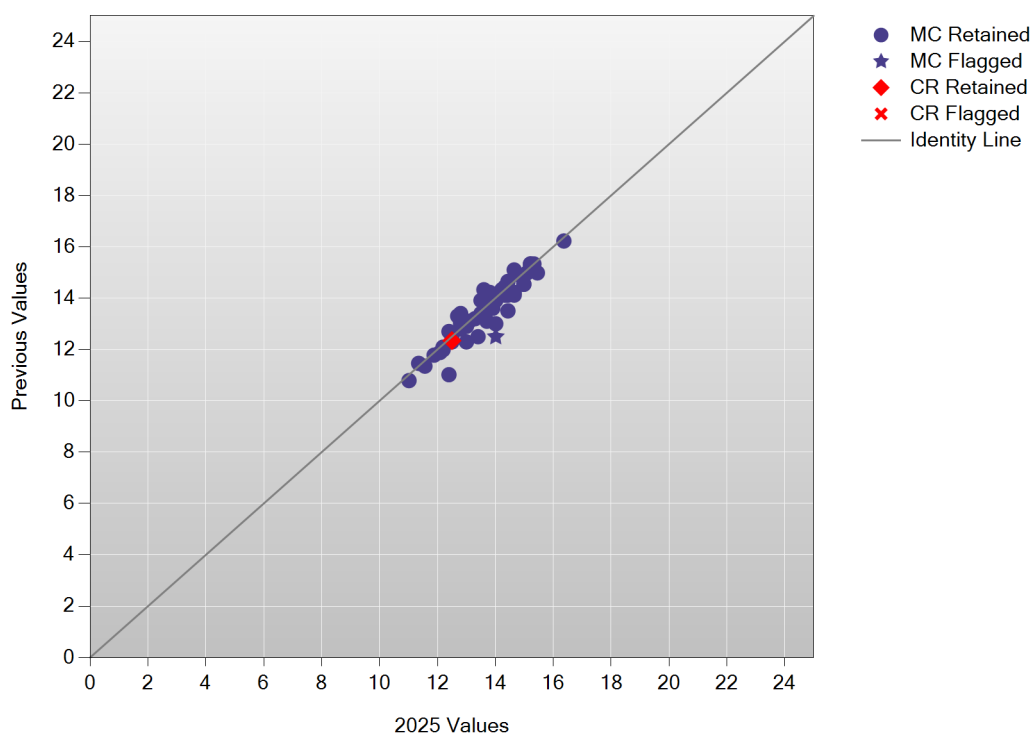
A/A Plot: Science Grade 11



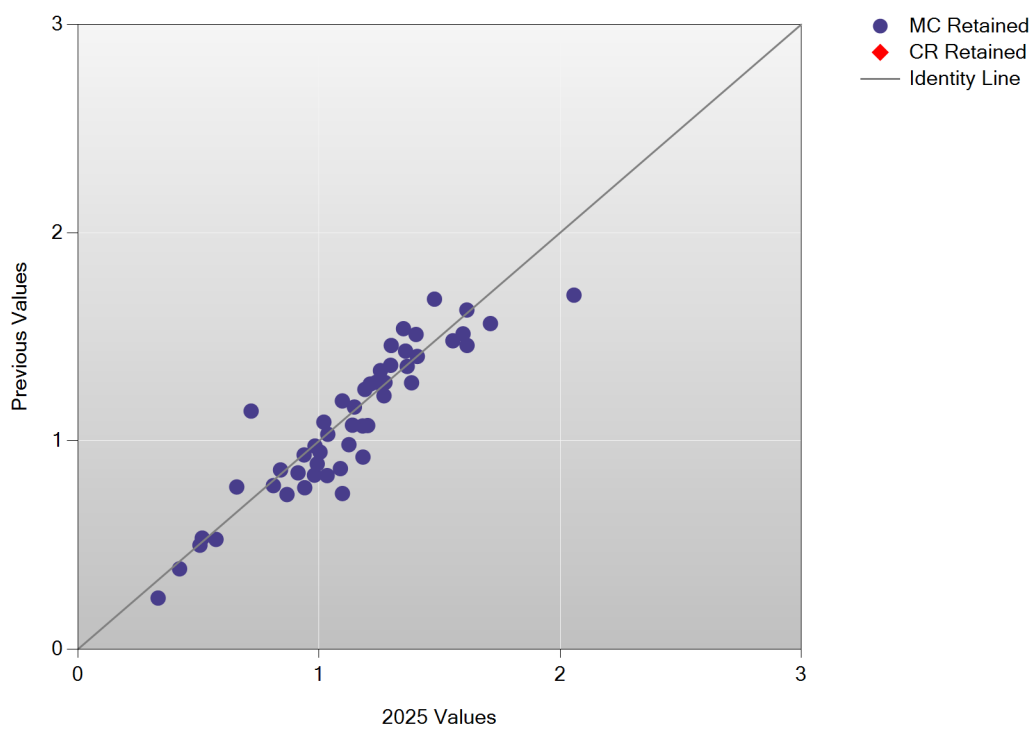
B/B Plot: Science Grade 11



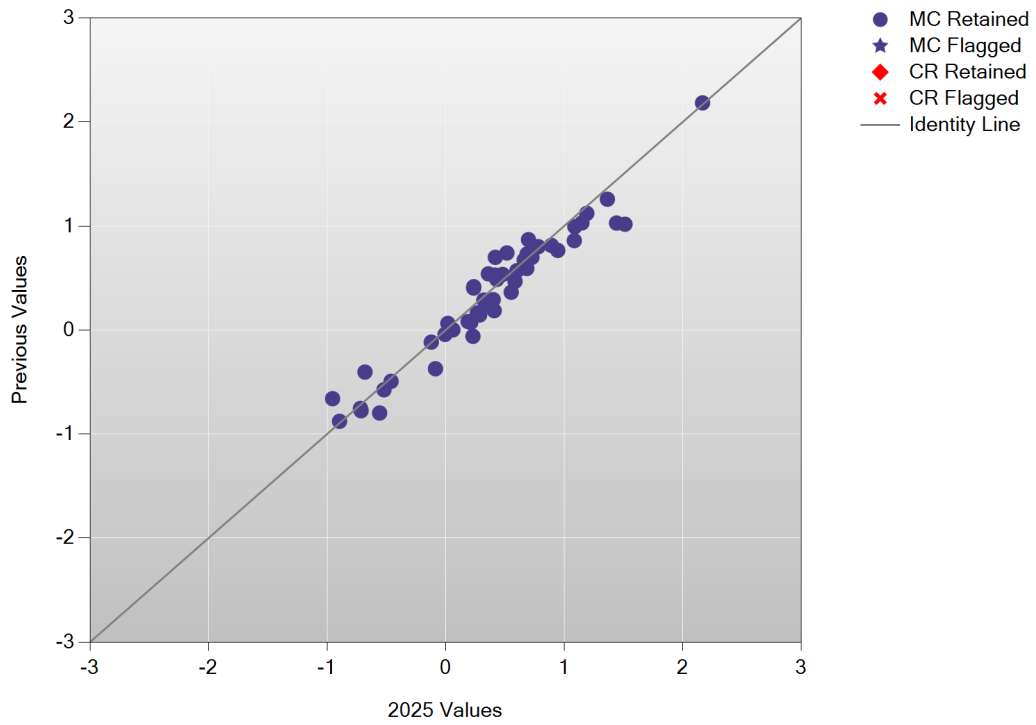
Delta Plot: Science Grade 11



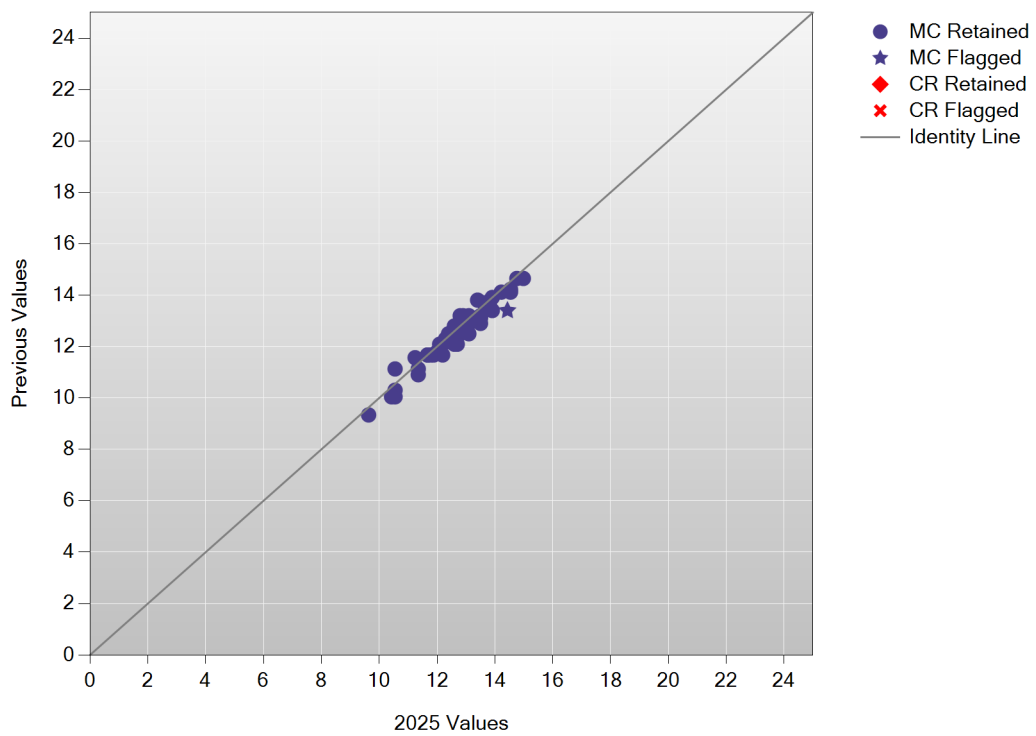
A/A Plot: US History Grade 11



B/B Plot: US History Grade 11



Delta Plot: US History Grade 11



Section 2.2

Cumulative Scale Score Distribution Tables

Table 2.2.1
Cumulative Scale Score Distribution
Science Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
212	BB	5	0.00010	0.00010
213	BB	17	0.00035	0.00045
214	BB	45	0.00093	0.00139
215	BB	55	0.00114	0.00252
216	BB	101	0.00209	0.00461
217	BB	129	0.00267	0.00728
218	BB	170	0.00352	0.01080
219	BB	194	0.00401	0.01481
220	BB	205	0.00424	0.01905
221	BB	225	0.00465	0.02370
222	BB	258	0.00534	0.02904
223	BB	316	0.00654	0.03557
224	BB	286	0.00591	0.04149
225	BB	409	0.00846	0.04995
226	BB	460	0.00951	0.05946
227	BB	413	0.00854	0.06800
228	BB	352	0.00728	0.07528
229	BB	427	0.00883	0.08411
230	BB	457	0.00945	0.09356
231	BB	456	0.00943	0.10299
232	BB	471	0.00974	0.11273
233	BB	469	0.00970	0.12243
234	BB	556	0.01150	0.13393
235	BB	513	0.01061	0.14454
236	BB	524	0.01084	0.15538
237	BB	535	0.01106	0.16644
238	BB	511	0.01057	0.17701
239	BB	547	0.01131	0.18832
240	BB	561	0.01160	0.19993
241	BB	487	0.01007	0.21000
242	BB	490	0.01013	0.22013
243	BB	506	0.01046	0.23060
244	BB	524	0.01084	0.24143
245	BB	507	0.01049	0.25192
246	BB	536	0.01109	0.26300
247	BB	523	0.01082	0.27382
248	BB	559	0.01156	0.28538
249	BB	563	0.01164	0.29702
250	BB	500	0.01034	0.30736
251	BB	511	0.01057	0.31793

Table 2.2.1 (continued)
Cumulative Scale Score Distribution
Science Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
252	BB	540	0.01117	0.32910
253	BB	546	0.01129	0.34039
254	BB	533	0.01102	0.35142
255	BB	530	0.01096	0.36238
256	BB	489	0.01011	0.37249
257	BB	519	0.01073	0.38322
258	BB	503	0.01040	0.39363
259	BB	505	0.01044	0.40407
260	BB	511	0.01057	0.41464
261	BB	476	0.00984	0.42448
262	BB	512	0.01059	0.43507
263	BB	503	0.01040	0.44547
264	BB	493	0.01020	0.45567
265	BB	465	0.00962	0.46529
266	BB	432	0.00893	0.47422
267	BB	484	0.01001	0.48423
268	BB	429	0.00887	0.49310
269	BB	480	0.00993	0.50303
270	BB	490	0.01013	0.51316
271	BB	475	0.00982	0.52299
272	BB	457	0.00945	0.53244
273	BB	483	0.00999	0.54243
274	BB	464	0.00960	0.55202
275	BB	450	0.00931	0.56133
276	BB	448	0.00927	0.57060
277	BB	376	0.00778	0.57837
278	B	575	0.01189	0.59026
279	B	471	0.00974	0.60000
280	B	439	0.00908	0.60908
281	B	435	0.00900	0.61808
282	B	434	0.00898	0.62706
283	B	439	0.00908	0.63613
284	B	487	0.01007	0.64621
285	B	463	0.00958	0.65578
286	B	463	0.00958	0.66536
287	B	431	0.00891	0.67427
288	B	481	0.00995	0.68422
289	B	410	0.00848	0.69270
290	B	460	0.00951	0.70221
291	B	422	0.00873	0.71094

Table 2.2.1 (continued)
Cumulative Scale Score Distribution
Science Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
292	B	398	0.00823	0.71917
293	B	444	0.00918	0.72835
294	B	412	0.00852	0.73687
295	B	423	0.00875	0.74562
296	B	417	0.00862	0.75424
297	B	406	0.00840	0.76264
298	B	430	0.00889	0.77153
299	B	601	0.01243	0.78396
300	P	206	0.00426	0.78822
301	P	401	0.00829	0.79652
302	P	374	0.00773	0.80425
303	P	369	0.00763	0.81188
304	P	344	0.00711	0.81900
305	P	388	0.00802	0.82702
306	P	348	0.00720	0.83422
307	P	412	0.00852	0.84274
308	P	368	0.00761	0.85035
309	P	329	0.00680	0.85715
310	P	317	0.00656	0.86371
311	P	300	0.00620	0.86992
312	P	315	0.00651	0.87643
313	P	284	0.00587	0.88230
314	P	291	0.00602	0.88832
315	P	287	0.00594	0.89426
316	P	257	0.00532	0.89957
317	P	278	0.00575	0.90532
318	P	260	0.00538	0.91070
319	P	248	0.00513	0.91583
320	P	236	0.00488	0.92071
321	P	225	0.00465	0.92536
322	P	215	0.00445	0.92981
323	P	196	0.00405	0.93386
324	P	220	0.00455	0.93841
325	P	208	0.00430	0.94271
326	P	96	0.00199	0.94470
327	A	224	0.00463	0.94933
328	A	167	0.00345	0.95278
329	A	157	0.00325	0.95603
330	A	143	0.00296	0.95899
331	A	130	0.00269	0.96168

Table 2.2.1 (continued)
Cumulative Scale Score Distribution
Science Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
332	A	130	0.00269	0.96437
333	A	129	0.00267	0.96703
334	A	126	0.00261	0.96964
335	A	101	0.00209	0.97173
336	A	104	0.00215	0.97388
337	A	93	0.00192	0.97580
338	A	91	0.00188	0.97768
339	A	101	0.00209	0.97977
340	A	84	0.00174	0.98151
341	A	78	0.00161	0.98312
342	A	63	0.00130	0.98443
343	A	74	0.00153	0.98596
344	A	69	0.00143	0.98738
345	A	62	0.00128	0.98867
346	A	62	0.00128	0.98995
347	A	55	0.00114	0.99109
348	A	35	0.00072	0.99181
349	A	27	0.00056	0.99237
350	A	34	0.00070	0.99307
351	A	41	0.00085	0.99392
352	A	31	0.00064	0.99456
353	A	33	0.00068	0.99524
354	A	25	0.00052	0.99576
355	A	24	0.00050	0.99626
356	A	19	0.00039	0.99665
357	A	18	0.00037	0.99702
358	A	21	0.00043	0.99746
359	A	13	0.00027	0.99773
360	A	13	0.00027	0.99799
361	A	13	0.00027	0.99826
362	A	14	0.00029	0.99855
363	A	12	0.00025	0.99880
364	A	11	0.00023	0.99903
365	A	8	0.00017	0.99919
366	A	4	0.00008	0.99928
367	A	8	0.00017	0.99944
368	A	6	0.00012	0.99957
369	A	1	0.00002	0.99959
370	A	5	0.00010	0.99969
371	A	2	0.00004	0.99973
372	A	1	0.00002	0.99975
373	A	2	0.00004	0.99979
374	A	2	0.00004	0.99983
375	A	1	0.00002	0.99986
376	A	1	0.00002	0.99988
399	A	6	0.00012	1.00000

Table 2.2.2
Cumulative Scale Score Distribution
U.S. History Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
240	BB	8	0.00017	0.00017
241	BB	18	0.00037	0.00054
242	BB	32	0.00066	0.00120
243	BB	105	0.00217	0.00337
244	BB	55	0.00114	0.00451
245	BB	102	0.00211	0.00662
246	BB	94	0.00195	0.00857
247	BB	138	0.00286	0.01143
248	BB	169	0.00350	0.01492
249	BB	168	0.00348	0.01840
250	BB	210	0.00435	0.02275
251	BB	217	0.00449	0.02724
252	BB	283	0.00586	0.03310
253	BB	282	0.00584	0.03894
254	BB	273	0.00565	0.04459
255	BB	302	0.00625	0.05084
256	BB	364	0.00753	0.05837
257	BB	373	0.00772	0.06610
258	BB	396	0.00820	0.07429
259	BB	362	0.00749	0.08179
260	BB	419	0.00867	0.09046
261	BB	354	0.00733	0.09779
262	BB	470	0.00973	0.10752
263	BB	467	0.00967	0.11718
264	BB	409	0.00847	0.12565
265	BB	454	0.00940	0.13505
266	BB	431	0.00892	0.14397
267	BB	474	0.00981	0.15378
268	BB	514	0.01064	0.16442
269	BB	519	0.01074	0.17516
270	BB	544	0.01126	0.18642
271	BB	544	0.01126	0.19769
272	BB	569	0.01178	0.20946
273	BB	555	0.01149	0.22095
274	BB	575	0.01190	0.23286
275	BB	538	0.01114	0.24399
276	BB	554	0.01147	0.25546
277	BB	576	0.01192	0.26738
278	BB	646	0.01337	0.28076
279	BB	600	0.01242	0.29318

Table 2.2.2 (continued)
Cumulative Scale Score Distribution
U.S. History Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
280	BB	623	0.01290	0.30607
281	BB	658	0.01362	0.31969
282	BB	629	0.01302	0.33271
283	BB	639	0.01323	0.34594
284	BB	652	0.01350	0.35944
285	BB	644	0.01333	0.37277
286	BB	641	0.01327	0.38604
287	BB	695	0.01439	0.40042
288	BB	684	0.01416	0.41458
289	BB	792	0.01639	0.43098
290	B	489	0.01012	0.44110
291	B	637	0.01319	0.45428
292	B	658	0.01362	0.46790
293	B	663	0.01372	0.48163
294	B	642	0.01329	0.49492
295	B	670	0.01387	0.50879
296	B	680	0.01408	0.52286
297	B	654	0.01354	0.53640
298	B	667	0.01381	0.55021
299	B	980	0.02029	0.57049
300	P	310	0.00642	0.57691
301	P	654	0.01354	0.59045
302	P	652	0.01350	0.60395
303	P	638	0.01321	0.61715
304	P	603	0.01248	0.62963
305	P	653	0.01352	0.64315
306	P	636	0.01317	0.65632
307	P	638	0.01321	0.66952
308	P	646	0.01337	0.68290
309	P	588	0.01217	0.69507
310	P	612	0.01267	0.70774
311	P	641	0.01327	0.72100
312	P	591	0.01223	0.73324
313	P	579	0.01199	0.74522
314	P	546	0.01130	0.75653
315	P	584	0.01209	0.76861
316	P	541	0.01120	0.77981
317	P	564	0.01167	0.79149
318	P	561	0.01161	0.80310
319	P	470	0.00973	0.81283

Table 2.2.2 (continued)
Cumulative Scale Score Distribution
U.S. History Grade 11

Scale Score	Performance Level	N	Proportion	Cumulative Proportion
320	P	534	0.01105	0.82388
321	P	526	0.01089	0.83477
322	P	476	0.00985	0.84463
323	P	465	0.00963	0.85425
324	P	473	0.00979	0.86404
325	P	434	0.00898	0.87303
326	P	392	0.00811	0.88114
327	P	403	0.00834	0.88948
328	P	383	0.00793	0.89741
329	P	522	0.01081	0.90822
330	A	181	0.00375	0.91196
331	A	331	0.00685	0.91881
332	A	319	0.00660	0.92542
333	A	291	0.00602	0.93144
334	A	270	0.00559	0.93703
335	A	239	0.00495	0.94198
336	A	256	0.00530	0.94728
337	A	231	0.00478	0.95206
338	A	209	0.00433	0.95638
339	A	211	0.00437	0.96075
340	A	192	0.00397	0.96473
341	A	179	0.00371	0.96843
342	A	142	0.00294	0.97137
343	A	153	0.00317	0.97454
344	A	151	0.00313	0.97766
345	A	130	0.00269	0.98036
346	A	126	0.00261	0.98296
347	A	100	0.00207	0.98503
348	A	87	0.00180	0.98683
349	A	85	0.00176	0.98859
350	A	71	0.00147	0.99006
351	A	67	0.00139	0.99145
352	A	57	0.00118	0.99263
353	A	58	0.00120	0.99383
354	A	67	0.00139	0.99522
355	A	32	0.00066	0.99588
356	A	11	0.00023	0.99611
357	A	31	0.00064	0.99675
358	A	15	0.00031	0.99706
359	A	43	0.00089	0.99795
360	A	29	0.00060	0.99855
398	A	70	0.00145	1.00000

Section 2.3

Tabled Delta Analysis Results

Table 2.3.1 Delta Analysis—Science Grade 11

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
186972A	0.34000	0.36000	14.64985	14.43384	1	False	0.08407
186989A	0.52000	0.52000	12.79939	12.79939	1	False	-0.40722
186992A	0.61000	0.59000	11.88272	12.08982	1	False	-0.97054
187933A	0.71000	0.69000	10.78646	11.01660	1	False	-0.92538
187935A	0.59000	0.58000	12.08982	12.19243	1	False	-0.66199
187938A	0.62000	0.61000	11.77808	11.88272	1	False	-0.62943
187996A	0.66000	0.64000	11.35015	11.56616	1	False	-0.94923
187999A	0.41000	0.40000	13.91018	14.01339	1	False	-0.89412
188008A	0.51000	0.50000	12.89972	13.00000	1	False	-0.75655
188454A	0.53000	0.56000	12.69892	12.39612	1	False	0.62203
188458A	0.41000	0.45000	13.91018	13.50265	1	False	0.82054
188459A	0.47000	0.53000	13.30108	12.69892	1	False	1.55092
439223	0.31000	0.27000	14.98340	15.45125	1	False	0.29682
439239	0.32000	0.32000	14.87080	14.87080	1	False	-0.66907
457186	0.45000	0.36000	13.50265	14.43384	1	False	1.66515
457197	0.39000	0.34000	14.11728	14.64985	1	False	0.40463
457199	0.35000	0.31000	14.54128	14.98340	1	False	0.15455
586027	0.31000	0.30000	14.98340	15.09760	1	False	-0.89045
586029	0.37000	0.38000	14.32741	14.22192	1	False	-0.24623
586031	0.43000	0.41000	13.70550	13.91018	1	False	-0.74823
586218	0.55000	0.55000	12.49735	12.49735	1	False	-0.36904
586649	0.44000	0.41000	13.60388	13.91018	1	False	-0.41991
586655	0.39000	0.36000	14.11728	14.43384	1	False	-0.32058
586659	0.56500	0.55000	12.34537	12.49735	2	False	-0.86008
586691	0.46000	0.44000	13.40173	13.60388	1	False	-0.79516
586693	0.48000	0.47000	13.20061	13.30108	1	False	-0.79522
586701	0.46000	0.52000	13.40173	12.79939	1	False	1.53884
586709	0.37000	0.44000	14.32741	13.60388	1	False	1.82867
586711	0.38000	0.42000	14.22192	13.80757	1	False	0.80401
591949	0.43000	0.42000	13.70550	13.80757	1	False	-0.86445
592069	0.55000	0.40000	12.49735	14.01339	1	True	3.50150
592071	0.69000	0.56000	11.01660	12.39612	1	False	2.85603
592073	0.29000	0.29000	15.21354	15.21354	1	False	-0.71239
593424	0.42000	0.41000	13.80757	13.91018	1	False	-0.87913
593426	0.49000	0.49000	13.10028	13.10028	1	False	-0.44525
639009	0.28000	0.28000	15.33137	15.33137	1	False	-0.72729
639014	0.30000	0.34000	15.09760	14.64985	1	False	0.80544
639018	0.28000	0.29000	15.33137	15.21354	1	False	-0.33172
701635	0.54000	0.55000	12.59827	12.49735	1	False	-0.04302
701641	0.50000	0.52000	13.00000	12.79939	1	False	0.24092

Table 2.3.1 (continued)

Delta Analysis

Science Grade 11

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
701654	0.65000	0.66000	11.45872	11.35015	1	False	0.12675
701674	0.57000	0.50000	12.29450	13.00000	1	False	0.75474
701698	0.50000	0.40000	13.00000	14.01339	1	False	1.87757
701703	0.55000	0.46000	12.49735	13.40173	1	False	1.44807
757839	0.38000	0.37000	14.22192	14.32741	1	False	-0.94119
757905	0.38000	0.34000	14.22192	14.64985	1	False	0.06654
786785	0.60000	0.58000	11.98661	12.19243	1	False	-0.96171
786787	0.49000	0.52000	13.10028	12.79939	1	False	0.56489
786789	0.42000	0.44000	13.80757	13.60388	1	False	0.14919
832326	0.49000	0.43000	13.10028	13.70550	1	False	0.51995
850055	0.36000	0.37000	14.43384	14.32741	1	False	-0.25655
850061	0.46000	0.45000	13.40173	13.50265	1	False	-0.82214
850065	0.57000	0.55000	12.29450	12.49735	1	False	-0.93274
850072	0.35000	0.31000	14.54128	14.98340	1	False	0.15455
850074	0.38000	0.36000	14.22192	14.43384	1	False	-0.65868
850076	0.39000	0.37000	14.11728	14.32741	1	False	-0.67786
850120	0.38000	0.35000	14.22192	14.54128	1	False	-0.29796
850122	0.21000	0.20000	16.22568	16.36648	1	False	-0.64412
850124	0.28000	0.29000	15.33137	15.21354	1	False	-0.33172

Table 2.3.2 Delta Analysis—U.S. History Grade 11

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
143252A	0.68000	0.73000	11.12920	10.54875	1	False	2.54359
143254A	0.41000	0.41000	13.91018	13.91018	1	False	-0.33158
143257A	0.39000	0.38000	14.11728	14.22192	1	False	-0.86574
143278A	0.55000	0.52000	12.49735	12.79939	1	False	-0.49363
143307A	0.38000	0.35000	14.22192	14.54128	1	False	-0.49606
143309A	0.46000	0.41000	13.40173	13.91018	1	False	0.53334
143323A	0.63000	0.61000	11.67259	11.88272	1	False	-0.92813
143337A	0.75000	0.73000	10.30204	10.54875	1	False	-0.66388
143349A	0.77000	0.73000	10.04461	10.54875	1	False	0.69139
143364A	0.55000	0.54000	12.49735	12.59827	1	False	-0.93337
143365A	0.44000	0.42000	13.60388	13.80757	1	False	-1.06553
143366A	0.59000	0.59000	12.08982	12.08982	1	False	-0.42945
143371A	0.55000	0.49000	12.49735	13.10028	1	False	1.07427
143416A	0.63000	0.63000	11.67259	11.67259	1	False	-0.45188
143447A	0.58000	0.56000	12.19243	12.39612	1	False	-0.98964
648621	0.48000	0.51000	13.20061	12.89972	1	False	1.19817
648623	0.52000	0.54000	12.79939	12.59827	1	False	0.65671
648625	0.53000	0.51000	12.69892	12.89972	1	False	-1.03195
648627	0.42000	0.46000	13.80757	13.40173	1	False	1.77768
648631	0.39000	0.35000	14.11728	14.54128	1	False	0.05487
648634	0.51000	0.45000	12.89972	13.50265	1	False	1.05264
648636	0.59000	0.54000	12.08982	12.59827	1	False	0.60388
648638	0.59000	0.53000	12.08982	12.69892	1	False	1.12838
652301	0.44000	0.42000	13.60388	13.80757	1	False	-1.06553
652304	0.68000	0.66000	11.12920	11.35015	1	False	-0.84261
652307	0.49000	0.45000	13.10028	13.50265	1	False	-0.00319
652332	0.63000	0.58000	11.67259	12.19243	1	False	0.68569
658018	0.43000	0.44000	13.70550	13.60388	1	False	0.18695
658053	0.63000	0.63000	11.67259	11.67259	1	False	-0.45188
658058	0.45000	0.43000	13.50265	13.70550	1	False	-1.06449
658060	0.64000	0.67000	11.56616	11.24035	1	False	1.24019
658072	0.46000	0.36000	13.40173	14.43384	1	True	3.26205
658076	0.48000	0.52000	13.20061	12.79939	1	False	1.72102
658078	0.43000	0.43000	13.70550	13.70550	1	False	-0.34258
658082	0.49000	0.48000	13.10028	13.20061	1	False	-0.89797
700021	0.53000	0.54000	12.69892	12.59827	1	False	0.12780
700082	0.82000	0.80000	9.33854	9.63352	1	False	-0.36055
700300	0.51000	0.49000	12.89972	13.10028	1	False	-1.04406
700377	0.77000	0.74000	10.04461	10.42662	1	False	0.05499
700443	0.54000	0.53000	12.59827	12.69892	1	False	-0.92662

Table 2.3.2 (continued)

Delta Analysis

U.S. History Grade 11

Item Id	Old P	New P	Old Delta	New Delta	Max	Discard	Std Dist
700938	0.70000	0.66000	10.90240	11.35015	1	False	0.35145
700979	0.63000	0.62000	11.67259	11.77808	1	False	-1.00158
755336	0.34000	0.33000	14.64985	14.75965	1	False	-0.86396
793774	0.50000	0.52000	13.00000	12.79939	1	False	0.66486
793793	0.62000	0.60000	11.77808	11.98661	1	False	-0.94215
793798	0.34000	0.31000	14.64985	14.98340	1	False	-0.44512
793805	0.57000	0.57000	12.29450	12.29450	1	False	-0.41845
793829	0.55000	0.56000	12.49735	12.39612	1	False	0.11997
824114	0.48000	0.49000	13.20061	13.10028	1	False	0.15313
824149	0.48000	0.45000	13.20061	13.50265	1	False	-0.53144

Section 2.4

Tabled B/B Analysis Results

Table 2.4.1 b/b Analysis—Science Grade 11

Item Id	Old b	New b	Std Dist	Flag
186972A	1.24985	1.09865	-0.41871	False
186989A	0.62135	0.43186	-0.22281	False
186992A	-0.30578	-0.33332	-0.97151	False
187933A	-0.54699	-0.51845	-1.09333	False
187935A	-0.12988	-0.29461	-0.32369	False
187938A	-0.58591	-0.52385	-0.93495	False
187996A	0.13611	0.05918	-0.74673	False
187999A	1.02990	0.95589	-0.78051	False
188008A	1.08854	1.03645	-0.88595	False
188454A	1.38177	0.86638	1.30848	False
188458A	1.68745	0.93624	2.42196	False
188459A	0.71993	-0.04090	2.48922	False
439223	1.23208	1.33364	-0.70680	False
439239	1.48985	1.55532	-0.87251	False
457186	0.46379	0.86830	0.71529	False
457197	0.99417	1.37722	0.62515	False
457199	1.21847	1.58427	0.54820	False
586027	1.37315	1.34735	-1.01719	False
586029	1.13823	1.06285	-0.77642	False
586031	0.88434	0.87944	-1.10558	False
586218	0.35537	0.49089	-0.56500	False
586649	0.95539	0.97560	-1.09943	False
586655	1.41097	1.31674	-0.69294	False
586659	-0.17767	-0.25686	-0.72900	False
586691	1.28866	1.31544	-1.06079	False
586693	0.72048	0.67777	-0.92231	False
586701	0.93163	0.50954	0.87528	False
586709	0.95681	0.55340	0.78597	False
586711	1.34333	0.80020	1.44112	False
591949	1.12262	1.10033	-1.02828	False
592069	0.06311	0.45300	0.63690	False
592071	-0.40102	-0.03020	0.53597	False
592073	1.56114	1.48655	-0.78959	False
593424	1.25439	1.20424	-0.89886	False
593426	0.56142	0.60736	-0.98597	False
639009	1.98423	1.91084	-0.80472	False
639014	1.52231	1.20659	0.35679	False
639018	1.95102	1.89732	-0.89752	False
701635	-0.81363	-0.19758	1.69177	False
701641	0.57557	0.11908	1.04663	False
701654	0.27225	-0.23833	1.31035	False
701674	0.35361	0.56311	-0.21359	False
701698	0.41976	0.73888	0.30865	False
701703	0.63360	0.95088	0.30467	False
757839	1.69241	1.56249	-0.52966	False
757905	1.77896	2.08385	0.27133	False
786785	-0.48964	0.14674	1.79557	False
786787	0.67591	0.57536	-0.64654	False
786789	0.59148	0.54813	-0.91640	False
832326	0.32725	0.94689	1.73425	False
850055	1.95681	1.39148	1.53291	False
850061	0.64693	0.55844	-0.70319	False
850065	0.48090	0.42007	-0.83089	False
850072	1.66247	1.95027	0.18755	False
850074	1.81251	2.37308	1.48672	False
850076	1.46265	1.72645	0.06908	False
850120	1.04027	1.27057	-0.09948	False
850122	1.56386	1.86716	0.25898	False
850124	1.75154	2.11144	0.53205	False

Table 2.4.2
b/b Analysis
U.S. History Grade 11

Item Id	Old b	New b	Std Dist	Flag
143252A	-0.40357	-0.68020	1.44811	False
143254A	1.03102	1.15007	-0.78966	False
143257A	0.69976	0.72851	-0.85423	False
143278A	0.29227	0.40151	-0.51576	False
143307A	1.02914	1.44250	1.95608	False
143309A	0.76630	0.94637	-0.08964	False
143323A	0.00240	0.06190	-0.83628	False
143337A	-0.77562	-0.71341	-0.42619	False
143349A	-0.49270	-0.46009	-0.84218	False
143364A	0.57098	0.60165	-0.93583	False
143365A	0.67368	0.66282	-0.49772	False
143366A	0.07378	0.21253	-0.13248	False
143371A	-0.05995	0.23088	1.35200	False
143416A	-0.37209	-0.08593	1.46284	False
143447A	0.18698	0.41021	0.59942	False
648621	0.48744	0.43182	-0.17239	False
648623	0.41675	0.23736	0.94696	False
648625	2.18459	2.16795	0.30349	False
648627	0.74156	0.51729	1.52618	False
648631	0.81418	0.89247	-1.06255	False
648634	0.86018	1.08598	0.29042	False
648636	0.16458	0.26689	-0.51724	False
648638	0.25379	0.33482	-0.75982	False
652301	0.65050	0.67647	-0.85267	False
652304	-0.75309	-0.71849	-0.69483	False
652307	0.46832	0.58576	-0.52636	False
652332	0.08207	0.19002	-0.42383	False
658018	0.73149	0.68735	-0.15874	False
658053	-0.04118	-0.00484	-1.03072	False
658058	1.12202	1.19126	-1.02300	False
658060	-0.65952	-0.95368	1.48501	False
658072	1.01779	1.51376	2.73214	False
658076	0.53493	0.48053	-0.16028	False
658078	0.59265	0.68492	-0.82260	False
658082	0.40413	0.23685	0.82778	False
700021	0.52842	0.41796	0.35933	False
700082	-0.87699	-0.89464	-1.12085	False
700300	0.80213	0.78217	-0.34931	False
700377	-0.57379	-0.51981	-0.60277	False
700443	0.28865	0.32297	-1.10951	False
700938	-0.79737	-0.55648	1.25098	False
700979	-0.11580	-0.12077	-0.94313	False
755336	1.25795	1.36511	-1.01279	False
793774	0.69956	0.41955	2.02525	False
793793	0.14521	0.28704	-0.13908	False
793798	0.99416	1.09092	-0.97931	False
793805	0.06444	0.01908	-0.47729	False
793829	0.54035	0.36079	1.00968	False
824114	0.86961	0.69937	1.08561	False
824149	0.36377	0.55331	0.19777	False

Section 2.5

Final Item Parameters

Table 2.5.1 IRT Parameters and Measures of Standard Error for Dichotomous Items—Science Grade 11

Item ID	a	SE(a)	b	SE(b)	c	SE(c)
186972A	0.88532	0.02200	1.09386	0.01516	0.16429	0.00501
186989A	1.21962	0.02351	0.43776	0.01279	0.24328	0.00526
186992A	0.82867	0.01146	-0.31515	0.01588	0.02255	0.00715
187933A	1.04389	0.01694	-0.49731	0.01915	0.12710	0.01024
187935A	0.74861	0.01089	-0.27706	0.01815	0.02280	0.00777
187938A	0.55837	0.00901	-0.50262	0.02813	0.02438	0.01031
187996A	1.62935	0.02907	0.07106	0.01102	0.30234	0.00534
187999A	1.11723	0.02529	0.95339	0.01260	0.21437	0.00432
188008A	0.74226	0.02535	1.03265	0.02525	0.31371	0.00750
188454A	0.52442	0.02446	0.86531	0.05553	0.33720	0.01406
188458A	0.27107	0.01493	0.93405	0.13850	0.07771	0.02871
188459A	0.67152	0.01313	-0.02742	0.02682	0.05319	0.01089
439223	1.33998	0.03071	1.32508	0.01089	0.14871	0.00282
439239	1.57528	0.04697	1.54320	0.01264	0.24054	0.00276
457186	1.69589	0.03098	0.86720	0.00781	0.16364	0.00292
457197	1.28194	0.03470	1.36796	0.01295	0.23278	0.00331
457199	1.17638	0.03582	1.57169	0.01574	0.21952	0.00329
586027	1.25320	0.03089	1.33857	0.01207	0.17930	0.00316
586029	1.51081	0.03358	1.05863	0.00990	0.23013	0.00325
586031	1.56822	0.03216	0.87816	0.00934	0.23437	0.00342
586218	1.74641	0.03413	0.49584	0.00958	0.31235	0.00405
586649	1.36710	0.03045	0.97278	0.01084	0.24479	0.00368
586655	0.67375	0.02193	1.30845	0.02194	0.16706	0.00693
586691	1.28156	0.03863	1.30717	0.01436	0.33309	0.00368
586693	1.12197	0.02361	0.67973	0.01319	0.23163	0.00499
586701	1.14008	0.02341	0.51419	0.01393	0.25444	0.00545
586709	1.90004	0.03167	0.55735	0.00718	0.18008	0.00322
586711	0.99366	0.02146	0.80019	0.01376	0.18092	0.00508
591949	1.40123	0.03479	1.09551	0.01158	0.28733	0.00359
592069	1.17074	0.01676	0.45856	0.00873	0.04995	0.00351
592071	1.23258	0.01846	-0.01689	0.01145	0.11756	0.00578
592073	1.21916	0.03265	1.47553	0.01352	0.17990	0.00309
593424	1.71190	0.04456	1.19775	0.01043	0.30355	0.00310
593426	1.54443	0.02981	0.61045	0.00984	0.26038	0.00400
639009	1.82004	0.07327	1.89302	0.01635	0.24368	0.00240
639014	1.64510	0.03812	1.20006	0.00954	0.21601	0.00288
639018	2.36844	0.10205	1.87971	0.01366	0.26079	0.00227
701635	0.37963	0.01053	-0.18158	0.07242	0.04951	0.01993
701641	0.45030	0.01395	0.13000	0.06273	0.07258	0.01932

Table 2.5.1 (continued)
IRT Parameters and Measures of Standard Error for Dichotomous Items
Science Grade 11

Item ID	a	SE(a)	b	SE(b)	c	SE(c)
701654	0.86346	0.01802	-0.22168	0.02777	0.22615	0.01159
701674	1.33092	0.02605	0.56691	0.01139	0.25238	0.00457
701698	1.58935	0.02850	0.73986	0.00834	0.17803	0.00331
701703	1.45270	0.03421	0.94846	0.01120	0.30490	0.00374
757839	1.44232	0.04712	1.55026	0.01453	0.29656	0.00307
757905	1.04547	0.05034	2.06325	0.03128	0.28719	0.00345
786785	0.73646	0.01795	0.15721	0.03049	0.21305	0.01116
786787	1.21144	0.02576	0.57896	0.01345	0.28777	0.00508
786789	1.01900	0.01881	0.55217	0.01274	0.13364	0.00520
832326	0.76830	0.02115	0.94453	0.02001	0.20021	0.00687
850055	0.34508	0.01704	1.38199	0.06611	0.06630	0.01847
850061	1.46146	0.02556	0.56231	0.00926	0.18872	0.00392
850065	1.78860	0.03372	0.42616	0.00928	0.30059	0.00410
850072	1.28056	0.05344	1.93182	0.02261	0.26588	0.00295
850074	0.67467	0.04285	2.34784	0.05567	0.28808	0.00557
850076	1.21492	0.04606	1.71159	0.01938	0.30844	0.00336
850120	1.14912	0.02936	1.26302	0.01308	0.21141	0.00369
850122	1.52918	0.04892	1.85004	0.01576	0.14368	0.00218
850124	1.33409	0.06230	2.09040	0.02650	0.25142	0.00273

Table 2.5.2
IRT Parameters and Measures of Standard Error for Polytomous Items
Science Grade 11

Item ID	a	SE(a)	b	SE(b)	d0	SE(d0)	d1	SE(d1)	d2	SE(d2)
586659	0.81960	0.00496	-0.23991	0.00655	0.89853	0.00869	-0.89853	0.00791	0.00000	0.00000

Table 2.5.3 IRT Parameters and Measures of Standard Error for Dichotomous Items—U.S. History Grade 11

Item ID	a	SE(a)	b	SE(b)	c	SE(c)
143252A	1.23355	0.01791	-0.74118	0.01472	0.10442	0.00860
143254A	0.49871	0.01920	1.14766	0.03655	0.15992	0.01129
143257A	1.00361	0.01862	0.71261	0.01178	0.11622	0.00435
143278A	1.22960	0.02269	0.37515	0.01218	0.22865	0.00500
143307A	0.40762	0.01805	1.44945	0.04321	0.09038	0.01331
143309A	1.35852	0.02963	0.93745	0.01079	0.23939	0.00361
143323A	1.14394	0.02096	0.02467	0.01545	0.25157	0.00663
143337A	1.11062	0.01626	-0.77545	0.01674	0.08688	0.00944
143349A	1.25636	0.02157	-0.51402	0.01696	0.25287	0.00866
143364A	0.78479	0.02118	0.58169	0.02428	0.29234	0.00777
143365A	1.17405	0.02146	0.64482	0.01082	0.15956	0.00415
143366A	1.99360	0.03469	0.18012	0.00829	0.27626	0.00411
143371A	0.83944	0.01468	0.19906	0.01625	0.08310	0.00670
143416A	1.06242	0.01917	-0.12789	0.01746	0.22624	0.00768
143447A	1.00102	0.02203	0.38413	0.01774	0.29234	0.00641
648621	1.10275	0.02124	0.40643	0.01369	0.22255	0.00543
648623	1.43258	0.02362	0.20575	0.01015	0.20334	0.00466
648625	0.32127	0.03363	2.19812	0.10047	0.36005	0.02085
648627	1.21518	0.02169	0.49463	0.01102	0.17943	0.00447
648631	1.05421	0.02051	0.88182	0.01134	0.12622	0.00388
648634	1.08853	0.02962	1.08152	0.01517	0.30924	0.00438
648636	1.65760	0.02727	0.23622	0.00885	0.21593	0.00418
648638	0.95198	0.01890	0.30633	0.01680	0.20476	0.00660
652301	1.56363	0.02773	0.65891	0.00852	0.18436	0.00339
652304	0.55418	0.01186	-0.78069	0.05147	0.06715	0.01938
652307	1.50657	0.02676	0.56530	0.00910	0.20223	0.00373
652332	0.63743	0.01742	0.15689	0.03831	0.22773	0.01241
658018	1.31631	0.02497	0.67014	0.01037	0.20412	0.00396
658053	1.54743	0.02604	-0.04420	0.01096	0.25990	0.00539
658058	1.34040	0.03620	1.19017	0.01291	0.30892	0.00355
658060	0.48957	0.00667	-1.02341	0.01602	0.00000	0.00000
658072	0.69533	0.02481	1.52299	0.02389	0.21553	0.00596
658076	1.25876	0.02434	0.45670	0.01218	0.25496	0.00482
658078	1.15275	0.02194	0.66763	0.01138	0.17720	0.00431
658082	1.19812	0.01771	0.20522	0.00997	0.09494	0.00442
700021	1.56275	0.02905	0.39213	0.01014	0.27882	0.00429
700082	1.30795	0.02134	-0.96248	0.01811	0.18305	0.01117
700300	1.16406	0.02693	0.76799	0.01369	0.29765	0.00463
700377	0.98781	0.01890	-0.57565	0.02550	0.26106	0.01146

Table 2.5.3 (continued)
IRT Parameters and Measures of Standard Error for Dichotomous Items
U.S. History Grade 11

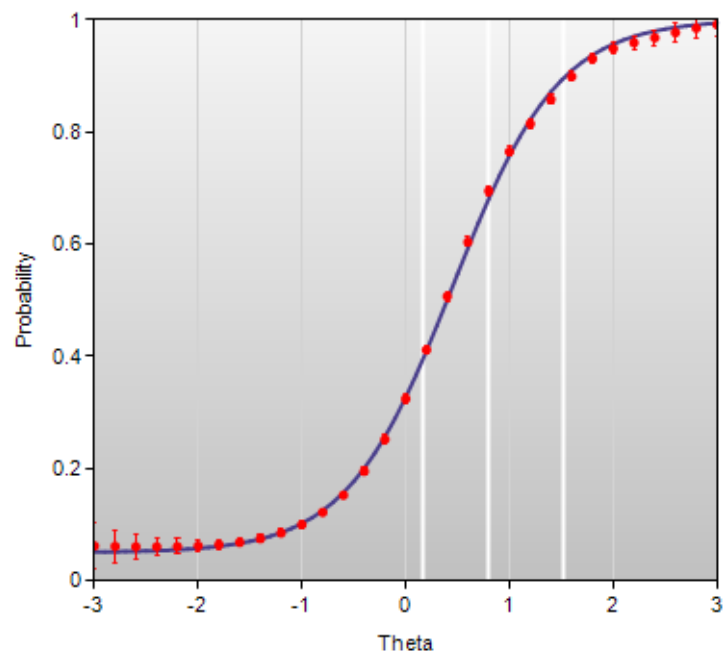
Item ID	a	SE(a)	b	SE(b)	c	SE(c)
700443	1.36346	0.02349	0.29410	0.01072	0.21427	0.00469
700938	0.91085	0.01307	-0.61350	0.01775	0.04319	0.00873
700979	1.32289	0.02038	-0.16384	0.01176	0.17037	0.00595
755336	0.97225	0.02642	1.36959	0.01568	0.19148	0.00394
793774	0.81350	0.01842	0.39377	0.02125	0.21256	0.00771
793793	0.94997	0.02151	0.25702	0.02070	0.31336	0.00732
793798	0.96106	0.02029	1.08662	0.01250	0.11014	0.00380
793805	0.88399	0.01635	-0.01952	0.01957	0.15562	0.00822
793829	1.06213	0.02175	0.33313	0.01593	0.26899	0.00610
824114	0.90841	0.02170	0.68254	0.01757	0.25325	0.00603
824149	1.14511	0.02102	0.53181	0.01168	0.17673	0.00463

Section 2.6

Fit Plots of Watchlist Items

Initial Calibration

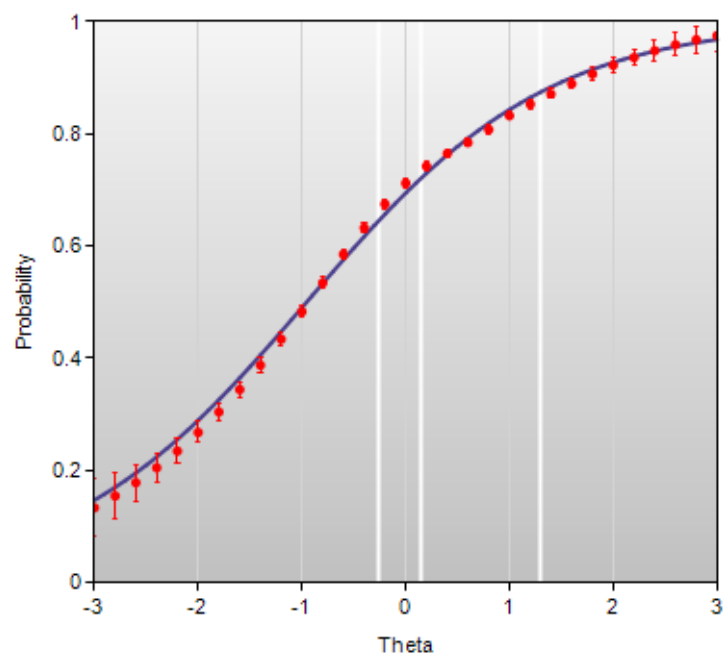
Science Grade 11: 592069



Beta Chart

Initial Calibration

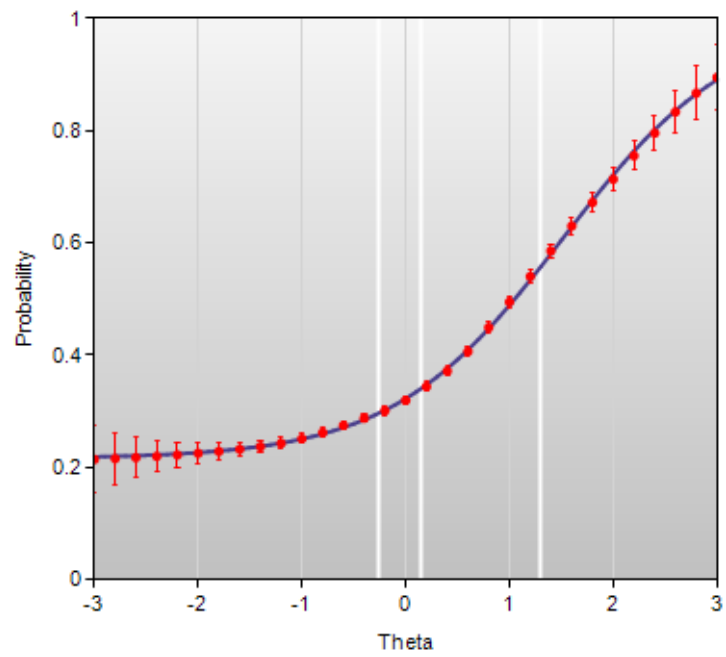
US History Grade 11: 658060



Beta Chart

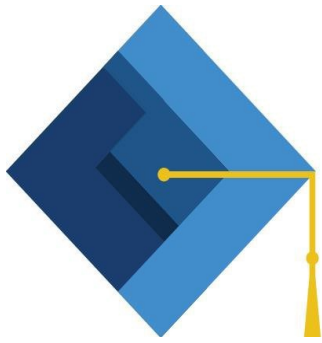
Initial Calibration

US History Grade 11: 658072



Beta Chart

APPENDIX N
2019 CCRA STANDARD SETTING
REPORT



OKLAHOMA STATE DEPARTMENT OF
EDUCATION
— CHAMPION EXCELLENCE —

Oklahoma College and Career Readiness

Standard-Setting Report

June 5-6, 2019

Oklahoma City, OK

TABLE OF CONTENTS

Chapter 1. Overview of Standard-Setting Procedures	4
Chapter 2. Tasks Completed Prior To Standard-Setting	4
2.1 Creation of Performance Level Descriptors	4
2.2 Preparation of Materials for Panelists	5
2.3 Preparation of Presentation Materials	5
2.4 Preparation of Instructions for Facilitators	5
2.5 Preparation of Systems and Materials for Analysis During the Meeting	6
2.6 Selection of Panelists	6
Chapter 3. Tasks Completed During the Standard-Setting Meeting	7
3.1 Overview of the Bookmark Method	7
3.2 General Orientation and Panelist Training	7
3.3 Lead Facilitator Training	7
3.4 Review of Assessment Materials	7
3.5 Completion of the Item Map Form	8
3.6 Review of Performance Level Descriptors	8
3.7 Review of Performance Level Descriptors and Definition of Borderline Students	9
3.8 Rating Rounds and Feedback	10
3.8.1 Practice Round	10
3.8.2 Round 1 Judgments and Results	10
3.8.3 Round 2 Judgments and Results	11
3.8.4 Round 3 Judgments and Results	12
3.8.5 Round 4 Judgments and Results	12
Chapter 4. Tasks Completed After the Standard-Setting Meeting	14
4.1 Analysis and Review of Panelists' Feedback	14
4.2 Policy Adjustments	14
4.3 Preparation of Standard-Setting Report	14
REFERENCES	15
APPENDICES	16

Chapter 1. Overview of Standard-Setting Procedures

The purpose of this report is to summarize the activities involved in the standard-setting process for the Oklahoma College and Career Readiness Assessment (CCRA) in high school science (SCI) on behalf of the Oklahoma State Department of Education (SDE). The need for standard setting arises from the fact that this is a new assessment that was administered for the first time in 2019. For such new assessments, performance standards must be set. The primary goal of the standard setting was to determine the knowledge, skills, and abilities (KSAs) that students must demonstrate to be classified into each of the student status levels (performance levels).

The standard-setting process used was the bookmark procedure (see, e.g., Lewis et al., 1996; Mitzel et al., 2000; Cizek & Bunch, 2007). There were two main reasons this method was chosen. First, the assessment consists primarily of multiple-choice items but also includes some constructed-response items, and the bookmark procedure is appropriate for use with assessments that contain primarily or exclusively multiple-choice items, scaled using item response theory (IRT; Cizek & Bunch, 2007). Second, the modified bookmark method has been used successfully to establish performance standards for Oklahoma in the past (CTB/McGraw-Hill, 2013, 2014; Measured Progress, 2015).

The standard-setting meeting was held from June 5th through June 6th of 2019. In all, 12 panelists participated in the process and were organized into 2 groups of 6 panelists each plus a facilitator provided by Measured Progress. In initial rounds, panelists were organized according to the domain (Life Sciences or Physical Sciences) in which each panelist had the most professional experience. In later rounds, panelists were organized into a single panel.

This report is organized into three major sections, describing tasks completed prior to, during, and after the standard-setting meeting.

Chapter 2. Tasks Completed Prior To Standard-Setting

2.1 Creation of Performance Level Descriptors

Oklahoma State Statute: Title 70. Schools, Chapter 22 – Testing and Assessment, Section 1210.541 – Student Performance Levels and Cut Scores – Accountability System mandates the adoption of “a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.” The law states that performance levels must be labeled and defined as follows:

1. Advanced, which shall indicate that students demonstrate superior performance on challenging subject matter;
2. Proficient, which shall indicate that students demonstrate mastery over appropriate grade-level subject matter and that students are ready for the next grade, course, or level of education, as applicable;
3. Basic, which shall indicate that students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level or course; and
4. Below Basic, which shall indicate that students have not performed at least at the limited knowledge level.

In 2016, the 29 Oklahoma educators who formed the science PLD committees, members of the Oklahoma SDE, and three Measured Progress staff members met for a three-day PLD writing meeting in Oklahoma City. The purpose of the meetings was to write PLDs for grades 5, 8 and high school that describe what students know and are able to display on a statewide assessment of the Oklahoma academic standards. The descriptors are used to provide a common understanding of each performance level for recommending cut scores during standard setting and to inform stakeholders on how to interpret student test scores.

After introductions of those in attendance at the PLD writing meetings, a brief overview of the purpose of PLDs, and an explanation of the PLD writing process, the Oklahoma PLD committees used the standards and the SDE test and item specifications document to begin development of the PLDs. To ensure that the committee members focused on the state-adopted standards and objectives, the committee members were not shown any items that appeared on the assessment.

Independently, PLD committee members filled in the PLD tables by writing down the skills and knowledge students would demonstrate in the Advanced, Proficient, and Limited Knowledge levels for each standard and objective. After the individual work was completed, the group discussed and arrived at a consensus on the wording for the performance levels. As a final step, the PLD committee members reviewed and revised the suggested wording for each level to ensure appropriateness and consistency, and that each level indicated a trajectory of students' knowledge of the content.

At this 2016 meeting the committee members dedicated to high school completed the PLDs for the Life Science domain of the assessment. In February of 2019 a second group was convened to define the PLDs for Physical Science. This meeting was conducted virtually, with some participants in the state department office in OK and the rest on a web conference with the facilitator from Measured Progress. Prior

to the meeting the participants were provided with materials to review, including the Life Science PLDs for reference. The same process was followed as described above to create the Physical Science PLD descriptions.

2.2 Preparation of Materials for Panelists

The following materials were assembled for presentation to the panelists at the standard-setting meeting in paper or digital form (as indicated):

- PLDs (paper)
- Meeting agendas (paper)
- Nondisclosure forms (paper)
- Test booklets (paper)
- Answer keys/scoring rubrics (paper)
- Ordered item booklets (paper)
- Item map forms (digital)
- Rating forms (digital)
- Evaluation forms (digital)

Copies of the PLDs, meeting agenda, nondisclosure form, sample item map form, sample rating form, and evaluation form are included in Appendices A through F.

2.3 Preparation of Presentation Materials

The PowerPoint presentation used in the opening session was prepared and approved by the SDE and TAC prior to the meeting. A copy of the presentation is included in Appendix A.

2.4 Preparation of Instructions for Facilitators

Scripts were created for the group facilitators to refer to while working through each step of the standard-setting process. This document is included in Appendix B. The facilitators also attended a training session, led by a Measured Progress psychometrician, approximately four weeks before the standard setting. The purpose of the training was to prepare the facilitators for the panel activities and to ensure consistency in the implemented procedures.

2.5 Preparation of Systems and Materials for Analysis During the Meeting

The computational programming used to calculate cutpoints and impact data during the standard-setting meeting was completed and thoroughly tested prior to the standard-setting meeting. See *Section 3.7.2, Round 1 Judgments and Results*, for a description of the analyses performed during standard setting.

2.6 Selection of Panelists

As emphasized in Cizek and Bunch (2007), regardless of the method used, the selection of panelists is an important factor in determining standard-setting outcomes and maximizing the validity of the standard-setting process. The guidance provided by *Standards for Educational and Psychological Testing* (AERA et al., 1999) states that “a sufficiently large and representative group of judges should be involved to provide reasonable assurance that results would not vary greatly if the process were repeated.” Consistent with the above guidance and respecting practical considerations regarding the maximum size of a group that can be successfully managed, the goal was to recruit a standard-setting panel of 10–12 members representing different stakeholder groups to set standards for the CCRA science. Additionally, in consideration of the distinct content of each domain, an attempt was made to ensure the panel equally represented experts in both the LS and PS domains. Targets for the size and composition of the panel were also consistent with federal guidelines as described in *Standards and Assessment Peer Review Guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001* (U.S. Department of Education, 2009).

The SDE selected panelists prior to the standard-setting meeting. The goal for panel selection was to include participants who are primarily teachers, but also to include school administrators, higher education personnel, and stakeholders from other interest groups. Moreover, to the extent possible, panelists were selected to reflect a balance of gender, race/ethnicity, and geographic location. Finally, panelists were selected who were familiar not only with the subject matter, but also with the grade for which they would be setting standards. A list of the panelists is included in Appendix C.

Chapter 3. Tasks Completed During the Standard-Setting Meeting

3.1 Overview of the Bookmark Method

The bookmark method (Lewis et al., 1996; Mitzel et al., 2000; Cizek & Bunch, 2007) involves rank ordering the items by difficulty and asking the panelists to identify the point in the ordered set of items at which the students at the borderline of two adjacent performance levels no longer have at least a two-thirds chance of answering the item correctly.

3.2 General Orientation and Panelist Training

Concerning panelist training, *Standards for Educational and Psychological Testing* (AERA et al., 2014) states the following:

Care must be taken to assure these persons understand what they are to do and that their judgments are as thoughtful and objective as possible. The process must be such that well-qualified participants can apply their knowledge and experience to reach meaningful and relevant judgments that accurately reflect their understandings and intentions. (p. 101)

The training of the panelists began with a general orientation session at the start of the standard-setting meeting. The purpose of the orientation was to ensure that all panelists received the same information about the need for and the goals of standard setting, and about their part in the process. The orientation consisted of three parts. First, Oklahoma Executive Director of State Assessments Craig Walker provided an overview of education policy in the state of Oklahoma, including additional context specific to the CCRA science assessment. Next, a Measured Progress psychometrician, Dr. Matthew Gushta, presented a brief overview of the bookmark procedure and the activities that would occur during the standard-setting meeting. Finally, Measured Progress Lead Program Manager Julie DiBona provided panelists with logistical information (e.g., materials review, content security, attendance).

Once the general orientation was complete, panelists broke out into domain specific groups, where they received more detailed training and completed the first two rounds of the standard-setting activities.

3.3 Lead Facilitator Training

Prior to Day 1, the two facilitators attended a brief training session led by Measured Progress psychometricians Dr. Matthew Gushta and Dr. Frank Padellaro. During this training, expectations for facilitators were set to include leading panelist review of the ordered item booklet, leading panelist development of borderline descriptors, facilitation of panel discussion, collection and review of standard-setting materials, and control of secure materials. Facilitators were separately expected to act as table leaders during the preliminary rounds, ensuring that discussion and logistics within each domain group were conducted fairly and efficiently.

3.4 Review of Assessment Materials

The first step after the opening session was for the panelists to take the test. The purpose of this step was to familiarize the panelists with the assessment and the test taking activities expected of students during

administration. Once panelists completed the test, the answer key was distributed. At this point, panelists were encouraged to discuss any issues regarding items or scoring.

3.5 Completion of the Item Map Form

Panelists were then split into two groups based on domain expertise and each panelist reviewed a domain-specific ordered item booklet item by item, considering the knowledge, skills, and abilities (KSAs) students needed to answer each one. The ordered item booklet contained one item per page, ordered from the easiest item to the most difficult item. The ordered item booklet was created by sorting the items according to their item response theory (IRT)-based difficulty values ($RP_{0.67}$ was used). A three-parameter logistic IRT model was used to calculate the $RP_{0.67}$ values for dichotomous items.

Panelists then completed the item map form using the provided laptop computers. The item map form listed the items in the same order as they were presented in the ordered item booklet. The form included space for the panelists to type in the KSAs required to answer each item correctly and to indicate why they believed each item was more difficult than the previous one. To ensure each panelist was comfortable using the provided laptop computers and understood the mechanics of data entry, Measured Progress Psychometrician Dr. Frank Padellaro reviewed the technology the panelists would use to complete their item maps.

Additionally, the item map form was shaded to show a projected range of expected proficiency, based on historic averages of student performance on state assessments from multiple grades and subjects. Item map entries that would produce percentages of students at or above Proficient comparable to those external assessments were identified as benchmarking locations. The shaded region on the item map form was then calculated as ± 2 standard errors around the IRT-based difficulty of the CCRA benchmarking locations. Table 3-1 identifies the benchmarking region for each booklet.

Table 3-1: CCR Standard-Setting Benchmarking Regions

<i>Subject</i>	<i>Grade</i>	<i>Percentage*</i>	<i>PS OIB Shaded Region</i>	<i>LS OIB Shaded Region</i>	<i>Complete OIB Shaded Region</i>
Science	11	18% – 50%	3-9	4-12	6-21

**OSTP historic % proficient and above grades 3–8 (ELA and mathematics) and grades 5 and 8 SCI were used to generate a predicted range of SCI 11 % proficient or above performance.*

After working individually, panelists had the opportunity to discuss the item map with members of their domain-specific group and make necessary additions or adjustments. The purpose of this step was to ensure that panelists became familiar with the ordered item booklet and understood the relationships among the ordered items.

3.6 Review of Performance Level Descriptors

Oklahoma State Statute: Title 70. Schools, Chapter 22 – Testing and Assessment, Section 1210.541 – Student Performance Levels and Cut Scores – Accountability System mandates the adoption of “a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.” The law states that performance levels must be labeled and defined as follows:

1. Advanced, which shall indicate that students demonstrate superior performance on challenging subject matter;
2. Proficient, which shall indicate that students demonstrate mastery over appropriate grade-level subject matter and that students are ready for the next grade, course, or level of education, as applicable;
3. Basic, which shall indicate that students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level or course; and
4. Below Basic, which shall indicate that students have not performed at least at the basic level.

In June of 2019, 12 Oklahoma educators, members of the SDE, and five Measured Progress staff members met for a two-day standard-setting meeting in Oklahoma City. Panelists discussed performance level descriptors (PLD), which describe what students know and are able to display on a statewide assessment of the Oklahoma academic standards. The descriptors are used to provide a common understanding of each performance level for recommending cut scores during standard setting and to inform stakeholders of how to interpret student test scores. Panelists then worked to define descriptors of a borderline level student. A borderline student is one who is minimally able to meet the requirements set by the descriptors for each performance level.

After introductions of those in attendance, a brief overview of the meeting's purpose, and an explanation of the standard-setting process, the panelists were organized into groups to begin setting standards for the Oklahoma CCR Science assessment. According to their professional experience, the panelists were organized into Life Science (LS) and Physical Science (PS) groups. Independently, standard-setting committee members filled in the item mapping tables by writing down the knowledge, skills, and abilities necessary for a student to be successful on each item within the subset of items relevant to the domain to which the group was assigned. After the individual work was completed, each group carefully reviewed and discussed the PLDs for Proficient, Advanced, Basic, and Below Basic as they applied to their domain. This understanding was used within the LS and PS groups to separately discuss and arrive at consensus on the definition of a borderline student for each of the Basic, Proficient, and Advanced performance levels. After developing a working understanding of the PLDs and defining borderline students at each cut, the panelists engaged in the standard-setting process in order to recommend the cuts between performance levels.

3.7 Review of Performance Level Descriptors and Definition of Borderline Students

Next, panelists reviewed the Performance Level Descriptors (PLDs). This important step was designed to ensure that panelists thoroughly understood the KSAs needed for students to be classified into performance levels (Below Basic, Basic, Proficient, and Advanced). Panelists first reviewed the PLDs on their own and then participated in group discussion of the PLDs, clarifying each level. Afterward, panelists developed consensus definitions of borderline students—that is, students who have only barely qualified for a

particular performance level. Bulleted lists of characteristics for each level were generated based on the whole-group discussion and posted in the room for reference throughout the bookmark process. Note that the purpose of this step was to clarify and add specificity to the PLDs based on the KSAs, paying particular attention to the definitions of the borderline students.

The bulleted lists were developed as working documents to be used by the panelists for the purposes of standard setting. They supplemented the PLDs, which provide the official definitions of each performance level, by specifically addressing the KSAs that define the borderline of each level.

The PLDs are provided in Appendix D.

3.8 Rating Rounds and Feedback

3.8.1 Practice Round

Next, the panelists completed a practice round of ratings. The purpose of the practice round was to familiarize the panelists with all the materials they would be using for the standard-setting process and to walk them through the process of placing bookmarks. In addition to the PLDs and borderline descriptions, panelists were given a practice ordered item booklet, which consisted of 10 items representing the range of difficulty on the test, and a practice rating form.

Within each domain-specific group, the facilitator explained what each of the materials was and how panelists would use it to make their ratings. Additionally, Measured Progress Psychometrician Dr. Frank Padellaro reviewed the technology panelists would use to complete their ratings, to ensure each panelist understood how to use the tools provided. Then, beginning with the first ordered item and considering the skills and abilities needed to complete it, panelists were instructed to ask themselves, “Would at least two out of three students performing at the borderline of Proficient answer this question correctly?” Panelists considered each ordered item in turn, asking themselves the same question until their answer changed from “yes” (or predominantly “yes”) to “no” (or predominantly “no”). Each panelist practiced placing the Proficient bookmark in the ordered item booklet. The facilitator then led the panelists in a readiness discussion, asking panelists to share the reasoning behind their bookmark placements with the group and assessing each panelist’s understanding of the rating task, borderline students, and the two-thirds rule. At the end of the practice round, panelists completed the practice evaluation form. The evaluation form was designed to ascertain whether the panelists were comfortable moving ahead to the rating task or whether they had lingering questions or issues that needed to be addressed before proceeding to the Round 1 ratings. Facilitators were instructed to glance over each panelist’s evaluation as he or she completed it, to make sure panelists were ready to move on. The results of the training evaluation can be found in Appendix E.

3.8.2 Round 1 Judgments and Results

In the first round, panelists worked individually with the borderline definitions, the item map form, and the ordered item booklet. Beginning with the first ordered item in the shaded region of the domain-specific OIB, described previously, and considering the skills and abilities needed to complete it, panelists asked themselves, “Would at least two out of three students performing at the borderline of Proficient answer this

question correctly?” Panelists considered each ordered item in turn, asking themselves the same question. They placed the bookmark between the two items where their answer changed from “yes” (or predominantly “yes”) to “no” (or predominantly “no”). For the identification of this Proficient cut point, panelists were instructed that placing a bookmark outside the shaded region required explicit written justification by the panelist. Panelists then repeated the process for the other two cut points and used the rating form to record their ratings for each cut point.

After the completion of each round, Measured Progress staff members calculated a variety of statistics which served various functions: feedback to panelists as part of the standard-setting method, reporting to Measured Progress and the SDE as intermediate evidence for the impact of panelists’ judgements, and as quality control metrics. While these statistics were available, only specific results were revealed to panelists as appropriate for the goals of the specific round.

Results for panelist ratings across all rounds are displayed in Appendix F. For each round, Measured Progress staff members calculated the median cut points for the group based on bookmark placements, theta scale cuts, the Median Absolute Deviation (MAD) of the panelists’ cut points, the conditional standard error of measurement (SEM) for each of the scale cuts, and impact data.

Each panelist’s theta scale cut points were found by averaging the $RP_{0.67}$ values of the items on either side of the bookmark placed by that panelist for each cut point. The /Round 1 overall cut points were then determined by calculating the median of the individual cut points obtained from each panelist. The MAD of the panelists’ cut points indicates the extent to which judgments were consistent across panelists and reflects the level of agreement among the ratings with each successive round of ratings. Conditional SEM characterizes the measurement precision for each of the scale cuts. Finally, impact data reflect the percentage of students across the state who would fall into each performance level category according to the total group median cut points.

3.8.3 Round 2 Judgments and Results

The purpose of Round 2 was for panelists to discuss their Round 1 placements and, if necessary, to revise their ratings. Prior to beginning their discussions, the panelists at each table were presented with the median cut points based on their Round 1 ratings for each cut point in that subject and grade. A Measured Progress psychometrician presented this information to the group using a projector and laptop and explained how to use it as they completed their Round 2 discussions. The distribution of panelists’ cut points was presented in terms of location in the ordered item booklet, both as numerical summaries of cut points ranges and graphically, as histograms.

Within both domain-specific groups, panelists were then given the opportunity to share their individual rationales for their bookmark placements in terms of the necessary knowledge and skills for each classification. Panelists were asked to pay particular attention to how their individual ratings compared to those of other panelists in their room to assess whether they were unusually stringent or lenient within the group. Once the discussions were complete, panelists were given the opportunity to revise their Round 1 ratings on the rating form. Panelists were told to set bookmarks according to their *individual* best judgments;

consensus among the panelists was not necessary. They were encouraged to listen to the points made by their colleagues but not to feel compelled to change their bookmark placements.

When Round 2 ratings were complete, Measured Progress staff members calculated the statistics described above and discussed the results with SDE staff. During this discussion, a lack of agreement was noted among some panelists, especially regarding the bookmark associated with the placement of the Advanced cut. This provided an opportunity for Measured Progress and SDE staff to return to the panels for the purpose of clarifying and confirming the judgmental task—answering for each item, "Would at least two out of three students performing at the borderline of the current PLD answer this item correctly?"

3.8.4 Round 3 Judgments and Results

The purpose of Round 3 was for panelists to gather in a single group, regardless of domain-specific expertise, to discuss their Round 2 placements and, if necessary, to revise their ratings. Prior to the discussions, the panelists were separated into domain-specific groups and presented with the median cuts based on Round 2 results. A Measured Progress psychometrician presented the information and explained how to use it, as described in Round 2. Additionally, SDE staff members presented condensed versions of the educational context information originally provided during the opening session.

Following the domain-specific presentations, the panelists were gathered into a single group. During this discussion, domain-specific information was combined and presented according to the entire CCRA Science assessment and content. The lead facilitator, David Harrison, led an extended discussion of the Round 2 results as they applied to the entire CCRA Science form: walking the panelists through the complete ordered item booklet (i.e., LS and PS items), focusing on the KSAs needed for each item and how they related to the overall PLDs, and facilitated synthesis of the borderline definitions into overall concepts of borderline students. In addition, the discussion explored the differences in cut point placement among panelists and across domains. After the discussions, panelists were given another opportunity to revise their bookmark placements, this time considering the entirety of CCRA Science. Once again, the facilitator reminded the panelists to place the bookmarks according to their individual best judgment, and that it was not necessary for them to reach a consensus. When Round 3 ratings were complete, Measured Progress staff members once again calculated the statistics described previously and reviewed these results with SDE staff.

When Round 3 ratings were complete, Measured Progress staff members calculated the usual statistics though in the context of CCRA Science and not separated by domain. The results were discussed with SDE staff, noting a lack of agreement among some panelists – though less so than round 2 – especially regarding the bookmark associated with the placement of the Advanced cut. This provided an opportunity for Measured Progress and SDE staff to return to the panels for the purpose of clarifying and confirming the judgmental task—answering for each item, "Would at least two out of three students performing at the borderline of the current PLD answer this item correctly?"

3.8.5 Round 4 Judgments and Results

Due to the separation of panelists into domain-specific groups in the first two rounds, a fourth round of judgments was planned as part of the standard-setting process, in order to review the results of Round 3

and introduce impact data (the percentage of students in each performance level using the Round 3 cuts). Following the introduction of impact data, the panelists met as a single group to discuss their Round 3 placements and, if necessary, revise their individual ratings

Prior to the discussions, a Measured Progress psychometrician presented the panelists with the median cuts based on Round 3 results, as well as the associated impact data. The lead facilitator then led an extended discussion of the Round 3 results. After discussion, panelists were given a final opportunity to revise their bookmark placements. When Round 4 ratings were complete, Measured Progress staff members once again calculated the various associated statistics.

A summary of the results is provided in Table 3-2, reporting final median cut points on the theta scale and impact data (percentage of students in performance level; percentage of students at-or-above performance level), respectively. Note that disaggregated impact data broken down by demographics are provided in Appendix G.

Table 3-2: CCRA Science Standard Setting: Round 4 Results

<i>Statistic</i>	<i>Below Basic</i>	<i>Basic</i>	<i>Proficient</i>	<i>Advanced</i>
Theta Scale Cuts	-1.52	0.17	0.80	1.53
Percentage of Students at/in Performance Level	53.30%	20.70%	18.10%	7.90%
Percentage of Students at/above Performance Level	100.00%	46.70%	26.00%	7.90%

Chapter 4. Tasks Completed After the Standard-Setting Meeting

Upon conclusion of the standard-setting meeting, several important tasks were completed. These tasks centered on the following: reviewing the standard-setting process and addressing issues presented by the outcomes; presenting the results to the SDE; and making any final revisions or adjustments based on policy considerations, under direction of the SDE.

4.1 Analysis and Review of Panelists' Feedback

The measurement literature sometimes considers the evaluation process to be another product of the standard-setting process (e.g., Reckase, 2001), as it provides important validity evidence supporting the cut points that are obtained. To provide evidence of the participants' views of the standard-setting process, panelists were asked to complete questionnaires after the practice round, after the completion of Round 1, and at the end of the meeting.

After the evaluation forms were completed, panelists' responses were reviewed. This review did not reveal any anomalies in the standard-setting process or indicate any reason that a particular panelist's data should not be included when the final cut points were calculated. In general, participants felt that the recommended cut points were appropriate and that their judgments were based on appropriate information and decision making. The results of the evaluations are presented in Appendix E.

4.2 Policy Adjustments

After all standard-setting activities had been completed and all materials reviewed, the SDE recommended no adjustments to the Round 4 cuts as recommended by panelists at the standard-setting meeting. The full set of cuts as shown in Table 3-2 were presented to the CEQA and approved for use assigning students to performance levels in the 2018–2019 CCRA science assessment.

4.3 Preparation of Standard-Setting Report

Following the final compilation of standard-setting results, Measured Progress prepared this report, which documents the procedures and results of the 2019 standard-setting meeting that was held to establish performance standards for the assessment.

REFERENCES

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: Establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Lewis, D.M., Mitzel, H.C., & Green, D.R. (1996). Standard setting: A bookmark approach. In D.R. Green (Chair), *IRT-based standard setting procedures utilizing behavioral anchoring*. Symposium conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.
- Measured Progress (2015). *Oklahoma Core Curriculum Tests – Geography: Standard Setting Report*. Dover, NH: Measured Progress.
- CTB/McGraw-Hill. (2013). *Oklahoma School Testing Program: Standard Setting Technical Report for OCCT Grades 5 and 8 Science and Writing*. Monterey, CA: Author.
- CTB/McGraw-Hill. (2014). *Oklahoma State Testing Program: Standard setting technical report for OSTP Grade 5 Social Studies, Grade 8 U.S. History, and End-of-Instruction U.S. History*. Monterey, CA: Author.
- Mitzel, H.C., Lewis, D.M., Patz, R.J., & Green, D.R. (2000). The Bookmark Procedure: Cognitive Perspectives on Standard Setting. In G.J. Cizek (Ed.), *Setting Performance Standards: Concepts, Methods, and Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Reckase, M.D. (2001). Innovative methods for helping standard-setting participants to perform their task: The role of feedback regarding consistency, accuracy, and impact. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives* (pp. 159–173). Mahwah, NJ: Lawrence Erlbaum Associates.
- U.S. Department of Education. (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: U.S. Department of Education Office of Elementary and Secondary Education. Retrieved June 10, 2010, from the World Wide Web: www2.ed.gov/policy/elsec/guid/saaprguidance.pdf.

APPENDICES

APPENDIX A—POWERPOINT PRESENTATION

Welcome!



Oklahoma Career and College Readiness Assessment (CCRA)

Standard Setting Science

June 5-6, 2019

Today's Agenda

1. Context and Policy Introduction
2. Standard Setting Process



Oklahoma State Department of Education Staff

Craig Walker
Executive Director of State Assessments



Assessment Report 2017

Oklahoma Legislature directed the State Board of Education to:

- Evaluate Oklahoma's current state assessment system, and make recommendations for its future.

As a result, Oklahoma State Department of Education:

- Held regional meetings across the state to determine stakeholder concerns
- Convened the Oklahoma Assessment & Accountability Task Force to develop recommendations
- Followed the federal requirements and rules as described in ESSA



Recommendations from the Task Force for CCR Assessments

- Score Interpretation
 - Support criterion-referenced interpretations (i.e., performance against the OAS) and report individual claims appropriate for high school students;
 - Provide a measure of performance indicative of being on track to College and Career Readiness (CCR).
 - (1) supported using theoretically related data in standard-setting activities (e.g., measures of college readiness and other nationally available data) and
 - (2) validated empirically using available postsecondary data linking to performance on the college-readiness assessment;



Goals for Oklahoma Schools

- Focus on college- and career- readiness:
College and career ready means that students graduate from high school prepared to enter and succeed in postsecondary opportunities whether college or career.
- **Students** should graduate high school ready for postsecondary success and need to demonstrate they are on-track toward that goal.



Individual Career Academic Plan

Student-driven, multi-measures approach representing indications of college- and career-readiness

- Students' coursework, learning and assessment results
- Students' postsecondary plans, aligned with their career, academic and personal/social goals and financial reality
- Students' records of college- and career-readiness activities



Oklahoma Statute on Performance Levels

The Commission for Educational Quality and Accountability shall determine and adopt a series of student **performance levels** and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.

- The Commission for Educational Quality and Accountability shall have the authority to set cut scores using any method which the State Board of Education was authorized to use in setting cut scores prior to July 1, 2013.

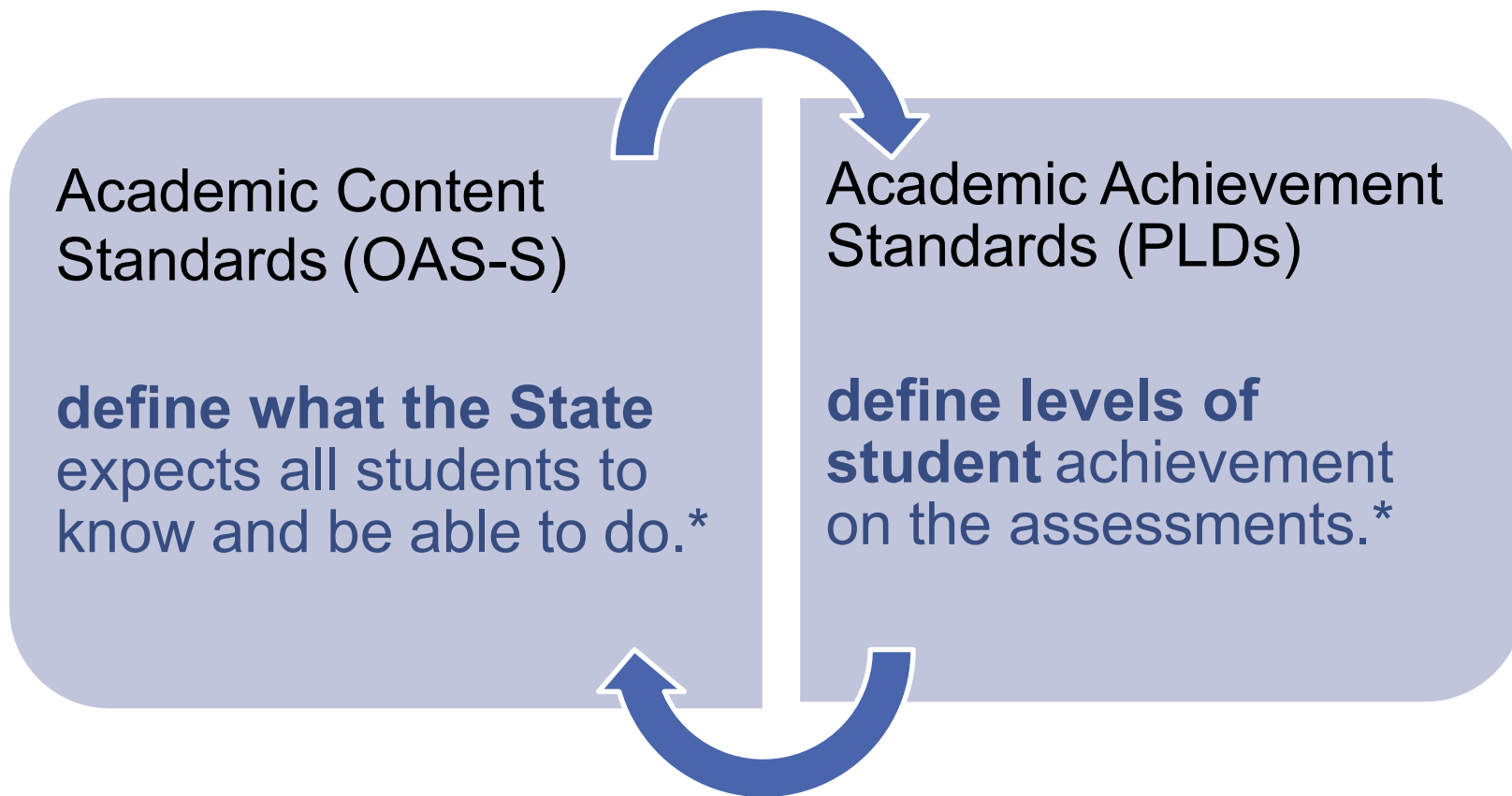


Oklahoma Statute on Performance Levels

- The **performance levels** shall be set by a method that indicates students are ready for the next grade, course, or level of education, as applicable.
- The Commission for Educational Quality and Accountability shall establish panels to review and revise the **performance level descriptors (PLDs)** for each subject and grade level.
- The Commission shall ensure that the criterion-referenced tests developed and administered by the State Board of Education pursuant to the Oklahoma School Testing Program Act in grades three through eight and the tests administered at the high school level are vertically aligned by content across grade levels to ensure consistency, continuity, alignment and clarity.



Content Standards and PLDs



**U.S. Department of Education Peer Review of State Assessment Systems Non-Regulatory Guidance for States, September 25, 2015*



More about PLDs

PLDs provide a narrative account of the knowledge, skills, and abilities **demonstrated by** students in each level of achievement.

PLDs describe what students *know and are able to do* based on the OAS.

PLDs inform stakeholders of how to interpret student test scores in relation to the OAS

PLDs are typically used for standard setting and score reporting.



Purpose and Use of PLDs

Purposes of PLDs

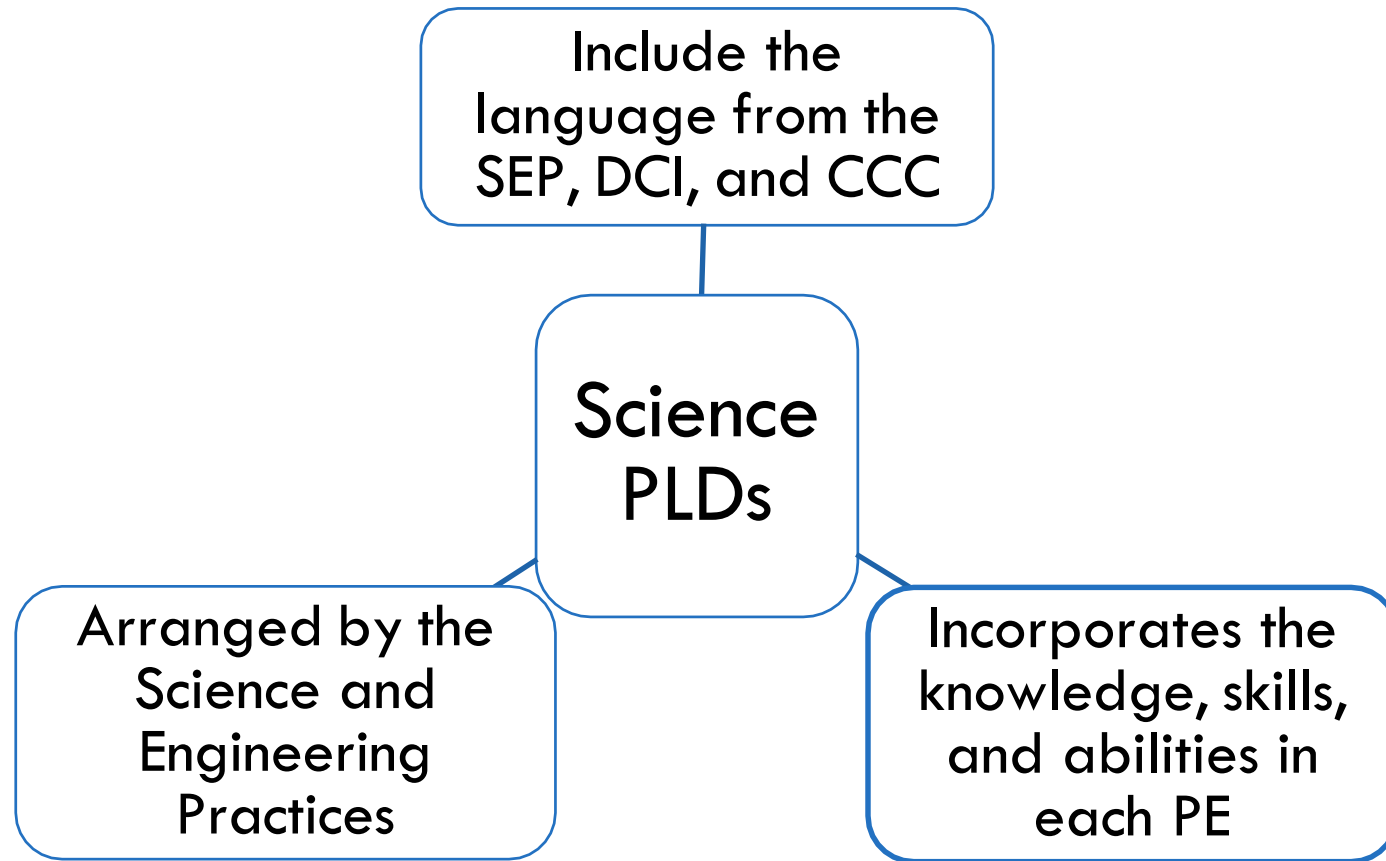
- Inform standard setting
- Inform score interpretation

OK SDE uses for PLDs

- Item and test development
- Standard setting
- Score interpretation



Structure of PLDs for Science



Anatomy of a Science PLD

<div>PS1-1 PS3-2</div> <div>Standard/s</div>	Proficient:
<p>SEP: Develop and Use Models</p> <p>DCI</p> <ul style="list-style-type: none">PS1.A Structure and Properties of MatterPS3.A Definitions of Energy <p>CCC</p> <ul style="list-style-type: none">PatternsEnergy and Matter	<p>Students demonstrate mastery with subject matter and exhibit readiness for college and career. Students scoring at the Proficient level typically use patterns and models to predict how components between or within systems are related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.</p> <p>PLD Knowledge, Skills, and Abilities (KSAs)</p>



OK CCRA Science

Standard Setting

Measured Progress Staff

- Julie DiBona – Lead Program Manager, Client Services
- Matthew Gushta – Director, Research & Analytics
- Frank Padellaro – Psychometrician
- David Harrison – Content Manager, Content Development – State
- Katie Schmidt – Content Specialist II, Content Development - State

Housekeeping

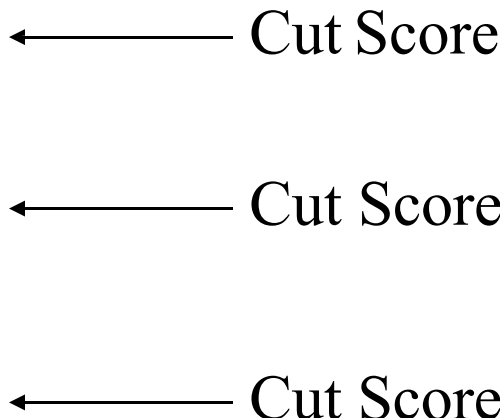
- Folder review
 - Content material
 - Administrative forms
- Secure materials
 - Signing out
 - No electronics
- Use of laptops
 - Only use sites you are directed to
 - Do not log out

The Standard Setting Process

Content Standards vs. Performance Standards

- Content standards = “What”
 - Describe the knowledge and skills students are expected to demonstrate by content area and grade
- Performance standards = “How well”
 - Describe attributes of student performance based on Performance Level descriptors

What is Your Job?

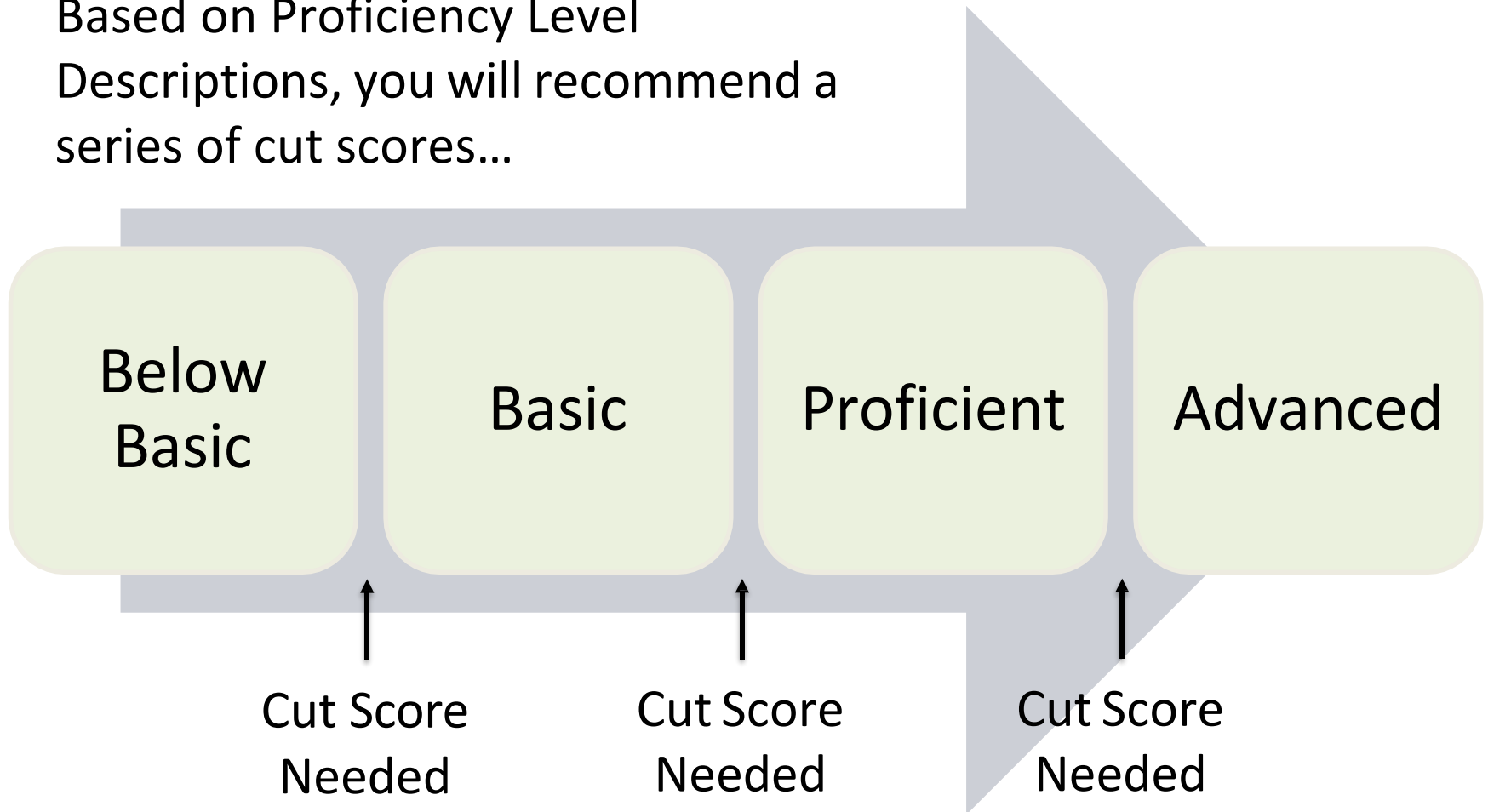
- To recommend cut scores for each of the performance levels that will be used to report results:
 - Below Basic
 - Basic
 - Proficient
 - Advanced
- 
- The diagram consists of three horizontal arrows pointing to the left, each aligned with a performance level. The first arrow is positioned between 'Below Basic' and 'Basic', the second between 'Basic' and 'Proficient', and the third between 'Proficient' and 'Advanced'. Each arrow is labeled 'Cut Score' at its right end.
- | | | |
|-------------|---|-----------|
| Below Basic | ← | Cut Score |
| Basic | ← | Cut Score |
| Proficient | ← | Cut Score |
| Advanced | | |

What are we Trying to Determine?

- What knowledge, skills, and abilities (KSAs) need to be demonstrated to be classified in each Performance Level?
- How much is enough?
- What test performance corresponds to:
 - Below Basic
 - Basic
 - Proficient
 - Advanced

Performance Continuum

Based on Proficiency Level
Descriptions, you will recommend a
series of cut scores...



General Phases of Standard Setting

- Data Collection
 - Your recommendations will be reviewed and presented to the policy makers responsible for final adoption of the cut scores.
- Policy/Decision Making
 - The recommendations may be accepted, rejected, or modified by the Commission for Educational Quality and Accountability (CEQA).

Overview of Standard Setting Method

- We will cover implementation of the Bookmark standard setting procedure
- This session is intended to be an overview
- Your facilitator will give you more details and guide you through the process step by step.

Factors that Influence Selection of Standard Setting Method

- Prior usage and history
- Recommendation or requirement by policy making authority
- Type of assessment(s)
 -
 -
 -
- Bookmark method chosen

What is the Bookmark Method and How Does It Work?

- A collection of test items is arranged in an Ordered Item Booklet (OIB)
 - Based on statistical analysis.
 - From easiest to most difficult.
- Panelists place one or more “bookmarks” in that OIB to recommend cut scores.

Important Terms to Know

- Performance Levels
- Test items
- “Borderline” students
- Knowledge, skills, and abilities (KSAs) needed to answer each test question
- Cut scores

Performance Levels

- Individual review of Performance Level Descriptors (PLDs)
- Group discussion of what student performance in each Performance Level looks like.
- Focus on the “borderline” students, i.e., students who just barely make it into Performance Level.

Develop Borderline Descriptions

- Create bulleted lists of
 - Knowledge, skills, and abilities (KSAs) a student must demonstrate to be classified in each Performance Level, and
 - Knowledge, skills, and abilities that distinguish one Performance Level from another.
- You must reach consensus as a group about the KSAs that define borderline student performance.

How to Place a Bookmark

- Start at the beginning of the OIB.
- Evaluate whether at least two thirds of the students who demonstrate knowledge and skills at the borderline of Proficient would correctly answer the item
- If Yes, move on to the next item.
- Place the bookmark where you think at least two thirds of the Proficient “borderline” students would no longer correctly answer the item.

How to Place a Bookmark

Item Number	Would at least two-thirds of borderline Proficient students correctly answer this item?
1	Yes
2	Yes
3	Yes
4	Yes
5	Yes
6	No
7	Yes
8	Yes
9	No
10	No
11	No
12	No
13	No
14	No
15	No
...	No

How to Place a Bookmark

- You will have opportunities to discuss your bookmark placements and change them, if desired.
- Place one bookmark for each cut score (between the Performance Levels).

Before You Place the Bookmarks

- Take the test to familiarize yourself with the test taking experience.
- Review the OIB.
- Use the item map form to identify KSAs specific to each item.

Item Map

Item Order	What knowledge and skills does this item measure?	Why is this item more difficult than the preceding item?
1		
2		
3		
4		
5		
6		
7		
8		

- Review and discuss Performance Levels.
- Develop definition of “borderline” for Below Basic, Proficient, and Advanced.

The Practice Round

- Before placing actual bookmarks, you will have an opportunity to practice the method with a set of practice items.
- You will be given an OIB with approximately 10 items to practice the bookmark placement for the cut point between Basic and Proficient.

Check for Understanding

- Your facilitator will check with you for understanding and answer any questions you may have during and after the practice round.
- You will then complete a training evaluation form which serves as readiness check before proceeding.

Domain-Specific Bookmark Placement

- Round 1 (Without Discussion)
 - Work through the ordered item booklet.
 - Place bookmarks between the items as appropriate.
- Round 2 (With Group Discussion)
 - Discuss the first-round bookmark placements (focus on the KSAs).
 - Examine your cut points in relation to the group results.
 - Review and revise placement of bookmarks as appropriate.

Overall Science Bookmark Placement

- Round 3 (With Group Discussion)
 - Discuss the second-round bookmark placements (focus on the KSAs).
 - Examine your cut points in relation to the group results and impact data.
 - Review and revise placement of bookmarks as appropriate.
- Round 4 (With Group Discussion)
 - Discuss the third bookmark placements (focus on the impact data).
 - Examine your cut points in relation to the group results and impact data.
 - Review and revise placement of bookmarks as appropriate.

External Assessment Data

- Information from prior OSTP assessments in grades 3-8 included as a validity check
- A region of the item map is shaded that corresponds to projected proficiency percentages, with a range of +/- 2SEMs around that point.
- Within this region is where the Proficient bookmark should be placed.
- Your facilitator will give additional training and guidance on the usage of this data.

External Assessment Data

Example Item Map with Shading

Item Order	What knowledge and skills does this item measure?	Why is this item more difficult than the preceding item?
18		
19		
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		
33		
34		

Role of the Facilitator

- Lead and keep the group on track.
- Ensure that all panelists clearly understand the procedures.
- Ensure that the evaluation forms are completed.
- Your honest feedback is important!

A Few Reminders

- It is not necessary for panelists to reach consensus as to how the items should be assigned to Performance Levels.
- You may or may not change your mind as a result of the discussions.
- Process is focused solely on recommending cut scores.
- The Performance Levels and their definitions are not open for debate.
- Items are operational and fixed.
- Panelists' recommendations are vital, but final cut score decisions will be made by the Commission of Educational Quality and Accountability (CEQA).

Each Panelist Must

- Use his or her own best judgment in each round of rating.
- Be open-minded when listening to your colleagues' rationales for their ratings.
- Complete an evaluation form at the end of the process.
- Participate in the entire process or his/her judgments will be discounted.
- Use cell phones only during breaks.
- Arrive on time after breaks and each day.

What's Next?

- Take the Test
- Domain-Specific Work
 - Complete Item Map Form
 - Discuss the Performance Levels
 - Practice, Rounds 1 & 2
- Overall Science Work
 - Rounds 3 & 4
- Final Evaluation

Any Questions?

Thank you.

APPENDIX B—INSTRUCTIONS FOR FACILITATORS

GENERAL INSTRUCTIONS FOR STANDARD SETTING GROUP FACILITATORS

CCRA Science
June 5-6, 2019

Single-Group Activity

General Orientation

The Standard Setting activities begin with all panelists in one large group, facilitated by the lead facilitator.

Take the Test

Overview: In order to establish an understanding of the test items and for panelists to gain an understanding of the experience of the students who take the test, each participant will take the test. Panelists may wish to discuss or take issue with the items in the test. Tell them we will gladly take their feedback to the SDE. However, this is the actual assessment that students took, and it is the set of items on which we must set standards.

Activities:

- 1) Introduce the assessment and convey/do each of the following:
 - a. Tell panelists that they are about to take an actual OSTP assessment.
 - b. The purpose of the exercise is to help them establish a good understanding of the test items and to gain an understanding of the experience of the students who take the assessment.
- 2) Distribute a computer to each panelist
- 3) Ensure each panelist is able to login to the eMetric Portal and begin the assessment
- 4) Tell panelists to try to take on the perspective of a student as they complete the test. The expectation is that they will spend no more than 30 minutes on this task.
- 5) When the majority of the panelists have finished, pass out the answer key/scoring rubrics.

Domain-Specific Panels: Preparation

Split into Smaller Panels

Overview: After the general orientation, panelists will convene into two smaller standard setting panels according to domain (Life Sciences or Physical Sciences) for which they will be setting standards. Domain-specific standard setting activities will first occur, allowing for close consideration of the distinct content within CCRA Science. These panels will reconvene at a later point in the meeting in order to set a single set of cut-points.

Preliminaries

1. Welcome group, introduce yourself (name, affiliation, a little selected background information).
2. Have each participant introduce him/herself.
3. Ask each participant to sign a nondisclosure form. Do not proceed until a signed nondisclosure form has been collected from each participant.
4. Note that while panelists are making their recommendation for the cut scores, the Commission for Education Quality and Accountability make the final cut decision. The decision is almost always within a range around the recommended cut.

Fill Out Item Map Form

Overview: The primary purpose of this activity is for panelists to think about what knowledge, skills and abilities (KSAs) are measured by each item as well as what makes one question harder or easier than another. The notes panelists take here will be useful in helping them place their bookmarks and in discussions during the rounds of ratings.

On the item map form there is a shaded region based on projections derived from previous assessment. This is the region panelists should consider for the placement of the Proficient bookmark. The shaded region corresponds to a projection based on expected proficiency with a range of ± 2 SEMs around that point.

Activities:

1. Prepare the materials
 - a. Ensure each panelist can open and view item map form (computer)
 - b. Distribute the domain-specific ordered item book
2. Review the domain-specific ordered item book and item map form (computer) with the panelists. Explain what each is, and point out the correspondence of the ordered items between the two. Explain that the items are statistically ordered from easiest to hardest, based on student performance from the most recent administration of the assessment.
3. Tell panelists that the shaded region represents a projection or expectation based on other assessment information, including prior-grade assessment results. During the actual standard setting activity, the Proficient bookmark placement should be set within this range. This information is not critical for the current activity.

4. Tell panelists they will work individually at first. After they have completed the item map form, they will then discuss it as a group.
5. Starting with the first item, they will record for each item:
 - a. The knowledge, skills and abilities (KSAs) the item measures, and
 - b. their thoughts about what makes that question harder than the previous question.
6. Panelists should not agonize over these decisions. It may be that the second item is only slightly harder than the first. Panelists should keep in mind that the purpose of the task is to record notes that will be useful to them in completing their ratings and not necessarily to fill in every space on the form.
7. Once panelists have completed the item map form, they should discuss them as a group.
8. Based on the group discussion, the panelists may modify their own item map form (make additional notes, cross things out, etc.)

Discuss Performance Level Definitions and Describe Characteristics of the “Borderline” Student

Overview: In order to establish an understanding of the expected performance of borderline students on the test, panelists must have a clear understanding of:

- 1) Specific interpretation of the performance levels within their current domain (Life Sciences or Physical Sciences), and
- 2) Characteristics of students who are “just able enough” to be classified into each level above Below Basic within a specific domain. These students will be referred to as borderline students, since they are right on the border between levels.

The purpose of this activity is for the panelists to obtain an understanding of the domain-specific Performance Level Definitions with an emphasis on characteristics that describe students at the borderline within a specific domain -- both what these students can and cannot do.

This activity is critical since the ratings panelists will be making will be based on these understandings.

Preparation:

1. Use 3 sheets of chart paper and label the top of each one: Borderline Basic, Borderline Proficient and Borderline Advanced.

Activities:

- 1) Introduce the task. In this activity they will:
 - a. individually review the domain-specific Performance Level Descriptors again as needed;

- b. generate group descriptions of borderline Basic, Proficient and Advanced students.

The facilitator should compile the descriptions as bulleted lists on chart paper; the chart paper will then be posted so the panelists can refer to the lists as they go through the bookmark process.

- 2) Check to see if panelists want to discuss the performance levels again. Once they have a solid understanding of the PLDs, have them focus their discussion on the knowledge, skills, and abilities of students who are in the Proficient category, but just barely. The focus should be on those characteristics and KSAs that best describe the lowest level of performance necessary to warrant Proficient classification.
- 3) After discussing Proficient, have the panelists discuss characteristics of the borderline Basic student and then characteristics of the borderline Advanced student. Panelists should be made aware of the importance of the Proficient cut. This is the cut from non-proficient to just barely proficient.
- 4) Using chart paper, generate a bulleted list of characteristics for each of the levels. Post these on the wall of the room. Make sure that panelists agree on the bulleted characteristics and have a common understanding.

Practice Round

Overview of Practice Round: The primary purpose of the Practice Round is for panelists to become familiar with the task of placing the bookmarks. The facilitator will walk the panelists through the Proficient bookmark placement on the practice set, engage the panelists in a readiness discussion and check for understanding. If any of the panelists indicate an incomplete understanding of the practice rating task, then the facilitator will continue to work with the panelists to clarify any misconceptions before proceeding to Round 1.

Activities:

1. Make sure panelists have the following materials:
 - a. Domain-specific practice ordered item set
 - b. Domain-specific Performance Level Definitions
 - c. Access to the domain-specific practice rating form (computer)
2. Orient panelists to the domain-specific practice ordered item set. Point out the following:
 - a. Only items from the current domain are included in the item set;
 - b. Items are organized by difficulty from easiest to hardest.
 - c. The items represent the full range of difficulty included on the test.
 - d. Identify the items on the item map form that correspond to the practice ordered item set. Panelists can see their item map form entries on the practice rating form.
 - e. Show panelists how to indicate their bookmark placement on the practice rating form (computer).

3. Give the panelists a few minutes to read through the items.
4. The facilitator leads the group through a discussion of the Proficient bookmark placement in the domain-specific practice OIB.
 - a. Referring to the ten ordered items in the practice set, the domain-specific Performance Level Definitions, and the bulleted lists of domain-specific borderline characteristics posted on chart paper, the facilitator will lead a discussion about the placement of the Proficient bookmark.
 - b. Panelists should consider the question: **would at least two-thirds of the students performing at the borderline of Proficient answer the item correctly?**
 - c. Where the answer changes from yes to no is where the bookmark should be placed.
 - d. Panelists should answer the above question for all items to check for anomalies.
 - e. Panelists should enter their bookmark placement on the practice rating form (computer)
 - f. Use the practice rating master sheet to show where each panelist placed their bookmark. Have a discussion of their ratings in the context of the ratings made by other members of their group. Ask the panelists to discuss the rationale for placement of the highest and lowest ratings. The group should get a sense of how much variation there is in the ratings.

Readiness Discussion

After the panelists have placed bookmarks in the domain-specific practice ordered item set, lead a readiness discussion by posing the following seven questions.

The purpose of this discussion is to determine how well each panelist understands the bookmark task, to correct any misunderstandings, and if necessary, to identify panelists whose ratings should be excluded from the standard setting if their understanding doesn't improve.

The "correct" answer for each of the question is listed directly under each question. Some common misunderstandings are also listed for questions one and two. Please watch for these typical misunderstandings and if they arise, redirect the panelists to the correct responses.

Make sure any questions or concerns are resolved prior to moving on.

1. What questions should you ask for each item?
 - Would at least two-thirds of the borderline students get this item correct?
 - Would at least two-thirds of the students who just barely fall in the criteria level of interest get this item correct?

Please watch for and correct the following misconceptions.

- Omission of two-thirds (stating all students is also incorrect)
- Omission of borderline (stating all students, or all students in the criteria level of interest is also incorrect)

2. What is meant by the “at least two-thirds” rule?
 - At least two-thirds of the borderline students would get items like this correct

Please watch for and correct the following misconceptions.

 - All students falling in the performance level of interest have a one out of two chance of getting this item correct.
3. What population of students should you consider for each item?
 - Borderline students
 - Students who just barely fall in the performance level of interest
 - a. Does the target population of borderline students change as I progress through the items for the first bookmark? (NO)
 - b. Does the target population change as I progress to the next bookmark? (YES)
4. As you approach a bookmark, how do answers change?
 - The answer to “Would at least two-thirds of the borderline students get this item correct” should change from a “yes” to a “no”
5. How should your confidence in the answers affect your bookmark placement?
 - As you become less confident in a “yes” answer, the bookmark placement should be approaching.
 - Where you are least confident in your “yes” answer, suggesting a “no”, is typically where the bookmark will be placed.
6. Does placing a bookmark after a certain page mean the student needs to get that many items correct on the assessment?
 - NO. The OIB page number is only an ordered index, and does not correspond to the number correct.
7. Should the population you are thinking about be the students in your classroom or school?
 - NO. You should be thinking about all of the students in the state.

NOTE: Make sure you collect all of the ‘training’ OIBs!

Standard Setting Practice Evaluation

After the panelists have placed bookmarks in the domain-specific practice ordered item set and you’ve completed the readiness discussion and answered any questions, have panelists fill out the training evaluation form. Before you start the Round 1 activities, scan the completed evaluations to see if there are any problems or concerns that need to be addressed before proceeding. **Make sure any questions or concerns are resolved prior to moving on.** Return the completed evaluations to the data analysis work room at the next convenient opportunity.

Domain-Specific Panels: Standard Setting

Round 1

Overview of Round 1: The primary purpose of Round 1 is to ask the panelists to make their initial judgments as to where the bookmark should be placed for each cut within their domain-specific OIB. For this round, panelists will work individually, without consulting with their colleagues. Beginning with the first ordered item in the domain-specific OIB, panelists will evaluate each item in turn. The panelists will gauge the level of difficulty of each of the items for those students who barely meet the definition of Proficient. The task that panelists are asked to do is to estimate whether a student performing at the borderline of Proficient, would answer each question correctly. More specifically, panelists should answer:

- Would *at least* two-thirds of the students performing at the borderline of Proficient answer the question correctly?

On the item map form there is a shaded region based on projections derived from previous assessments. This is the region panelists should consider for the placement of the Proficient bookmark. The shaded region corresponds to a projection based on expected proficiency with a range of ± 2 SEMs around that point.

The Proficient bookmark placement must be between two shaded items. Should a panelist desire to set the bookmark outside the shaded region they will be asked to provide written justification.

The same process is then repeated for the [Below Basic/Basic] and [Proficient/Advanced] cuts.

Activities:

1. Panelists should have their domain-specific ordered item booklets, and Performance Level Definitions. Instruct the panelists to open the procedural rating form (computer) and show how details from their individual item map descriptions have been carried forward to the rating form. Ensure each panelist is able to open their rating form before proceeding.
2. Have panelists confirm their ID number matches the ID number on their procedural rating form and item map form. The ID number is on the back of their table tent.
3. Provide an overview of Round 1, covering each of the following:
 - a. Orient panelists to the domain-specific ordered-item book. Remind them that the items are presented in order of difficulty, from easiest to hardest, for their current domain only.
4. Remind panelists the shaded region is derived from growth projections, and that the Proficient bookmark placement should be set in this range. Placing the bookmark outside the shaded region will require that the panelist provide brief written justification.

- b. The primary purpose of this activity is for the panelists to make their initial determination as to whether students whose performance is barely Proficient would correctly answer each item, and to place their bookmark where they believe the answer of 'yes' turns to 'no'. Remind panelists that they should be thinking about at least two-thirds of the borderline students. Once they have completed the process for the [Basic/Proficient] cut, they will proceed to the remaining two cut points starting with [Below Basic/Basic] and then the [Proficient/Advanced] cut.
 - c. Each panelist needs to base his/her judgments on his/her experience with the content, understanding of students, and the definitions of the borderline students generated previously.
 - d. One bookmark will be placed for each cut point. For CCRA there are 3 cut points and, therefore, three bookmarks will be placed ["Basic", "Proficient", "Advanced"]. Place the cut point number on the procedural rating form in the RND 1 column.
 - e. If panelists are struggling with placing a particular bookmark, they should use their best judgment and move on. They will have an opportunity to discuss their ratings and make revisions in Rounds 2 and 3.
5. Tell panelists that they will be placing the bookmarks individually; they will have the option to discuss each cut point with the other panelists during Round 2. **It is not necessary for the panelists to come to consensus about where the bookmarks should be placed.**
6. Go over the rating form with panelists.
 - a. Lead panelists through a step-by-step demonstration of how to fill in the rating form.
 - b. Answer questions the panelists may have about the work in Round 1.
 - c. Once everyone understands what they are to do in Round 1, tell them to begin.
7. Starting with the first ordered item in the OIB and proceeding up to their bookmark placement for the [Basic/Proficient] cut point, the panelists will work through the OIB item by item and make their initial bookmark placements. Have panelists continue to examine five items past their placement to check for anomalies.
8. As panelists complete the task, ask them to carefully inspect their rating forms to ensure they are filled out properly.
 - a. The ID number must be filled in.
 - b. Exactly three cuts must be entered and identified "Basic", "Proficient" and "Advanced" on the procedural rating form in the RND 1 column.

- a. The cut points must be entered sequentially on the rating form (e.g., the bookmark for “Proficient” cannot be placed on an easier item in the OIB than the bookmark for “Basic” on the rating sheet).
- b. The “Proficient” bookmark placement should be between two shaded items on the item map form, or a written justification must be provided.
- c. Check each panelist’s rating form before you allow them to leave for a short break.
- d. When all the rating forms have been validated, the group will take a break. Rating information for round 1 will be locked, so it cannot be changed.

Complete Procedural Evaluation Form

Make sure panelists fill out the procedural evaluation for the grade. Emphasize that their honest feedback is important. Return the completed evaluations to the data analysis work room at the next convenient opportunity. Collect the materials from the grade and mark them off on the Materials Tracking sheet.

Tabulation of Round 1 Results

Tabulation of Round 1 results will be completed by the data analysis team as quickly as possible after processing the rating forms.

Round 2

Overview of Round 2: In Round 2, the panelists will discuss their Round 1 placements as a group and then revise their ratings on the basis of that discussion. They will discuss their ratings in the context of the ratings made by other members of their group. Panelists should discuss the rationale for placement of the highest and lowest ratings. The group should get a sense of how much variation there is in the ratings. Panelists should also consider the question, “How tough or easy a rater are you?” The purpose here is to allow panelists to examine their individual expectations (in terms of their experiences) and to share these expectations and experiences in order to attain a better understanding of how their experiences impact their decision-making.

To aid with the discussion, the panelists will be provided with the median Round 1 bookmark placements for their group.

Once panelists have reviewed and discussed their bookmark placements, they will be given the opportunity to change or revise their Round 1 ratings.

Activities:

1. Make sure the panelists have their domain-specific ordered item booklets, item map forms (computer), and Performance Level Definitions. Ensure each panelist is able to open their rating form.
2. A psychometrician will present and explain the following information to the panelists:
 - a. the median bookmark placements for the group based on the Round 2 ratings. This information is provided so panelists can get a sense of where they fall.

relative to the group median –if they are more stringent or more lenient than other panelists.

3. Provide an overview of Round 2. Remind panelists of the following:
 - a. As in Round 1, the primary purpose is to place bookmarks where you feel the criteria levels are best distinguished, considering the additional information and discussion.
 - b. Each panelist needs to base his/her judgments on his/her experience with the content area and specific domain, understanding of students, the definitions of the borderline students generated previously, discussions with other panelists and the knowledge, skills, and abilities (KSAs) required to answer each item.
4. The panelists will discuss their Round 1 ratings as a group, beginning with the Proficient cut point and followed by the Basic and Advanced cuts.
 - a. The discussion should focus on differences in where individual panelists in the group placed their bookmarks.
 - b. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
 - c. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
 - d. On the basis of the discussions, panelists should make a second round of ratings.
 - e. Remind panelists the shaded region is derived from growth projections and that the Proficient bookmark placement will be set in this range. The Proficient bookmark should be between two shaded items.
 - f. When placing their Round 2 bookmarks, panelists should not feel compelled to change their ratings.
 - g. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.

Encourage the panelists to use the discussion and feedback to assess how stringent or lenient a judge they are. If a panelist is consistently higher or lower than the group, they may have a different understanding of the borderline student than the rest of the group, or a different understanding of the Performance Level Definitions, or both. **It is O.K. for panelists to disagree, but that disagreement should be based on a common understanding of the Performance Level Definitions.**

5. As the group is conducting their discussions, circulate around the room to ensure that the discussions are staying on topic, the panelists understand the task, and that all panelists are participating appropriately in the discussion.
6. When all panelists in each group have completed their second ratings, **carefully inspect the rating forms** to ensure they are filled out properly.
 - a. The ID number must be filled in correctly.
 - b. Exactly three cuts must be entered and identified “Basic”, “Proficient” and “Advanced” on the procedural rating form.
 - c. The cut points must be entered sequentially on the rating form (e.g., the bookmark for “Proficient” can’t come before the bookmark for “Basic” on the rating sheet).
 - d. The “Proficient” bookmark placement should be between two shaded items on the item map form. If it is outside the shaded region, a written justification must be provided.
 - e. Check each panelist’s rating form before you allow them to leave for a short
 - f. break.
 - g. When all the rating forms have been validated, the group will take a break. Rating information for round 2 will be locked, so it cannot be changed.

Single-Group Activity: Standard Setting

Round 3

Overview of Round 3: At the conclusion of Round 2 discussions, the complete Science panel will be reassembled from the domain-specific Life Sciences and Physical Sciences panels. Subsequent standard setting activities will be conducted with the entire panel. The primary purpose of Round 3 is to ask the complete Science panel to discuss their Round 2 placements as a group. However, unlike in Round 2, in Round 3 the panelists will have the opportunity to discuss the impact of their domain-specific bookmark placements against overall Science performance and to revise the cut-points based on that discussion. The goal of these discussions is for panelists to resolve the cut-points determined separately by domain, considering whether the percentage of students in each achievement level category seems reasonable.

To aid with the discussion, a psychometrician will present the following information to the panelists:

1. The group median Round 2 bookmark placements for each domain;
2. Impact data, showing the approximate percentage of students statewide that would be classified into each performance level category based on the room median bookmark placements from Round 2 for each domain; and
3. Standard error information, this will demonstrate to the panelists the amount of variability present in the cut scores expressed in real-world terms. Both Median Absolute Deviation (How much disagreement among panelists) and Conditional Standard Error (Measure of error in assessment) data will be provided. A range of impact data for each cut will be determined for ± 1 SE around the cut score for each of these.

Once panelists have reviewed and discussed their bookmark placements and the impact data, they will be given the opportunity to change or revise their Round 2 ratings.

Activities:

1. Make sure the panelists have their complete Science ordered item booklets, item map forms (computer), and Performance Level Definitions. Ensure each panelist is able to open and access their Round 3 and 4 procedural rating form.
 - a. The rating form for Rounds 3 and 4 (computer) is a different worksheet than for Rounds 1 and 2.
 - b. The rating form continues to include the shaded region for guiding placement of the Proficient bookmark and includes colored regions for the range of domain-specific bookmark placements. Yellow indicates the range of Basic bookmarks, green indicates the range of Proficient bookmarks, and blue indicates the range of Advanced bookmarks. For example, a yellow region indicates the Life Sciences bookmark placement, the Physical Sciences bookmark placement, and any pages or items that are between the two.

2. A psychometrician will present and explain the following information to the panelists:
 - a. the median bookmark placements for the group based on the Round 2 ratings. Based on their Round 2 rating form, panelists will know where they fall relative to the group median. This information is provided so panelists can get a sense if they are more stringent or more lenient than other panelists.
 - b. Impact data, showing the approximate percentage of students statewide that would be classified into each performance level category based on the room median bookmark placements for each domain. Panelists will use this information as a “reasonableness check.” In other words, they will discuss whether the percentages in each level seem reasonable, based on their knowledge of the test and the current status of students across the state relative to the Performance Level Definitions. If the answer is no, panelists may choose to make adjustments to one or more of their bookmark placements. Panelists may decide to select bookmarks resulting from either domain or select an entirely new bookmark between the domain-specific bookmarks. To facilitate these discussions and decisions, the panelists will be provided with an overall Science OIB which will include both Life and Physical Science items as administered on the core operational form. To facilitate the identification of an appropriate bookmark, panelists will be instructed to consider only those items in the OIB that appear between the domain-specific bookmarks.
 - c. Standard error information, this will demonstrate to the panelists the amount of variability present in the cut scores expressed in real-world terms. Both Median Absolute Deviation (How much disagreement among panelists) and Conditional Standard Error (Measure of error in assessment) data will be provided. A range of impact data for each cut will be determined for ± 1 SE around the cut score for each of these.
3. Provide an overview of Round 3. Remind panelists of the following:
 - a. As in Round 2, the primary purpose is to place bookmarks where you feel the performance levels are best distinguished, considering the additional information and further discussion.
 - b. Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definitions of the borderline students generated previously, discussions with other panelists, the knowledge, skills, and abilities required to answer each item, and the consensus and impact data.
 - c. The panelists will discuss their domain-specific ratings, beginning with the Proficient cut point and followed by the Basic and Advanced cuts.
 - d. The discussion should focus on differences in where individual panelists placed their bookmarks.

- e. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
 - f. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
 - g. On the basis of the discussions, panelists should make a third round of ratings.
 - h. Remind panelists the shaded region is derived from growth projections and that the Proficient bookmark placement will be set in this range. The Proficient bookmark must be between two shaded items or a written justification must be provided by the panelist.
 - i. Remind panelists additionally that the yellow, green, and blue shaded regions indicate the domain-specific bookmark placements. The complete Science bookmarks for Basic, Proficient, and Advanced should be placed within those ranges, or a written justification must be provided.
 - j. Because of the combination of domain-specific OIBs and the need to make a judgement about overall Science performance, it is likely that panelists will change their bookmark placement from the previous round.
 - k. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.
 - l. Write brief notes on any notable discussions of the process, any particular sticking points or issues, or key rationales had in their judgments. These do not need to be formal but will be useful if the client has questions regarding the process.
4. When all panelists have completed their second ratings, **carefully inspect the rating forms (computer)** to ensure they are filled out properly.
- a. The ID number must be filled in correctly.
 - b. Exactly three cuts must be entered and identified “Basic”, “Proficient” and “Advanced” on the procedural rating form.
 - c. The cut points must be entered sequentially on the rating form (e.g., the bookmark for “Proficient” can’t come before the bookmark for “Basic” on the rating sheet).
 - d. The “Proficient” bookmark placement should be between two shaded items on the item map form. If it is outside the shaded region, a written justification must be provided.

- g. Check each panelist's rating form before you allow them to leave for a short
- h. break.
- i. When all the rating forms have been validated, the group will take a break. Rating information for round 3 will be locked, so it cannot be changed.

Round 4

Overview of Round 4: The primary purpose of Round 4 is to ask the panelists to discuss their Round 3 placements as a group and to give them one last opportunity to revise their ratings on the basis of that discussion. As in Round 3, they will discuss their ratings in the context of the ratings made by other members of the group.

To aid with the discussion, a psychometrician will present the following information to the panelists:

1. The group median Round 3 bookmark placements for Science overall, and
2. Impact data, showing the approximate percentage of students statewide that would be classified into each performance level category based on the room median bookmark placements from Round 3 for Science overall.
3. Standard error information, as before.

Once panelists have reviewed and discussed their bookmark placements and the impact data, they will be given the opportunity to change or revise their Round 3 ratings.

Activities:

1. Make sure the panelists have their ordered item booklets, item map forms (computer), and Performance Level Definitions. Ensure each panelist can open their procedural rating form.
2. A psychometrician will present and explain the following information to the panelists:
 - a. the median bookmark placements for the group based on the Round 3 ratings. Based on their Round 3 rating form, panelists will know where they fall relative to the group median. This information is provided so panelists can get a sense if they are more stringent or more lenient than other panelists.
 - b. Impact data, showing the approximate percentage of students statewide that would be classified into each performance level category based on the room median bookmark placements. Panelists will use this information as a "reasonableness check." In other words, they will discuss whether the percentages in each level seem reasonable, based on their knowledge of the test and the current status of students across the state relative to the Performance Level Definitions. If the answer is no, panelists may choose to make adjustments to one or more of their bookmark placements.

Standard error information, this will demonstrate to the panelists the amount of variability present in the cut scores expressed in real-world terms. Both Median Absolute Deviation (How much disagreement among panelists) and Conditional Standard Error (Measure of error in assessment) data will be provided. A range of impact data for each cut will be determined for ± 1 SE around the cut score for each of these.

3. Provide an overview of Round 4. Remind panelists of the following:
 - a. As in Round 3, the primary purpose is to place bookmarks where you feel the performance levels are best distinguished, considering the additional information and further discussion.
 - b. Each panelist needs to base his/her judgments on his/her experience with the content area, understanding of students, the definitions of the borderline students generated previously, discussions with other panelists and the knowledge, skills, and abilities required to answer each item.
 - c. The panelists will discuss their Round 3 ratings, beginning with the Proficient cut point and followed by the Basic and Advanced cuts.
 - d. The discussion should focus on differences in where individual panelists placed their bookmarks.
 - e. Panelists should be encouraged to listen to their colleagues as well as express their own points of view.
 - f. If the panelists hear a logic/rationale/argument that they did not consider and that they feel is compelling, then they may adjust their ratings to incorporate that information.
 - g. On the basis of the discussions, panelists should make a fourth round of ratings.
 - h. Remind panelists that the shaded regions for Proficient, Basic, and Advanced should guide placement of their bookmarks. Placement outside these ranges will require brief written justification.
 - i. When placing their Round 4 bookmarks, panelists should not feel compelled to change their ratings.
 - j. The group does not have to achieve consensus. If panelists honestly disagree, that is fine. We are trying to get the best judgment of each panelist. Panelists should not feel compelled or coerced into making a rating they disagree with.
 - k. Write brief notes on any notable discussions of the process, any particular sticking points or issues, or key rationales had in their judgments. These do not

need to formal but will be useful if the client has questions regarding the process.

4. When all panelists have completed their fourth ratings, **carefully inspect the rating forms (computer)** to ensure they are filled out properly.
 - a. The ID number must be filled in.
 - b. Exactly three cuts must be entered and identified “Basic”, “Proficient” and “Advanced” on the procedural rating form.
 - c. The cut points must be entered sequentially on the rating form (e.g., the
 - d. bookmark for “Proficient” can’t come before the bookmark for “Basic” on the
 - e. rating sheet).
 - f. The “Proficient” bookmark placement should be between two shaded items on the item map form. If it is outside the shaded region, a written justification must be provided.
 - g. The standard setting team will now lock the round 4 ratings.

Complete Final Evaluation Forms

Make sure panelists fill out the final evaluations before they leave. Emphasize that their honest feedback is important.

APPENDIX C—PANELISTS

Panelists

Table C-1. 2019 OK Standard Setting Report: Science Panelists

<i>Full Name</i>	<i>Email Address</i>	<i>Company</i>	<i>Invitation List</i>	<i>Status</i>
Peters, Chanda	cpeters@woodwardps.net	Woodward	Physical Science	Accepted
Wright, Gayla	docgayla@cox.net	OERB	Physical Science	Accepted
Jones, Vanessa (cancelled)	jonesv@bethel.k12.ok.us	Bethel High School	Life Science	Accepted
Chaisson, Leiha	lchaisson1@cox.net	Mustang	Life Science	Accepted
Will, Tammy	tammywill@morrisonps.com	Morrison Public School	Physical Science	Accepted
Tamez, Jeramey	Jeramey.Tamez@yukonps.com	Yukon	Life Science	Accepted
Zumwalt, Ruth	ruth.zumwalt@edmondschools.net	Edmond Public Schools	Physical Science	Accepted
Richardson, Traci	trichardson@stillwaterschools.com	Currently Stillwater, but that will change	Life Science	Accepted
Schweitzer, Dawna	schweitzer.dawna@gmail.com	Retired	Life Science	Accepted
Shrauner, Jennifer	jshrauner@putnamcityschools.org	Putnam City	Life Science	Accepted
Gilmore, Paul	pgilmore@putnamcityschools.org	Putnam City	Physical Science	Accepted
Maier, Steve	sjmaier@nwosu.edu	Alva	Physical Science	Accepted

APPENDIX D—PERFORMANCE LEVEL DESCRIPTORS

Oklahoma School Testing Program
Performance-Level Descriptors
Grade 7 Geography: Eastern Hemisphere

ADVANCED: Students demonstrate superior understanding of challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level will

- infer and apply information using a variety of geographic sources
- analyze the importance of Celebrate Freedom Week
- compare and contrast cultural, physical, and political regions; urban areas and countries
- analyze how human and physical characteristics affect regions over time
- evaluate the role of international organizations in conflict and cooperation
- identify and describe major landforms and bodies of water
- identify the causes of natural disasters and analyze their effects on human populations and the environment
- summarize and evaluate how countries/regions are categorized based on cultures, population locations, economic development, social and political structures, and standard of living measures
- analyze and predict the distribution of natural resources and the three sectors of the economy
- analyze how humans adapt to and change the natural environment
- evaluate governmental policies that address regional resource issues

PROFICIENT: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade. Students scoring at the Proficient level will

- interpret information using a variety of geographic sources
- explain the importance of Celebrate Freedom Week
- identify and describe cultural, physical, and political regions; urban areas and countries
- explain how human and physical characteristics affect regions over time

- describe the role of international organizations in conflict and cooperation
- identify and describe major landforms and bodies of water
- identify the causes of natural disasters and explain their effects on human populations and the environment
- compare and contrast how countries/regions are categorized based on cultures, population locations, economic development, social and political structures, and standard of living measures
- identify and describe the distribution of natural resources and the three sectors of the economy
- explain how humans adapt to and change the natural environment
- describe governmental policies that address regional resource issues

LIMITED KNOWLEDGE: Students demonstrate partial mastery of the essential grade-level knowledge and skills. Students at the Limited Knowledge level will

- identify some information using a variety of geographic sources
- identify the importance of Celebrate Freedom Week
- identify or describe some of the cultural, physical, and political regions; urban areas and countries
- identify how some human and physical characteristics affect regions over time
- identify the involvement of some international organizations in conflict and cooperation
- identify and locate some major landforms and bodies of water
- identify some of the causes of natural disasters and explain some of their effects on human populations and the environment
- compare or contrast how some countries/regions are categorized based on cultures, population locations, economic development, social and political structures, and standard of living measures
- identify or describe the distribution of some natural resources and some sectors of the economy
- identify some ways humans adapt to and change the natural environment
- identify or describe some governmental policies that address regional resource issues

UNSATISFACTORY: Students have not performed at least at the Limited Knowledge level. Students at the Unsatisfactory level have not demonstrated grade-level knowledge and skills.

Grade 3 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically complete complex addition, subtraction, and multiplication problems and model division facts. Students order fractions using models and compose and decompose fractions related to the same whole. Students extend patterns and generate real-world situations to represent number sentences. Students determine volume and elapsed time. Students summarize complex data sets and analyze the data to solve problems. Students solve complex and non-routine real-world problems, draw logical conclusions, and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically compare and order whole numbers. Students complete addition, subtraction, and multiplication problems and recognize the relationship between multiplication and division. Students construct and compare fractions using models. Students select the fewest number of coins for a given amount of money. Students determine rules to describe basic patterns. Students determine unknowns in equations and apply number properties. Students classify angles. Students sort three-dimensional figures and determine the perimeter of polygons. Students determine the area of two-dimensional figures. Students read and analyze length, temperature, and time. Students summarize a data set and analyze the data to solve problems. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level represent whole numbers. Students complete simple addition, subtraction, and multiplication problems. Students read and write fractions. Students determine the value of a set of coins or bills. Students determine rules to describe simple patterns. Students determine unknowns in simple equations. Students identify right angles. Students choose an appropriate instrument to measure an object. Students read and write time from a digital clock.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 3 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level consistently choose the best summary of the text and identify the main idea and key details. Students compare and contrast details in literary and nonfiction/informational texts to describe genres. Students frequently identify literary elements, literary devices, and author's purpose and frequently distinguish fact from opinion. Students consistently infer whether a text is written in first or third person point of view.

Students consistently engage in a recursive writing process to create organized written works with a purpose that is clearly communicated for an appropriate audience. Students skillfully use details that support the writing task.

Students skillfully use vocabulary knowledge and resources to analyze complex text through word parts, word relationships, and context clues. Students consistently use appropriate and meaningful vocabulary to enhance clarity and effectiveness in their writing.

Students consistently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students generate a question on a specific topic and consistently locate and use information, including graphic features, to understand the text. Students determine the relevance and reliability of information. Students clearly summarize and present information in an organized and cohesive way.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically choose the best summary of the text and identify the main idea and key details. Students compare and contrast details to classify genres. Students identify literary elements, literary devices, and author's purpose and distinguish fact from opinion. Students infer whether a text is written in first or third person point of view.

Students engage in a recursive writing process to create organized written works. Students create written works for specific purposes and audiences using details that support the writing task.

Students use vocabulary knowledge and resources to interpret text through word parts, word relationships, and context clues. Students use appropriate vocabulary to write clearly and effectively.

Students frequently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students generate a question on a specific topic and locate and use information, including graphic features, to understand the text. Students summarize and present information in an organized way.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level inconsistently choose the best summary of the text and have difficulty identifying main ideas and key details. Students compare and contrast but inconsistently classify genres. Students inconsistently identify literary elements, literary devices, author's purpose, or points of view or inconsistently distinguish fact from opinion.

Students inconsistently engage in a recursive writing process to create written works that lack organization. Students write for a specific purpose but seldom consider the audience. Students inconsistently support their ideas with details.

Students inconsistently use vocabulary knowledge and resources to interpret text through word parts, word relationships, or context clues. Students inconsistently use appropriate vocabulary in written works.

Students inconsistently identify and apply appropriate use of grammar and mechanics.

Students generate a question on a topic but ineffectively locate and use information, or imprecisely use graphic features, to understand the text. Students provide an incomplete summary and present information with lack of clarity.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.

Grade 4 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically estimate and solve complex mathematical problems and determine the unknown in non-equivalent expressions. Students compare decimals and fractions. Students solve complex money problems. Students determine a rule and extend a complex pattern. Students determine and represent unknown values in complex problems. Students determine volume. Students solve complex measurement problems. Students represent complex data sets and solve problems involving the data. Students solve complex and non-routine real-world problems, draw logical conclusions, and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically estimate and solve mathematical problems. Students use models to determine equivalent fractions, compare and order fractions, and add and subtract fractions. Students read and write decimals and make connections between decimals and fractions. Students determine change using coins. Students determine rules and extend patterns. Students determine unknown values in mathematical problems. Students describe parts of geometrical figures and identify similarities in three-dimensional figures. Students decompose and determine the area of polygons. Students solve measurement problems. Students represent data sets and solve problems involving the data. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level demonstrate the ability to estimate and solve simple mathematical problems. Students use models to determine simple equivalent fractions, compare and order whole numbers and simple fractions, and decompose fractions. Students read and write simple decimals and compare and order whole numbers and decimals. Students determine change using whole dollars. Students determine a rule and extend a simple pattern. Students determine unknown values in simple mathematical problems. Students identify quadrilaterals and determine the area of simple polygons. Students identify appropriate units and tools to measure. Students solve simple problems given a data set.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 4 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level consistently choose the best summary of the text and explain how the details support the main idea. Students compare and contrast details in literary and nonfiction/informational texts to describe and analyze genres. Students consistently recognize the paraphrase of original text. Students consistently identify and describe literary elements, literary devices, author's purpose, accuracy of facts, and text structure in various texts. Students consistently infer meaning from increasingly complex text including author's purpose and points of view.

Students consistently engage in a recursive writing process to create purposeful and organized written works. Students create fully developed and engaging written works for specific purposes and audiences using details that support the writing task.

Students efficiently use vocabulary knowledge and resources to analyze complex text through word parts, word relationships, and context clues. Students consistently use appropriate and meaningful vocabulary to enhance clarity and effectiveness in their writing.

Students consistently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students generate a viable research question on a specific topic and consistently locate and use information, including graphic features, to interpret the text. Students organize and synthesize relevant and reliable information in order to present findings.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically choose the best summary of the text and identify the details that support the main idea. Students compare and contrast details in literary and nonfiction/informational texts to classify genres. Students recognize the paraphrase of original text most of the time. Students identify and describe literary elements, literary devices, author's purpose, accuracy of facts, and text structure in various texts. Students infer meaning from a text including author's purpose and points of view.

Students engage in a recursive writing process to create purposeful written works. Students select and apply the organizational structure that best fits the mode, purpose, and audience.

Students use vocabulary knowledge and resources to interpret text through word parts, word relationships, and context clues. Students use appropriate vocabulary to write clearly and effectively.

Students frequently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students generate a viable research question on a specific topic and adequately locate and use information, including graphic features, to interpret the text. Students organize relevant and reliable information in order to present findings.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level inconsistently choose the best summary of the text and have difficulty differentiating main ideas from details. Students compare and contrast details in literary and nonfiction/informational texts but inconsistently classify genres. Students seldom identify the paraphrase of original text. Students inconsistently identify and describe literary elements, literary devices, author's purpose, points of view, or accuracy of fact.

Students inconsistently engage in a recursive writing process to create written works. Students' writing lacks organizational structure. Students create underdeveloped written works for specific purposes and audiences with inconsistent use of details.

Students inconsistently use vocabulary knowledge and resources to interpret text through word parts, word relationships, or context clues. Students inconsistently use appropriate vocabulary in written works.

Students inconsistently identify and apply appropriate use of grammar and mechanics.

Students generate a research question on a topic but ineffectively locate and use information, or imprecisely use graphic features, to interpret the text.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.

Grade 5 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically interpret the remainder of division problems within the context of the problem. Students order decimals, fractions, and whole numbers. Students evaluate complex expressions, equations, and inequalities. Students construct geometric figures and identify them in various contexts. Students compare the volume, perimeter, or surface area of geometric figures. Students analyze complex graphs. Students solve complex and non-routine real-world problems, draw logical conclusions, and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically estimate and solve division problems with the remainder represented as a fraction or decimal. Students generate equivalent decimals and fractions, represent whole numbers or decimals, and compare fractions and decimals, including mixed numbers. Students estimate, add, and subtract decimals and fractions. Students describe patterns of change and graph these patterns as ordered pairs on a coordinate plane. Students evaluate expressions, equations, and inequalities. Students solve volume and perimeter problems and simple surface area problems. Students determine reasonable values for the perimeter of shapes with curves. Students compare angles. Students recognize relationships within a measurement system. Students determine the mean, median, mode, and range of a data set and analyze simple graphs. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level estimate and solve division problems with remainders and solve addition and subtraction real-world problems. Students recognize basic equivalent decimals and fractions, represent whole numbers, and compare and order fractions or decimals. Students add and subtract decimals and fractions with like denominators. Students describe simple patterns of change and identify ordered pairs on a coordinate plane. Students evaluate simple equivalent numerical expressions or equations. Students describe and classify geometric figures. Students solve simple volume and perimeter problems. Students choose an appropriate instrument to measure objects and read and analyze the length of objects. Students read and analyze the measure of angles. Students read simple graphs.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 5 Science Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically analyze scale, proportion, quantity and patterns when performing computational thinking to complex data as it pertains to distribution of water on Earth, conservation of matter, and Earth's relationship with the sun, moon and stars. Students predict, modify, and extend complex models at various scales to analyze the movement of matter and energy between organisms, ecosystems, and Earth's systems, and analyze the outcomes of these interactions. Students analyze and compare evidence, data, and models to engage in argument to explain the cause and effect relationships between an object and Earth's gravity, how scale and proportion affect the apparent brightness of the sun and other stars/ and/or how plants use matter (chiefly air and water) to grow. Students observe and measure phenomenon to interpret and evaluate patterns that classify materials based on properties. Students can describe complex cause and effect relationships when mixing substances within an investigation framework.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically describe, use and/or develop basic models at various scales to explain the movement of matter and energy between organisms, ecosystems, and Earth's systems and explain the outcomes of these interactions. Students apply scale, proportion, quantity, and/or patterns when performing computational thinking to data as it pertains to distribution of water on Earth, conservation of matter, and Earth's relationship with the sun, moon, and stars. Students use evidence, data, and/or models to engage in argument to explain the cause and effect relationships between an object and Earth's gravity, how scale and proportion affect the apparent brightness of the sun and other stars, or how plants use matter (chiefly air and water) to grow. Students observe and measure phenomenon to identify patterns that classify materials based on properties. Students can describe cause and effect relationships when mixing substances within an investigation framework.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level identify basic models to represent common features of matter and/or energy, ecosystems, and/or Earth's systems. Students recognize scale, proportion, quantity, or patterns when performing basic computations with data as it pertains to distribution of water on Earth, conservation of matter, and/or Earth's relationship with the sun, moon, and stars. Students identify evidence, data, or models to distinguish relationships between an object and Earth's gravity, how basic scale and proportion affect the brightness of the sun and other stars, or how plants use air and water. Students will observe or measure phenomenon to recognize patterns of materials. Students can identify basic relationships when mixing substances within an investigation framework.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive science instruction.

Grade 5 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level analyze how summaries reflect a meaningful, text-based sequence of the main idea and supporting details. Students compare and contrast details in literary and nonfiction/informational texts to describe and analyze genres. Students consistently recognize the paraphrase of original text. Students evaluate and analyze literary devices, author's purpose, point of view, and accuracy of fact to interpret the meaning of the text as a whole. Students consistently compare and contrast texts, and ideas within and between texts, to support inferences.

Students consistently engage in a recursive writing process to create purposeful and organized written works. Students create thoroughly organized and engaging written works by selecting and applying the organizational structure that best fits the mode, purpose, and audience.

Students skillfully use vocabulary knowledge and resources to analyze complex text through word parts, word relationships, and context clues. Students consistently use appropriate and meaningful vocabulary to enhance clarity and effectiveness in their writing.

Students consistently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students consistently locate, record, and organize relevant and reliable information on a topic in order to synthesize and clearly present findings.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically identify objective text-based summaries that include main idea, supporting details, and a logical sequence of events. Students compare and contrast details in literary and nonfiction/informational texts to classify genres. Students recognize the paraphrase of original text most of the time. Students explain how literary elements, literary devices, author's purpose, point of view, accuracy of facts, and text structure contribute to the meaning of the text. Students compare and contrast texts and ideas within and between texts.

Students engage in a recursive writing process to create purposeful written works. Students select and apply the organizational structure that best fits the mode, purpose, and audience.

Students use vocabulary knowledge and resources to interpret text through word parts, word relationships, and context clues. Students use appropriate vocabulary to write clearly and effectively.

Students frequently identify and apply appropriate use of grammar and mechanics to provide clarity and enhance communication.

Students adequately locate, record, and organize relevant and reliable information on a topic in order to present findings.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level inconsistently choose the best summary of the text and have difficulty differentiating main ideas from details. Students compare and contrast details in literary and nonfiction/informational texts but inconsistently classify genres. Students seldom identify the paraphrase of original text. Students identify literary elements, literary devices, author's purpose, point of view, or accuracy of fact. Students inconsistently compare and contrast texts and ideas within or between texts.

Students inconsistently engage in a recursive writing process to create written works. Students create written works for various purposes and audiences but inconsistently select and apply an organizational structure that fits the writing task.

Students inconsistently use vocabulary knowledge and resources to interpret text through word parts, word relationships, or context clues. Students inconsistently use appropriate vocabulary in written works.

Students inconsistently identify and apply appropriate use of grammar and mechanics.

Students ineffectively locate, record, and organize information on a topic in order to present findings.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.

Grade 6 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically estimate and solve complex problems requiring unit conversions. Students use the distance between points and transformations to solve complex problems involving congruent figures. Students analyze the differences between two outcomes of simple experiments. Students solve complex and non-routine real-world problems, draw logical conclusions, and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level estimate, illustrate, and simplify the addition and subtraction of integers and assess the reasonableness of an answer. Students solve ratio and unit rate problems. Students estimate and illustrate the multiplication and division of non-negative rational numbers. Students evaluate the validity of the value of a variable. Students generate expressions, equations, and inequalities. Students interpret the solution of an equation and assess the reasonableness of the solution. Students determine the area of polygons and composite figures. Students use relationships between angles and the triangle sum theorem to solve problems. Students estimate and solve problems requiring unit conversion. Students predict transformations, analyze lines of symmetry, and use the distance between points and transformations to solve problems involving congruent figures. Students explain and justify which measure of central tendency provides the most descriptive information for a data set. Students create and analyze box-and-whisker plots and explain and compare possible outcomes of simple experiments. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level read, order, represent, and explain rational numbers expressed as fractions, decimals, percents, and ratios. Students write positive integers as products of factors. Students illustrate or simplify the addition and subtraction of integers. Students identify and compare quantities, determine unit rates, and find equivalent fractions and percents. Students multiply and divide non-negative rational numbers. Students graph ordered pairs in all quadrants. Students represent reflective relationships between varying quantities. Students evaluate the value of a variable in expressions, equations, and inequalities. Students use number sense and properties of operations to solve equations and graph the solution. Students determine the area of parallelograms and triangles. Students identify angle relationships by name. Students identify and display the effect of transformations. Students identify lines of symmetry. Students calculate measures of central tendency, determine the sample space of simple experiments, and identify possible outcomes.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 6 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level will thoroughly comprehend, interpret, evaluate, and respond to a variety of increasingly complex texts of all literary and informational genres. Students skillfully create an objective summary including main idea and supporting details. Students effectively paraphrase main ideas with supporting details in a text. Students thoroughly compare and contrast stated or implied purposes of authors' writing. Students thoroughly evaluate literary devices, points of view, and perspectives, and they explicitly analyze how authors use key literary elements to contribute to the meaning of the text. Students consistently categorize facts included in an argument. Students analyze and evaluate complex textual evidence to support inferences and understanding within and between varied texts.

Students effectively engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In opinion writing, students strategically state an opinion supported with facts and details. Students use fully developed, complex ideas, thorough organization, purposeful word choice, a variety of fluent sentences, and appropriate voice.

Students skillfully use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer complex relationships among words with multiple meanings. Students select precise vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students intentionally apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a strong command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students thoroughly comprehend, evaluate, and synthesize resources. Students skillfully summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically comprehend, interpret, evaluate, and respond to a variety of complex texts of all literary and informational genres. Students create an objective summary including main idea and supporting details.

Students paraphrase main ideas with supporting details in a text. Students compare and contrast stated or implied purposes of authors' writing. Students evaluate literary devices, points of view, and perspectives, and they analyze how authors use key literary elements to contribute to the meaning of the text. Students categorize facts included in an argument. Students analyze textual evidence to support inferences and understanding within and between texts.

Students engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim and organize reasons and evidence. Students use fully developed ideas, strong organization, well-chosen words, fluent sentences, and appropriate voice.

Students use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer the relationships among words with multiple meanings. Students select vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions to find information on a topic. Students record and organize information from various sources. Students comprehend, evaluate, and synthesize resources. Students summarize and integrate information following a citation style with guidance and support. Students summarize and present information in a report.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level partially comprehend, interpret, evaluate, and respond to literary and informational texts, applying limited critical thinking skills. Students create a summary including main idea and limited supporting details. Students inconsistently paraphrase main ideas with limited supporting details in a text. Students inconsistently compare and contrast stated or implied purposes of authors' writing. Students inconsistently identify literary devices, points of view, and perspectives, and they describe how authors use key literary elements. Students inconsistently categorize facts included in an argument. Students inconsistently identify limited textual evidence to support inferences between texts.

Students inconsistently engage in a writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In opinion writing, students inconsistently state an opinion supported with limited facts and details. Students use partially developed ideas, weak organization, and ineffective word choice, sentences, and voice.

Students ineffectively use context clues, word parts, and reference tools to determine the meaning of words. Students may or may not infer the relationships among words with multiple meanings. Students use a limited vocabulary to communicate ideas in writing and to create an effect according to a purpose.

Students inconsistently apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a limited command of Standard English grammar, mechanics, and usage.

Students may not recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students partially comprehend, evaluate, and synthesize resources. Students ineffectively summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.

Grade 7 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically interpret equations and inequalities involving variables and rational numbers. Students make connections between circumference and area to solve problems involving circles. Students analyze, apply, and display the effect of dilations and multiple transformations. Students use central tendencies and range, predict data and select an appropriate data display, and predict theoretical probability. Students solve complex and non-routine real-world problems, draw logical conclusions, and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically estimate solutions of problems involving rational numbers and assess the reasonableness of the solutions. Students differentiate between proportional and inversely proportional relationships and identify the constant of proportionality. Students represent proportional relationships in a variety of ways. Students use representations to identify and compare unit rates. Students solve problems involving proportional relationships and assess the reasonableness of solutions. Students represent, solve, and write equations. Students solve simple inequalities. Students generate and evaluate equivalent expressions with justification of steps. Students interpret theoretical probability and draw conclusions. Students apply the effect of dilations and transformations. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level recognize, compare, and order rational numbers. Students create equivalent representations of rational numbers. Students calculate and model mathematical problems involving rational numbers and exponents. Students calculate the absolute value of a rational number. Students describe and identify a proportional relationship.

Students identify and solve problems involving ratios and unit rates. Students represent, solve, and write simple equations. Students represent, write, and graph simple inequalities. Students evaluate expressions using the order of operations. Students determine the surface area and volume of rectangular prisms and calculate the area and perimeter of trapezoids. Students calculate the circumference and area of circles. Students describe the effect of dilations and transformations.

Students calculate the measures of central tendencies and range and determine appropriate data displays. Students calculate theoretical probability.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 7 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level thoroughly comprehend, interpret, evaluate, and respond to a variety of increasingly complex texts of all literary and informational genres. Students skillfully create an objective summary including main idea and supporting details. Students effectively paraphrase main ideas with supporting details in a text. Students thoroughly compare and contrast stated or implied purposes of authors' writing. Students thoroughly evaluate literary devices, points of view, and perspectives, and they explicitly analyze how authors use key literary elements to contribute to the meaning of the text. Students consistently distinguish factual claims from opinions. Students analyze and evaluate complex textual evidence to support inferences and draw logical conclusions between and across multiple and varied texts.

Students effectively engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students strategically introduce a claim and organize well-developed reasons and evidence. Students use fully developed, complex ideas, thorough organization, purposeful word choice, a variety of fluent sentences, and appropriate voice.

Students skillfully use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer complex relationships among words with multiple meanings. Students select precise vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students intentionally apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a strong command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students thoroughly comprehend, evaluate, and synthesize resources. Students skillfully summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically read and comprehend increasingly complex literary and informational texts. Students create an objective summary including main idea and supporting details. Students paraphrase main ideas with supporting details in a text. Students compare and contrast stated or implied purposes of authors' writing. Students evaluate literary devices, points of view, and perspectives, and they analyze how authors use key literary elements to contribute to the meaning of the text. Students distinguish factual claims from opinions. Students analyze and evaluate textual evidence to support inferences and draw simple, logical conclusions between and across multiple texts.

Students engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim and organize reasons and evidence. Students use fully developed ideas, strong organization, well-chosen words, fluent sentences, and appropriate voice.

Students use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer the relationships among words with multiple meanings. Students select vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students comprehend, evaluate, and synthesize resources. Students summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level partially comprehend, interpret, evaluate, and respond to literary and informational texts, applying limited critical thinking skills. Students create a summary including main idea and limited supporting details. Students inconsistently paraphrase main ideas with limited supporting details in a text. Students inconsistently compare and contrast stated or implied purposes of authors' writing. Students inconsistently identify literary devices, points of view, and perspectives, and they describe how authors use key literary elements. Students inconsistently distinguish factual claims from opinions. Students inconsistently identify limited textual evidence to support inferences and draw weak conclusions between texts.

Students inconsistently engage in a writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim, reasons, and evidence. Students use partially developed ideas, weak organization, and ineffective word choice, sentences, and voice.

Students ineffectively use context clues, word parts, and reference tools to determine the meaning of words. Students may or may not infer the relationships among words with multiple meanings. Students use a limited vocabulary to communicate ideas in writing and to create an effect according to a purpose.

Students inconsistently apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a limited command of Standard English grammar, mechanics, and usage.

Students may not recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students partially comprehend, evaluate, and synthesize resources. Students ineffectively summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.

Grade 8 Mathematics Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically generate, simplify, and evaluate complex equivalent expressions. Students make connections between volume and surface area to solve problems involving solids and compare the volume and surface area of different solids. Students describe the impact on central tendencies of a data set with multiple outliers and when inserting or deleting multiple data points. Students solve complex and non-routine real-world problems, draw logical conclusions and justify solutions.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically generate, simplify, and evaluate equivalent expressions. Students classify and explain operational closure of rational and irrational numbers. Students distinguish between a linear and nonlinear function. Students identify independent and dependent variables. Students describe, analyze, and represent linear functions with two variables and translate between representations. Students use and apply the Pythagorean Theorem. Students describe the impact on central tendencies of a data set with an outlier and when inserting or deleting a data point. Students interpret a scatterplot, determine the rate of change, and use a line of best fit to make predictions. Students calculate, interpret, and predict experimental probability and generalize samples to populations. Students solve real-world problems and employ problem-solving strategies of identifying and using appropriate information.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level simplify and generate simple equivalent expressions, including expressions in scientific notation. Students translate between standard form and scientific notation. Students identify and compare real numbers. Students recognize if a graph represents a linear function. Students identify intercepts and slope from the graph of a line. Students identify the effect on the graph of a linear function when characteristics are changed. Students solve and graph equations and inequalities. Students use the Pythagorean Theorem to identify right triangles and to find the length of the hypotenuse. Students calculate the surface area and volume of solids. Students identify the outliers of a data set. Students identify the line of best fit from a given scatterplot and determine if the rate of change is positive or negative. Students calculate the experimental probability of single events, identify sample spaces, and classify events as independent or dependent.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive mathematical instruction.

Grade 8 Science Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically evaluate, revise, or develop a model from evidence, or apply models to complex concepts involving conservation of matter in chemical reactions, patterns in the structure and function of waves, or stability and change at varying scales in Earth's systems. Students design, evaluate, or modify investigations about stability and change of forces and motion, or analyze and draw conclusions from patterns in data about common ancestry and diversity of organisms, the geologic history of Earth, or natural hazards. Students modify, synthesize, or apply a design solution, or evaluate evidence of relationships within a design solution in various systems involving energy transfer in chemical reactions or forces in collisions. Students analyze, infer, relate, or identify complex relationships within a system to construct or evaluate explanations for evidence of anatomy and common ancestry of organisms, or aspects of Earth systems including geologic history, materials and processes, natural resources, or human impacts on those systems using the concept of patterns in cause and effect relationships or the concept of scale and proportion.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically make predictions about, describe, develop, or use a given model involving conservation of matter in chemical reactions, patterns in the structure and function of waves, or stability and change at varying scales in Earth's systems. Students identify, describe, or explain how to plan or perform investigations about stability and change of forces and motion, or identify and apply patterns in data about common ancestry and diversity of organisms, the geologic history of Earth, or natural hazards. Students use, describe, or explain a design solution, or identify evidence of relationships within a design solution in various systems involving energy transfer in chemical reactions or forces in collisions. Students construct explanations by identifying, describing, or comparing evidence of anatomy and common ancestry of organisms, or aspects of Earth systems including geologic history, materials and processes, natural resources, or human impacts on those systems using the concept of patterns in cause and effect relationships or the concept of scale and proportion.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level identify or describe basic components or concept(s) of a model involving conservation of matter in chemical reactions, patterns in the structure and function of waves, or stability and change at varying scales in Earth's systems. Students identify or describe basic steps or processes within investigations about stability and change of forces and motion, or identify and define patterns in data about common ancestry and diversity of organisms, the geologic history of Earth, or natural hazards. Students identify components of a design solution or describe simple relationships within a design solution in various systems involving energy transfer in chemical reactions or forces in collisions. Students identify or describe basic relationships shown in

evidence of anatomy and common ancestry of organisms, or aspects of Earth systems, including geologic history, materials and processes, natural resources, or human impacts on those systems using the concept of patterns in cause and effect relationships or the concept of scale and proportion.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive science instruction.

Grade 8 English Language Arts Performance Level Descriptors

Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the **Advanced** level typically thoroughly comprehend, interpret, evaluate, and respond to literary and informational texts, applying critical thinking skills. Students skillfully evaluate literary devices, points of view, and perspectives, and they skillfully analyze how authors use key literary elements to contribute to the meaning of the text. Students explicitly analyze and evaluate textual evidence to support inferences and conclusions between and across multiple texts.

Students effectively engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim, counterclaim, and support with logical reasons and evidence. Students synthesize fully developed ideas, strong organization, well-chosen words, fluent sentences, and appropriate voice.

Students skillfully use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer complex relationships among words with multiple meanings. Students select precise vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students intentionally apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a strong command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students thoroughly comprehend, evaluate, and synthesize resources. Students skillfully summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the **Proficient** level typically read, comprehend, interpret, evaluate, and respond to literary and informational texts, applying critical thinking skills. Students evaluate literary devices, points of view, and perspectives, and they analyze how authors use key literary elements to contribute to the meaning of the text. Students analyze and evaluate textual evidence to support inferences and conclusions between and across multiple texts.

Students engage in a recursive writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim, recognize a claim from an opposing viewpoint, and organize reasons and evidence. Students use fully developed ideas, strong organization, well-chosen words, fluent sentences, and appropriate voice.

Students use context clues, word parts, and reference tools to determine or clarify the meaning of words. Students infer the relationships among words with multiple meanings. Students select vocabulary to communicate ideas in writing and to create a specific effect according to a purpose.

Students apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a command of Standard English grammar, mechanics, and usage.

Students recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students comprehend, evaluate, and synthesize resources. Students summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the **Basic** level partially comprehend, interpret, evaluate, and respond to literary and informational texts, applying limited critical thinking skills. Students inconsistently evaluate literary devices, points of view, and perspectives, and they inconsistently analyze how authors use key literary elements to contribute to the meaning of the text. Students inconsistently analyze and evaluate textual evidence to support inferences and conclusions between or across multiple texts.

Students inconsistently engage in a writing process to compose narrative, informative, and argumentative responses for varied purposes and audiences. In argumentative writing, students introduce a claim and provide reasons and evidence. Students use partially developed ideas, weak organization, ineffective word choice, basic sentences, or inconsistent voice.

Students ineffectively use context clues, word parts, and reference tools to determine the meaning of words. Students may or may not infer the relationships among words with multiple meanings. Students use a limited vocabulary to communicate ideas in writing and to create an effect according to a purpose.

Students inconsistently apply knowledge of grammar and rhetorical style to analyze and evaluate a variety of texts in reading and writing. Students demonstrate a limited command of Standard English grammar, mechanics, and usage.

Students may not recognize viable research questions and well-developed thesis statements to find information on a specific topic. Students partially comprehend, evaluate, and synthesize resources. Students ineffectively summarize and paraphrase, integrate evidence, and cite sources to create written works for multiple purposes.

Below Basic: Students have not performed at least at the Basic level. Students scoring at the **Below Basic** level should be given comprehensive reading instruction.



**Oklahoma Grade 11 Life Science
Performance Level Descriptor Tables**

Advanced

Students demonstrate superior performance on challenging subject matter and clearly exhibit readiness for college and career. Students scoring at the Advanced level:

- develop and use models to interpret and evaluate components and relationships among components within and between complex systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems and/or energy in chemistry processes.
- plan and conduct investigations to produce reliable data considering the types, amounts, and accuracy of data needed; analyze and interpret complex data sets to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.
- ask questions to analyze relationships about the effect of structure and function on inheritance of traits; or support and/or evaluate the merits of arguments to synthesize and communicate understanding and defend them based on empirical evidence about stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.
- construct, evaluate, make inferences and revise an explanation based on valid and reliable evidence from a variety of sources regarding the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or evaluate or refine explanations derived from evidence from a variety of sources for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.

Proficient

Students demonstrate mastery with subject matter and exhibit readiness for college and career. Students scoring at the Proficient Level:

- develop and use models to describe components and relationships among the components of a system, related to structure and function, growth and development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, and energy in chemistry processes, including hierarchical structures and inputs and outputs of a system. Use the models to represent basic aspects of phenomena that result from changes in energy and matter.
- plan and conduct investigations to produce reliable data; analyze and interpret provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.
- ask questions to clarify relationships about the effect of structure and function on inheritance of traits; or evaluate arguments based on evidence as

students synthesize and communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.

- construct an explanation based on valid and reliable evidence from sources of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or construct and revise explanations derived from evidence from a variety of sources for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.

Basic:

Students demonstrate partial mastery with subject matter and may not exhibit readiness for college and career. Students scoring at the Basic level typically:

- identify or describe basic components or relationships among components within systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, or energy in chemistry processes.
- conduct investigations to produce data; use provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.
- ask questions to identify relationships about the effect of structure and function on inheritance of traits; or describe arguments based on evidence as students communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.
- identify and describe basic relationships based on evidence of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or identify and describe explanations from evidence for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.

Below Basic

Students scoring **Below Basic** have not demonstrated they can perform at the Basic level. Students scoring at the Basic Level:

- identify or describe basic components or relationships among components within systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, or energy in chemistry processes.
- conduct investigations to produce data; use provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.
- ask questions to identify relationships about the effect of structure and function on inheritance of traits; or describe arguments based on evidence

as students communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.

- identify and describe basic relationships based on evidence of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or identify and describe explanations from evidence for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.

LS1-2 LS1-4 LS1-5 LS1-7 LS2-5	Below Basic Students have not performed at least at the Basic level.	Basic Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR)	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR)	Advanced: Students demonstrate superior performance on challenging subject matter.
Develop and Use Models DCI <ul style="list-style-type: none"> LS1.A Structure and function LS1.B Growth and Development of Organisms LS1.C Organization for Matter and Energy Flow in Organisms LS2.B Cycles of matter and Energy Transfer In Ecosystems PS3.D Energy in Chemistry Processes CCC <ul style="list-style-type: none"> Systems and System Models Energy and matter 		Students scoring at the Basic level typically identify or describe basic components or relationships among components within systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, or energy in chemistry processes.	Students scoring at the Proficient level typically develop and use models describing components and relationships among components of a system, related to structure and function, growth and development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer In ecosystems, and energy in chemistry processes, including hierarchical structures and inputs and outputs of a system. Use the models to represent basic aspects of phenomena that result from changes in energy and matter.	Students scoring at the Advanced level typically develop and use models to interpret and evaluate components and relationships among components within and between complex systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, and/or energy in chemistry processes.

LS1-3 LS2-1 LS2-2 LS2-4 LS3-3 LS4-1 LS4-3	Below Basic: Students have not performed at least at the Limited Knowledge level.	Basic Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR)	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR)	Advanced: Students demonstrate superior performance on challenging subject matter.
<p>Planning and Carrying Out Investigations, Using Mathematics and Computational Thinking, Analyzing and Interpreting Data</p> <p>DCI</p> <ul style="list-style-type: none"> • LS1.A Structure and Function • LS2.A Interdependent Relationships in Ecosystems • LS2.B Cycles of Matter and Energy Transfer in Ecosystems • LS2.C Ecosystem Dynamics, Functioning and Resilience • LS3.B Variation of Traits • LS4.A Evidence of Common Ancestry and Diversity • LS4.B Natural Selection • LS4.C Adaptation <p>CCC</p> <ul style="list-style-type: none"> • Patterns • Scale, Proportion, Quantity • Energy and matter • Stability and Change 		<p>Students scoring at the Limited Knowledge level typically conduct investigations to produce data; use provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.</p>	<p>Students scoring at the Proficient level typically plan and conduct investigations to produce reliable data; analyze and interpret provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.</p>	<p>Students scoring at the Advanced level typically plan and conduct investigations; produce reliable data considering the types, amounts, and accuracy of data needed; analyze and interpret complex data sets to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of common ancestry and diversity, natural selection, or adaptation.</p>

<p>LS2-6 LS2-8 LS3-1 LS3-2 LS4-5</p>	<p>Limited Knowledge: Students have not performed at least at the Limited Knowledge level.</p>	<p>Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR)</p>	<p>Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR)</p>	<p>Advanced: Students demonstrate superior performance on challenging subject matter.</p>
<p>Asking Questions, Engaging in Argument from Evidence (make and defend a claim, evaluate a claim)</p> <p>DCI</p> <ul style="list-style-type: none"> LS2.C Ecosystem dynamics, functioning and resilience LS2.D Social interactions and group behavior LS3.A Inheritance of traits LS1.A Structure and function LS3.B Variation of traits LS4.C Adaptation <p>CCC</p> <ul style="list-style-type: none"> Stability and change Cause and effect 		<p>Students scoring at the Basic level typically ask questions to identify relationships demonstrating how cause of structure and function affect inheritance of traits; or describe arguments based on evidence to communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.</p>	<p>Students scoring at the Proficient level typically ask questions to clarify relationships demonstrating how cause of structure and function affect inheritance of traits; or evaluate arguments based on evidence as students synthesize and communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.</p>	<p>Students scoring at the Advanced level typically ask questions to analyze relationships demonstrating how cause of structure and function affect inheritance of traits; or support, evaluate, and defend arguments based on evidence as students synthesize and communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause and effect relationships of social interactions, group behaviors, adaptation, and variation of traits.</p>

LS1-1 LS1-6 LS2-3 LS4-2 LS4-4	Below Basic: Students have not performed at least at the Basic level.	Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR)	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR)	Advanced: Students demonstrate superior performance on challenging subject matter.
Constructing Explanations DCI <ul style="list-style-type: none"> LS1.A Structure and function LS1.C Organization for matter and energy flow in organisms LS2.B Cycles of matter and energy transfer in ecosystems LS4.B Natural selection LS4.C Adaptation CCC <ul style="list-style-type: none"> Structure and function Energy and matter Cause and effect 		Students scoring at the Basic level typically identify and describe basic relationships based on evidence of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or identify and describe explanations from evidence for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.	Students scoring at the Proficient level typically construct an explanation based on valid and reliable evidence from sources of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or construct and revise explanations from evidence from sources for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.	Students scoring at the Advanced level typically construct, evaluate, or draw inferences from an explanation based on valid and reliable evidence from a variety of sources of the cause and effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or evaluate or refine explanations from evidence from a variety of sources for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.

NAEP grade 12 Performance Level Descriptors with content extracted. NAEP only assesses science at grade 12 in high school.

Limited Knowledge	Proficient	Advanced
Students performing at the Limited Knowledge level should be able to describe, measure, classify, explain, and predict phenomena at multiple scales, from atomic/molecular to interstellar. They should be able to design and critique observational and experimental studies, and they should be able to propose and critique solutions to problems at local or regional scales.	Students performing at the Proficient level should be able to demonstrate relationships and compare alternative models, predictions, and extrapolations. They should be able to design and critique observational and experimental studies, controlling multiple variables; use scientific models to explain results; and choose among alternative conclusions based on the arguments from evidence. They should be able to compare scientific costs or risks and benefits of alternative solutions to problems at local or regional scales.	Students performing at the Advanced level should be able to use alternative models to generate predictions and explanations. They should be able to explain differences, use evidence, and be able to design and critique investigations that relate data to alternative models of phenomena. They should be able to compare costs or risks and benefits of alternative solutions to problems at local, regional, and global scales.



Oklahoma Grade 11 Physical Science Performance Level Descriptor Tables

Name: _____

Advanced

Students demonstrate superior performance on challenging subject matter and clearly exhibit readiness for college and career. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:

- **evaluate multiple patterns to develop and use models to predict how components between or within systems are related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.**
- **use complex mathematical models and plan and conduct investigations to produce and refine reliable data considering the types, amounts, accuracy, and limitations of data needed; analyze and interpret complex data sets to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.**
- **evaluate the validity and reliability of complex claims about the effects of electromagnetic radiation on matter from a variety of published sources, including complex texts.**
- **construct, evaluate, make inferences, and revise an explanation based on scientific principles using valid and reliable evidence obtained from a variety of sources to identify patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design, refine, and evaluate solutions, taking into account unanticipated effects around defining and delimiting engineering problems and interdependence of science, engineering, and technology.**

Proficient

Students demonstrate mastery with subject matter and exhibit readiness for college and career. In addition to demonstrating understanding and application of all skills in the Basic Level, students scoring at the Proficient Level typically:

- **use patterns and models to predict how components between or within systems are related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.**
- **use mathematical models and plan and conduct investigations to produce and use reliable data to serve as a basis for evidence to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.**
- **evaluate the validity and reliability of claims about the effects of electromagnetic radiation on matter from a variety of published sources.**
- **construct and revise an explanation based on scientific principles using valid and reliable evidence obtained from a variety of sources to identify patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design and refine solutions around defining and delimiting engineering problems and interdependence of science, engineering, and**

technology.

Basic

Students demonstrate partial mastery with subject matter and may not exhibit readiness for college and career. Students scoring at the Basic level typically:

- use basic patterns and models to identify and describe components between or within systems related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.
- use simple mathematical models and conduct investigations to produce data or use provided data to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.
- evaluate the validity and/or reliability of a simple claim about the effects of electromagnetic radiation on matter from a published source.
- identify and describe basic relationships and construct explanations based on evidence from a variety of sources about patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design solutions around defining and delimiting engineering problems and interdependence of science, engineering, and technology.

Below Basic

Students scoring **Below Basic** have not demonstrated they can perform at the Basic level. Students scoring at the Basic Level:

- use basic patterns and models to identify and describe components between or within systems related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.
- use simple mathematical models and conduct investigations to produce data or use provided data to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.
- evaluate the validity and/or reliability of a simple claim about the effects of electromagnetic radiation on matter from a published source.
- identify and describe basic relationships and construct explanations based on evidence from a variety of sources about patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design solutions around defining and delimiting engineering problems and interdependence of science, engineering, and technology.

PS1-1 PS3-2	Below Basic: Students have not performed at least at the Basic level.	Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR).	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR).	Advanced: Students demonstrate superior performance on challenging subject matter.
Develop and Use Models DCI <ul style="list-style-type: none"> PS1.A Structure and Properties of Matter PS3.A Definitions of Energy CCC <ul style="list-style-type: none"> Patterns Energy and Matter 		Students scoring at the Basic level typically use basic patterns and models to identify and describe components between or within systems related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.	Students scoring at the Proficient level typically use patterns and models to predict how components between or within systems are related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.	Students scoring at the Advanced level typically evaluate multiple patterns to develop and use models to predict how components between or within systems are related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.

PS1-7 PS2-5 PS3-1 PS3-4 PS4-1	Below Basic: Students have not performed at least at the Basic level.	Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR).	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR).	Advanced: Students demonstrate superior performance on challenging subject matter.
<p>Planning and Carrying Out Investigations, Using Mathematics and Computational Thinking</p> <p>DCI</p> <ul style="list-style-type: none"> • PS1.B Chemical Reactions • PS2.B Types of Interactions • PS3.A Definitions of Energy • PS3.B Conservation of Energy and Energy Transfer • PS4.A Wave Properties <p>CCC</p> <ul style="list-style-type: none"> • Energy and Matter • Cause and Effect • Systems and System Models 		<p>Students scoring at the Basic level typically use simple mathematical models and conduct investigations to produce data or use provided data to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.</p>	<p>Students scoring at the Proficient level typically use mathematical models and plan and conduct investigations to produce and use reliable data to serve as a basis for evidence to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.</p>	<p>Students scoring at the Advanced level typically use complex mathematical models and plan and conduct investigations to produce and refine reliable data considering the types, amounts, accuracy and limitations of data needed; analyze and interpret complex data sets to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different type of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.</p>

PS4-4	Below Basic: Students have not performed at least at the Basic level.	Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR).	Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR).	Advanced: Students demonstrate superior performance on challenging subject matter.
<p>Obtaining, Evaluating, and Communicating Information</p> <p>DCI</p> <ul style="list-style-type: none"> PS4.B Electromagnetic Radiation <p>CCC</p> <ul style="list-style-type: none"> Cause and Effect 		<p>Students demonstrate partial mastery of the essential knowledge and skills appropriate to college and career readiness. Students scoring at the Basic level typically evaluate the validity and/or reliability of a simple claim about the effects of electromagnetic radiation on matter from a published source.</p>	<p>Students demonstrate mastery with subject matter and exhibit readiness for college and career. Students scoring at the Proficient level typically evaluate the validity and reliability of claims about the effects of electromagnetic radiation on matter from a variety of published sources.</p>	<p>Students demonstrate superior performance on challenging subject matter and clearly exhibit readiness for college and career. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically evaluate the validity and reliability of complex claims about the effects of electromagnetic radiation on matter from a variety of published sources, including complex texts.</p>

<p>PS1-2 PS1-5 PS3-3</p>	<p>Below Basic: Students have not performed at least at the Basic level.</p>	<p>Basic: Students demonstrate partial mastery of the essential knowledge and skills that are foundational for proficient work at their grade level or course and that students are not on track to be career and college ready (CCR).</p>	<p>Proficient: Students demonstrate mastery over challenging grade-level subject matter, can analyze and apply such knowledge to real-world situations, are ready for the next grade, course, or level, and are on-track to be career and college ready (CCR).</p>	<p>Advanced: Students demonstrate superior performance on challenging subject matter.</p>
<p>Constructing Explanations and Designing Solutions</p> <p>DCI</p> <ul style="list-style-type: none"> PS1.A Structure and Properties of Matter PS1.B: Chemical Reactions PS3.A Definitions of Energy ETS1.A Defining and Delimiting Engineering Problems ETS2.B Interdependence of Science, Engineering, and Technology <p>CCC</p> <ul style="list-style-type: none"> Patterns Energy and Matter 		<p>Students scoring at the Basic level typically identify and describe basic relationships and construct explanations based on evidence from a variety of sources about patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design solutions around defining and delimiting engineering problems and interdependence of science, engineering, and technology.</p>	<p>Students scoring at the Proficient level typically construct and revise an explanation based on scientific principles using valid and reliable evidence obtained from a variety of sources to identify patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design and refine solutions around defining and delimiting engineering problems and interdependence of science, engineering, and technology.</p>	<p>Students scoring at the Advanced level typically construct, evaluate, make inferences, and revise an explanation based on scientific principles using valid and reliable evidence obtained from a variety of sources to identify patterns relating to the structure and properties of matter and chemical reactions; and define energy and matter in order to design, refine, and evaluate solutions taking into account unanticipated effects around defining and delimiting engineering problems and interdependence of science, engineering, and technology.</p>

APPENDIX E—EVALUATION RESULTS

Training Evaluation Results

I understand the goals of the standard setting meeting
 I understand the procedures we are using to set standards
 I understand how to use the standard setting materials
 I understand the differences between the performance levels
 I understand how to make the bookmark placements
 I know what tasks to expect for the remainder of the meeting
 I am confident in my understanding of the standard setting task
 I am ready to proceed with the standard setting process

N	Average	%SD	%D	%N	%A	%SA
12	4.92	0%	0%	0%	8%	92%
12	4.67	0%	0%	0%	33%	67%
12	4.67	0%	0%	0%	33%	67%
12	4.58	0%	0%	0%	42%	58%
12	4.83	0%	0%	0%	17%	83%
12	4.42	0%	0%	8%	42%	50%
12	4.67	0%	0%	0%	33%	67%
12				100%		

Procedural Evaluation Results

I understood how to make the bookmark placements
 I understood how to use the materials provided
 I understood how to record my judgments
 I thought the procedures made sense
 I was sufficiently familiar with the assessment
 I understood the differences between the performance levels

N	Average	%SD	%D	%N	%A	%SA
12	4.92	0%	0%	0%	8%	92%
12	4.92	0%	0%	0%	8%	92%
12	4.75	0%	0%	0%	25%	75%
12	4.67	0%	0%	8%	17%	75%
12	4.5	0%	0%	8%	33%	58%
12	4.67	0%	0%	0%	33%	67%

Final Evaluation Results

Please rate the usefulness of each of the following

The opening session
 Completing the practice test
 Completing the item map
 Discussions with other participants
 Impact data

N	Average	Not Useful at All			Extremely Useful	
		1	2	3	4	5
12	4.92	0%	0%	0%	8%	92%
12	4.92	0%	0%	0%	8%	92%
12	4.92	0%	0%	0%	8%	92%
12	4.92	0%	0%	0%	8%	92%
12	4.92	0%	0%	0%	8%	92%

Please rate the usefulness of each of the following

The Performance Level Definitions
 My expectations of students
 The difficulty of the test materials
 My experience in the field
 Discussions with other participants
 Decisions of other participants
 Impact data

N	Average	Not at all influential				Extremely Influential
		1	2	3	4	5
12	4.92	0%	0%	0%	8%	92%
12	4.33	0%	0%	17%	33%	50%
12	4.17	8%	0%	8%	33%	50%
12	4.58	0%	0%	8%	25%	67%
12	4.67	0%	0%	8%	17%	75%
12	4	0%	8%	8%	58%	25%
12	3.92	8%	0%	25%	25%	42%

I understood the goals of the standard setting meeting
 The facilitator helped me understand the process
 The materials contained the information needed to set standards
 I understood how to use the impact data
 I understood how the cut scores were calculated
 The facilitator was able to provide answers to my questions
 Sufficient time was allotted for training on the standard setting tasks
 Sufficient time was allotted to complete the standard setting tasks
 The facilitator helped the standard setting process run smoothly
 Overall, the standard setting process produced credible results

N	Average	%SD	%D	%N	%A	%SA
12	4.75	0%	0%	0%	25%	75%
12	4.92	0%	0%	0%	8%	92%
12	4.83	0%	0%	8%	0%	92%
12	4.58	0%	0%	8%	25%	67%
12	4.42	0%	8%	0%	33%	58%
12	5	0%	0%	0%	0%	100%
12	4	0%	8%	17%	42%	33%
12	4.25	0%	8%	8%	33%	50%
12	5	0%	0%	0%	0%	100%
12	4.92	0%	0%	0%	8%	92%

Do you believe the final recommended cut score for each performance level was Too Low, Somewhat Low, About Right, Somewhat High, or Too High?

Advanced / Proficient
 Proficient / Basic
 Basic / Below Basic

N	Average	%TL	%SL	%AR	%SH	%TH
12	3	0%	0%	100%	0%	0%
12	2.92	0%	8%	92%	0%	0%
12	2.92	0%	8%	92%	0%	0%

Demographics and Professional Experience

Panelist Demographics

Gender:

Male	3	25.00%
Female	9	75.00%

Race/Ethnicity:

White	11	91.67%
Black		0.00%
Hispanic		0.00%
Asian		0.00%
Pacific Islander		0.00%
American Indian	1	8.33%

Professional Experience:

Students with Disabilities	1	8.33%
Students with Limited English Proficiency	1	8.33%
Economically Disadvantaged Students	3	25.00%
Gifted and Talented Students	7	58.33%
General Education	12	100.00%

APPENDIX F—STANDARD SETTING RESULTS

Table F-1. 2017 OK Standard Setting Report: Round 1 CCRA Physical Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	0.3442	0.1432	0.0964	16.55%	40.49%
Proficient	0.8227	0.0838	0.4785	13.83%	23.94%
Advanced	1.3836	0.1416	0.2322	10.11%	10.11%

Table F-2. 2017 OK Standard Setting Report: Round 1 CCRA Life Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	-0.2795	0.2642	0.4274	31.00%	66.03%
Proficient	0.5126	0.0472	1.0483	26.00%	35.03%
Advanced	1.4509	0.1408	0.1886	9.03%	9.03%

Table F-3. 2017 OK Standard Setting Report: Round 2 CCRA Physical Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	0.3442	0.0604	0.1351	16.55%	40.49%
Proficient	0.8577	0.0567	0.4960	16.03%	23.94%
Advanced	1.5050	0.0319	0.0183	7.91%	7.91%

Table F-4. 2017 OK Standard Setting Report: Round 2 CCRA Life Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	0.1684	0.1064	0.0825	16.33%	46.69%
Proficient	0.6290	0.0576	0.6404	20.25%	30.36%
Advanced	1.4265	0.0246	0.0527	10.11%	10.11%

Table F-5. 2017 OK Standard Setting Report: Round 3 CCRA Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	0.3056	0.0258	0.1017	17.48%	43.47%
Proficient	0.8021	0.0211	0.4965	18.08%	25.99%
Advanced	1.5289	0.0053	0.0000	7.91%	7.91%

Table F-6. 2017 OK Standard Setting Report: Round 4 CCRA Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>SE</i>	<i>MAD</i>	<i>At %</i>	<i>At or Above %</i>
Limited Knowledge	0.1684	0.0114	0.0000	20.70%	46.69%
Proficient	0.8021	0.0131	0.6337	18.08%	25.99%
Advanced	1.5289	0.0047	0.0000	7.91%	7.91%

APPENDIX G—DISAGGREGATED IMPACT DATA

Table G-1. 2019 OK Standard Setting Report: Round 1—Physical Science

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	25,968	0.5951	7,222	0.1655	6,036	0.1383	4,412	0.1011
ELL	2,027	1,874	0.9245	116	0.0572	33	0.0163	4	0.0020
ELL w Acc	461	438	0.9501	17	0.0369	5	0.0108	1	0.0022
ELL wo Acc	1,566	1,436	0.9170	99	0.0632	28	0.0179	3	0.0019
Black African American	3,751	2,945	0.7851	409	0.1090	279	0.0744	118	0.0315
American Indian Alaskan Native	6,154	4,008	0.6513	995	0.1617	754	0.1225	397	0.0645
Hispanic or Latino	7,097	4,969	0.7002	1,044	0.1471	720	0.1015	364	0.0513
Asian	1,000	442	0.4420	156	0.1560	173	0.1730	229	0.2290
Native Hawaiian or Other Pacific Islander	136	104	0.7647	18	0.1324	9	0.0662	5	0.0368
White Caucasian	22,053	11,477	0.5204	4,006	0.1817	3,609	0.1637	2,961	0.1343
Multi Racial	3,404	1,994	0.5858	588	0.1727	487	0.1431	335	0.0984
No Response	43	29	0.6744	6	0.1395	5	0.1163	3	0.0698
Foster	166	123	0.7410	16	0.0964	19	0.1145	8	0.0482
Non Foster	43,472	25,845	0.5945	7,206	0.1658	6,017	0.1384	4,404	0.1013
Female	21,813	12,994	0.5957	3,898	0.1787	3,086	0.1415	1,835	0.0841
Male	21,788	12,948	0.5943	3,319	0.1523	2,947	0.1353	2,574	0.1181
Not Indicated	37	26	0.7027	5	0.1351	3	0.0811	3	0.0811
IEP	5,971	5,169	0.8657	447	0.0749	214	0.0358	141	0.0236
IEP w Accommm	2,689	2,361	0.8780	189	0.0703	85	0.0316	54	0.0201
IEP w o Accommm	3,282	2,808	0.8556	258	0.0786	129	0.0393	87	0.0265
Military	291	133	0.4570	61	0.2096	55	0.1890	42	0.1443
Non Military	43,347	25,835	0.5960	7,161	0.1652	5,981	0.1380	4,370	0.1008
ELL 1st Yr Proficient	159	88	0.5535	37	0.2327	24	0.1509	10	0.0629
ELL 2nd Yr Proficient	87	49	0.5632	20	0.2299	10	0.1149	8	0.0920
Econ Disadv	22,230	15,306	0.6885	3,328	0.1497	2,315	0.1041	1,281	0.0576
Non Econ Disadv	21,408	10,662	0.4980	3,894	0.1819	3,721	0.1738	3,131	0.1463
Migrant	13	8	0.6154	2	0.1538	2	0.1538	1	0.0769
Non Migrant	43,625	25,960	0.5951	7,220	0.1655	6,034	0.1383	4,411	0.1011
Plan 504	1,201	674	0.5612	207	0.1724	174	0.1449	146	0.1216
Plan 504 w Accommm	167	82	0.4910	28	0.1677	30	0.1796	27	0.1617
Plan 504 w o Accommm	1,034	592	0.5725	179	0.1731	144	0.1393	119	0.1151

Table G-2. 2019 OK Standard Setting Report: Round 1—Life Science

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	14,822	0.3397	13,529	0.3100	11,348	0.2600	3,939	0.0903
ELL	2,027	1,434	0.7074	490	0.2417	99	0.0488	4	0.0020
ELL w Acc	461	341	0.7397	99	0.2148	20	0.0434	1	0.0022
ELL wo Acc	1,566	1,093	0.6980	391	0.2497	79	0.0504	3	0.0019
Black African American	3,751	1,966	0.5241	1,127	0.3005	552	0.1472	106	0.0283
American Indian Alaskan Native	6,154	2,265	0.3681	2,051	0.3333	1,497	0.2433	341	0.0554
Hispanic or Latino	7,097	3,075	0.4333	2,272	0.3201	1,444	0.2035	306	0.0431
Asian	1,000	237	0.2370	240	0.2400	311	0.3110	212	0.2120
Native Hawaiian or Other Pacific Islander	136	70	0.5147	40	0.2941	21	0.1544	5	0.0368
White Caucasian	22,053	6,061	0.2748	6,728	0.3051	6,598	0.2992	2,666	0.1209
Multi Racial	3,404	1,131	0.3323	1,057	0.3105	916	0.2691	300	0.0881
No Response	43	17	0.3953	14	0.3256	9	0.2093	3	0.0698
Foster	166	73	0.4398	53	0.3193	32	0.1928	8	0.0482
Non Foster	43,472	14,749	0.3393	13,476	0.3100	11,316	0.2603	3,931	0.0904
Female	21,813	6,953	0.3188	7,329	0.3360	5,915	0.2712	1,616	0.0741
Male	21,788	7,853	0.3604	6,188	0.2840	5,427	0.2491	2,320	0.1065
Not Indicated	37	16	0.4324	12	0.3243	6	0.1622	3	0.0811
IEP	5,971	3,776	0.6324	1,566	0.2623	500	0.0837	129	0.0216
IEP w Accommm	2,689	1,747	0.6497	681	0.2533	210	0.0781	51	0.0190
IEP w o Accommm	3,282	2,029	0.6182	885	0.2697	290	0.0884	78	0.0238
Military	291	67	0.2302	82	0.2818	107	0.3677	35	0.1203
Non Military	43,347	14,755	0.3404	13,447	0.3102	11,241	0.2593	3,904	0.0901
ELL 1st Yr Proficient	159	35	0.2201	63	0.3962	51	0.3208	10	0.0629
ELL 2nd Yr Proficient	87	27	0.3103	28	0.3218	26	0.2989	6	0.0690
Econ Disadv	22,230	9,367	0.4214	7,099	0.3193	4,657	0.2095	1,107	0.0498
Non Econ Disadv	21,408	5,455	0.2548	6,430	0.3004	6,691	0.3125	2,832	0.1323
Migrant	13	6	0.4615	3	0.2308	3	0.2308	1	0.0769
Non Migrant	43,625	14,816	0.3396	13,526	0.3101	11,345	0.2601	3,938	0.0903
Plan 504	1,201	340	0.2831	409	0.3405	318	0.2648	134	0.1116
Plan 504 w Accommm	167	36	0.2156	55	0.3293	51	0.3054	25	0.1497
Plan 504 w o Accommm	1,034	304	0.2940	354	0.3424	267	0.2582	109	0.1054

Table G-3. 2019 OK Standard Setting Report: Round 2—Physical Science

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	25,968	0.5951	7,222	0.1655	6,997	0.1603	3,451	0.0791
ELL	2,027	1,874	0.9245	116	0.0572	33	0.0163	4	0.0020
ELL w Acc	461	438	0.9501	17	0.0369	5	0.0108	1	0.0022
ELL wo Acc	1,566	1,436	0.9170	99	0.0632	28	0.0179	3	0.0019
Black African American	3,751	2,945	0.7851	409	0.1090	304	0.0810	93	0.0248
American Indian Alaskan Native	6,154	4,008	0.6513	995	0.1617	850	0.1381	301	0.0489
Hispanic or Latino	7,097	4,969	0.7002	1,044	0.1471	832	0.1172	252	0.0355
Asian	1,000	442	0.4420	156	0.1560	205	0.2050	197	0.1970
Native Hawaiian or Other Pacific Islander	136	104	0.7647	18	0.1324	13	0.0956	1	0.0074
White Caucasian	22,053	11,477	0.5204	4,006	0.1817	4,227	0.1917	2,343	0.1062
Multi Racial	3,404	1,994	0.5858	588	0.1727	560	0.1645	262	0.0770
No Response	43	29	0.6744	6	0.1395	6	0.1395	2	0.0465
Foster	166	123	0.7410	16	0.0964	20	0.1205	7	0.0422
Non Foster	43,472	25,845	0.5945	7,206	0.1658	6,977	0.1605	3,444	0.0792
Female	21,813	12,994	0.5957	3,898	0.1787	3,533	0.1620	1,388	0.0636
Male	21,788	12,948	0.5943	3,319	0.1523	3,460	0.1588	2,061	0.0946
Not Indicated	37	26	0.7027	5	0.1351	4	0.1081	2	0.0541
IEP	5,971	5,169	0.8657	447	0.0749	243	0.0407	112	0.0188
IEP w Accommm	2,689	2,361	0.8780	189	0.0703	94	0.0350	45	0.0167
IEP w o Accommm	3,282	2,808	0.8556	258	0.0786	149	0.0454	67	0.0204
Military	291	133	0.4570	61	0.2096	63	0.2165	34	0.1168
Non Military	43,347	25,835	0.5960	7,161	0.1652	6,934	0.1600	3,417	0.0788
ELL 1st Yr Proficient	159	88	0.5535	37	0.2327	26	0.1635	8	0.0503
ELL 2nd Yr Proficient	87	49	0.5632	20	0.2299	14	0.1609	4	0.0460
Econ Disadv	22,230	15,306	0.6885	3,328	0.1497	2,649	0.1192	947	0.0426
Non Econ Disadv	21,408	10,662	0.4980	3,894	0.1819	4,348	0.2031	2,504	0.1170
Migrant	13	8	0.6154	2	0.1538	2	0.1538	1	0.0769
Non Migrant	43,625	25,960	0.5951	7,220	0.1655	6,995	0.1603	3,450	0.0791
Plan 504	1,201	674	0.5612	207	0.1724	207	0.1724	113	0.0941
Plan 504 w Accommm	167	82	0.4910	28	0.1677	36	0.2156	21	0.1257
Plan 504 w o Accommm	1,034	592	0.5725	179	0.1731	171	0.1654	92	0.0890

Table G-4. 2019 OK Standard Setting Report: Round 2—Life Science

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	23,265	0.5331	7,124	0.1633	8,837	0.2025	4,412	0.1011
ELL	2,027	1,809	0.8925	144	0.0710	70	0.0345	4	0.0020
ELL w Acc	461	431	0.9349	16	0.0347	13	0.0282	1	0.0022
ELL wo Acc	1,566	1,378	0.8799	128	0.0817	57	0.0364	3	0.0019
Black African American	3,751	2,747	0.7323	460	0.1226	426	0.1136	118	0.0315
American Indian Alaskan Native	6,154	3,591	0.5835	1,012	0.1644	1,154	0.1875	397	0.0645
Hispanic or Latino	7,097	4,550	0.6411	1,066	0.1502	1,117	0.1574	364	0.0513
Asian	1,000	375	0.3750	156	0.1560	240	0.2400	229	0.2290
Native Hawaiian or Other Pacific Islander	136	96	0.7059	17	0.1250	18	0.1324	5	0.0368
White Caucasian	22,053	10,120	0.4589	3,808	0.1727	5,164	0.2342	2,961	0.1343
Multi Racial	3,404	1,760	0.5170	597	0.1754	712	0.2092	335	0.0984
No Response	43	26	0.6047	8	0.1860	6	0.1395	3	0.0698
Foster	166	107	0.6446	24	0.1446	27	0.1627	8	0.0482
Non Foster	43,472	23,158	0.5327	7,100	0.1633	8,810	0.2027	4,404	0.1013
Female	21,813	11,502	0.5273	3,903	0.1789	4,573	0.2096	1,835	0.0841
Male	21,788	11,739	0.5388	3,214	0.1475	4,261	0.1956	2,574	0.1181
Not Indicated	37	24	0.6486	7	0.1892	3	0.0811	3	0.0811
IEP	5,971	4,924	0.8247	540	0.0904	366	0.0613	141	0.0236
IEP w Accommm	2,689	2,263	0.8416	211	0.0785	161	0.0599	54	0.0201
IEP w o Accommm	3,282	2,661	0.8108	329	0.1002	205	0.0625	87	0.0265
Military	291	115	0.3952	63	0.2165	71	0.2440	42	0.1443
Non Military	43,347	23,150	0.5341	7,061	0.1629	8,766	0.2022	4,370	0.1008
ELL 1st Yr Proficient	159	74	0.4654	34	0.2138	41	0.2579	10	0.0629
ELL 2nd Yr Proficient	87	44	0.5057	17	0.1954	18	0.2069	8	0.0920
Econ Disadv	22,230	13,947	0.6274	3,419	0.1538	3,583	0.1612	1,281	0.0576
Non Econ Disadv	21,408	9,318	0.4353	3,705	0.1731	5,254	0.2454	3,131	0.1463
Migrant	13	8	0.6154	1	0.0769	3	0.2308	1	0.0769
Non Migrant	43,625	23,257	0.5331	7,123	0.1633	8,834	0.2025	4,411	0.1011
Plan 504	1,201	591	0.4921	219	0.1823	245	0.2040	146	0.1216
Plan 504 w Accommm	167	73	0.4371	28	0.1677	39	0.2335	27	0.1617
Plan 504 w o Accommm	1,034	518	0.5010	191	0.1847	206	0.1992	119	0.1151

Table G-5. 2019 OK Standard Setting Report: Round 3—Combined

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	24,671	0.5654	7,626	0.1748	7,890	0.1808	3,451	0.0791
ELL	2,027	1,849	0.9122	131	0.0646	43	0.0212	4	0.0020
ELL w Acc	461	433	0.9393	18	0.0390	9	0.0195	1	0.0022
ELL wo Acc	1,566	1,416	0.9042	113	0.0722	34	0.0217	3	0.0019
Black African American	3,751	2,851	0.7601	456	0.1216	351	0.0936	93	0.0248
American Indian Alaskan Native	6,154	3,797	0.6170	1,066	0.1732	990	0.1609	301	0.0489
Hispanic or Latino	7,097	4,781	0.6737	1,108	0.1561	956	0.1347	252	0.0355
Asian	1,000	402	0.4020	178	0.1780	223	0.2230	197	0.1970
Native Hawaiian or Other Pacific Islander	136	100	0.7353	20	0.1471	15	0.1103	1	0.0074
White Caucasian	22,053	10,818	0.4905	4,180	0.1895	4,712	0.2137	2,343	0.1062
Multi Racial	3,404	1,895	0.5567	611	0.1795	636	0.1868	262	0.0770
No Response	43	27	0.6279	7	0.1628	7	0.1628	2	0.0465
Foster	166	114	0.6867	21	0.1265	24	0.1446	7	0.0422
Non Foster	43,472	24,557	0.5649	7,605	0.1749	7,866	0.1809	3,444	0.0792
Female	21,813	12,263	0.5622	4,164	0.1909	3,998	0.1833	1,388	0.0636
Male	21,788	12,383	0.5683	3,456	0.1586	3,888	0.1784	2,061	0.0946
Not Indicated	37	25	0.6757	6	0.1622	4	0.1081	2	0.0541
IEP	5,971	5,058	0.8471	512	0.0857	289	0.0484	112	0.0188
IEP w Accommm	2,689	2,324	0.8643	199	0.0740	121	0.0450	45	0.0167
IEP w o Accommm	3,282	2,734	0.8330	313	0.0954	168	0.0512	67	0.0204
Military	291	123	0.4227	65	0.2234	69	0.2371	34	0.1168
Non Military	43,347	24,548	0.5663	7,561	0.1744	7,821	0.1804	3,417	0.0788
ELL 1st Yr Proficient	159	84	0.5283	34	0.2138	33	0.2075	8	0.0503
ELL 2nd Yr Proficient	87	47	0.5402	19	0.2184	17	0.1954	4	0.0460
Econ Disadv	22,230	14,670	0.6599	3,581	0.1611	3,032	0.1364	947	0.0426
Non Econ Disadv	21,408	10,001	0.4672	4,045	0.1889	4,858	0.2269	2,504	0.1170
Migrant	13	8	0.6154	2	0.1538	2	0.1538	1	0.0769
Non Migrant	43,625	24,663	0.5653	7,624	0.1748	7,888	0.1808	3,450	0.0791
Plan 504	1,201	636	0.5296	223	0.1857	229	0.1907	113	0.0941
Plan 504 w Accommm	167	80	0.4790	24	0.1437	42	0.2515	21	0.1257
Plan 504 w o Accommm	1,034	556	0.5377	199	0.1925	187	0.1809	92	0.0890

Table G-6. 2019 OK Standard Setting Report: Round 4—Combined

	<i>Total N</i>	<i>Below Basic N</i>	<i>Below Basic %</i>	<i>Basic N</i>	<i>Basic %</i>	<i>Prof N</i>	<i>Prof %</i>	<i>Adv N</i>	<i>Adv %</i>
Total	43,638	23,265	0.5331	9,032	0.2070	7,890	0.1808	3,451	0.0791
ELL	2,027	1,809	0.8925	171	0.0844	43	0.0212	4	0.0020
ELL w Acc	461	431	0.9349	20	0.0434	9	0.0195	1	0.0022
ELL wo Acc	1,566	1,378	0.8799	151	0.0964	34	0.0217	3	0.0019
Black African American	3,751	2,747	0.7323	560	0.1493	351	0.0936	93	0.0248
American Indian Alaskan Native	6,154	3,591	0.5835	1,272	0.2067	990	0.1609	301	0.0489
Hispanic or Latino	7,097	4,550	0.6411	1,339	0.1887	956	0.1347	252	0.0355
Asian	1,000	375	0.3750	205	0.2050	223	0.2230	197	0.1970
Native Hawaiian or Other Pacific Islander	136	96	0.7059	24	0.1765	15	0.1103	1	0.0074
White Caucasian	22,053	10,120	0.4589	4,878	0.2212	4,712	0.2137	2,343	0.1062
Multi Racial	3,404	1,760	0.5170	746	0.2192	636	0.1868	262	0.0770
No Response	43	26	0.6047	8	0.1860	7	0.1628	2	0.0465
Foster	166	107	0.6446	28	0.1687	24	0.1446	7	0.0422
Non Foster	43,472	23,158	0.5327	9,004	0.2071	7,866	0.1809	3,444	0.0792
Female	21,813	11,502	0.5273	4,925	0.2258	3,998	0.1833	1,388	0.0636
Male	21,788	11,739	0.5388	4,100	0.1882	3,888	0.1784	2,061	0.0946
Not Indicated	37	24	0.6486	7	0.1892	4	0.1081	2	0.0541
IEP	5,971	4,924	0.8247	646	0.1082	289	0.0484	112	0.0188
IEP w Accommm	2,689	2,263	0.8416	260	0.0967	121	0.0450	45	0.0167
IEP w o Accommm	3,282	2,661	0.8108	386	0.1176	168	0.0512	67	0.0204
Military	291	115	0.3952	73	0.2509	69	0.2371	34	0.1168
Non Military	43,347	23,150	0.5341	8,959	0.2067	7,821	0.1804	3,417	0.0788
ELL 1st Yr Proficient	159	74	0.4654	44	0.2767	33	0.2075	8	0.0503
ELL 2nd Yr Proficient	87	44	0.5057	22	0.2529	17	0.1954	4	0.0460
Econ Disadv	22,230	13,947	0.6274	4,304	0.1936	3,032	0.1364	947	0.0426
Non Econ Disadv	21,408	9,318	0.4353	4,728	0.2209	4,858	0.2269	2,504	0.1170
Migrant	13	8	0.6154	2	0.1538	2	0.1538	1	0.0769
Non Migrant	43,625	23,257	0.5331	9,030	0.2070	7,888	0.1808	3,450	0.0791
Plan 504	1,201	591	0.4921	268	0.2231	229	0.1907	113	0.0941
Plan 504 w Accommm	167	73	0.4371	31	0.1856	42	0.2515	21	0.1257
Plan 504 w o Accommm	1,034	518	0.5010	237	0.2292	187	0.1809	92	0.0890

APPENDIX H—SAMPLE RATING FORM

ID	EXAMPLE_01	Example Domain 1 Panelist						
Procedural Round 384								
Directions:	For Each Round, in the column marked "Bookmark", indicate YOUR BOOKMARK PLACEMENT PAGE in the ordered item book. YELLOW AREA=BASIC, GREEN AREA=PROFICIENT, BLUE AREA=ADVANCED							
Warning:								
Item order	Item ID	RND 3 Bookmark	RND 3 Level	RND 4 Bookmark	RND 4 Level	What knowledge and skills does this item measure?	Rationale for placements outside shaded areas	
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								
21								
22								
23								
24								
25								
26								
27								
28								
29								
30								
31								
32								
33								
34								
35								
36								
37								
38								
39								
40								
41								
42								
43								
44								
45								
46								
47								
48								
49								
50								
51								
52								
53								
54								
55								
56								
57								
58								
59								
60								
61								
62								

APPENDIX I—EVALUATION FORM

Procedural Evaluation Form

OK CCRA SCI 11

The purpose of this evaluation form is to obtain your feedback about the Standard Setting process. Please complete the information below. Do not put your name on the form. We want your feedback to be confidential.

* Required

1. Please mark the appropriate circle for each statement. *

Mark only one oval per row.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I understood how to make the bookmark placements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood how to use the materials provided.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood how to record my judgements.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I thought the procedures made sense.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I was sufficiently familiar with the assessment.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood the differences between the performance levels.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2. What materials, information, or procedures were most influential in your placement of the cut scores? Why? *

3. Please provide any additional comments about the cut score placements.

Final Evaluation Form

OK CCRA SCI 11

The purpose of this evaluation form is to obtain your feedback about the Standard Setting process. Please complete the information below. Do not put your name on the form. We want your feedback to be confidential.

*** Required**

1. *Mark only one oval per row*.*

	Male	Female
Gender	<input type="radio"/>	<input type="radio"/>

2. *Mark only one oval per row.*

	White	Black	Hispanic	Asian	Pacific Islander	American Indian
Race Ethnicity	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Area of expertise (check all that apply)

- ☐ Students with Disabilities
- ☐ Students with Limited English Proficiency
- ☐ Economically Disadvantaged Students
- ☐ Gifted and Talented Students
- ☐ General Education

4. Please rate the usefulness of each of the following*

Mark only one oval per row.

	Not at all Useful	Somewhat not Useful	Neutral	Somewhat Useful	Extremely Useful
The opening session	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Completing the practice test	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Completing the item map	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Discussions with other participants	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Impact data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Please rate the influence of the following when setting standards: *

Mark only one oval per row.

	Not at all influential	Somewhat not influential	Neutral	Somewhat influential	Extremely influential
The Performance Level Definitions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My expectations of students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The difficulty of the test materials	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
My experience in the field	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Discussions with other participants	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Decisions of other participants	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Impact data	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Please select the appropriate circle for each statement. *

Mark only one oval per row.

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
I understood the goals of the standard setting meeting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The facilitator helped me understand the process.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The materials contained the information needed to set standards.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood how to use the impact data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I understood how the cut scores were calculated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The facilitator was able to provide answers to my questions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sufficient time was allotted for training on the standard setting tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sufficient time was allotted to complete the standard setting tasks.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The facilitator helped the standard setting process run smoothly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall, the standard setting process produced credible results.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Do you believe the final recommended cut scores for each performance level was

*Too Low, Somewhat Low, About Right, Somewhat High, or Too High? **

	Too Low	Somewhat Low	About Right	Somewhat High	Too High
Advanced / Proficient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Proficient / Basic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Basic / Below Basic	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Please provide any additional comments about the standard setting process or suggestions as to how the training and process could be improved.

APPENDIX J—SAMPLE ITEM LIST FORM

ID	EXAMPLE_01		
DOMAIN	1		
Directions:	Enter your notes for knowledge / skills and rationale for increased difficulty in the columns below		
Item order	Item ID	What knowledge and skills does this item measure?	Why is this item more difficult than the preceding item?
1	586659-1		
2	592071		
3	592069		
4	586636		
5	586031		
6	586218		
7	593426		
8	586106		
9	586029		
10	594357		
11	586649		
12	586701		
13	586709		
14	586693		
15	586659-2		
16	594361		
17	586108		
18	594375		
19	594354		
20	591949		
21	593424		
22	586655		
23	586691		
24	586711		
25	586027		
26	594373		
27	592073		
28	586631		
29	586110		
30	594379		
31	586640		

APPENDIX K—NONDISCLOSURE FORM

Nondisclosure Agreement
CCRA – Science Standard Setting
June 5–6, 2019

The undersigned is an employee, contractor, assessment committee member, or person otherwise authorized to view secure state assessment materials. The undersigned hereby agrees to be bound to the terms of this agreement restricting the disclosure of said materials.

It is essential to the integrity of this item development project and testing program that all test items remain secure. To maintain this security, only authorized persons are permitted to view the test questions. With the exception of materials released by the Oklahoma State Department of Education for informational purposes, all test questions (draft or final) in hardcopy or electronic format and associated materials must be regarded as secure documents. As a result, such materials may not be reproduced, electronically transmitted, discussed, used in classroom instruction, or in any way released or distributed to unauthorized persons. All materials including items and item drafts must be returned at the end of the meeting.

I understand that I am responsible for test materials security. By breaching test materials security as described here, I am breaching professional testing ethics and may be subject to additional penalties under law.

Name: _____

Signature: _____

Date: _____

APPENDIX L—MEETING AGENDA

CCRA Science Content Standard Setting Meeting June 5-6, 2019

Agenda—Day 1: Wednesday, June 5, 2019

8:15 am	Registration/Breakfast
9:00 am	Welcome and Introductions Review of Agenda and Materials Overview of the Standard Setting Process
9:45 am	Take the Test
10:15 am	Break
10:30 am	Split into Domain-Specific Groups Fill Out Item Map
11:15 am	Discuss PLDs and Describe Characteristics of “Borderline” Students
12:00 pm	Lunch in Hotel Restaurant
1:00 pm	Practice Round
1:30 pm	Readiness Discussion
2:15 pm	Training Evaluation
2:30 pm	Break
2:45 pm	Round 1
4:15 pm	Round 1 questions and discussions
5:00 pm	Adjourn

All times are approximate Breaks will take place as needed

CCRA Science Content Standard Setting Meeting June 5-6, 2019

Agenda—Day 2: Thursday, June 6, 2019

8:00 am	Breakfast and sign in
9:00 am	Introduction to Day 2
9:15 am	Round 2
10:15 am	Break
10:30 am	Reconvene as Single Group Review of PLDs and borderline definitions Round 3
12:00 pm	Lunch in Hotel Restaurant
1:00 pm	Round 4
2:15 pm	Break
3:00 pm	Round 4 questions and discussions
4:15 pm	Final Evaluation
4:30 pm	Adjourn

All times are approximate Breaks will take place as needed

APPENDIX M—FINAL CUTPOINTS

Table M-1. 2019 OK Standard Setting Report: Final Cutpoints—CCRA Science

<i>Performance Level</i>	<i>Theta Cut</i>	<i>At %</i>	<i>At or Above %</i>
Below Basic	0.1684	53.31%	100.00%
Basic	0.8021	20.70%	46.69%
Proficient	1.5289	18.08%	25.99%
Advanced	0.1684	7.91%	7.91%

APPENDIX O
2022 CCRA US HISTORY
STANDARD SETTING REPORT



2022 Oklahoma Standard-Setting Report

US History

June 23–24, 2022—Oklahoma City, Oklahoma

Prepared by Cognia for the Oklahoma Department of Education



TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF STANDARD-SETTING PROCEDURES	3
CHAPTER 2. TASKS COMPLETED PRIOR TO STANDARD SETTING	4
2.1 CREATION OF PERFORMANCE LEVEL DESCRIPTORS	4
2.2 PREPARATION OF MATERIALS	4
2.3 PREPARATION OF THE STANDARD-SETTING TOOLKIT FOR USE DURING THE MEETING	5
2.4 SELECTION OF PANELISTS	6
CHAPTER 3. TASKS COMPLETED DURING THE STANDARD-SETTING MEETING.....	7
3.1 OVERVIEW OF THE ID MATCHING METHOD	7
3.2 GENERAL ORIENTATION AND PANELIST TRAINING	7
3.3 BECOMING FAMILIAR WITH THE TEST ITEMS AND CONTENT	8
3.4 USE OF THE STANDARD-SETTING TOOLKIT	8
3.5 REVIEW OF BORDERLINE PERFORMANCE LEVEL DESCRIPTORS	8
3.6 JUDGMENT ROUNDS AND FEEDBACK	8
3.6.1 Modeling and Practice	9
3.6.2 Round 1 Judgments and Results	9
3.6.3 Round 2 Judgments and Results	10
3.6.4 Round 3 Judgments and Results	10
3.6.5 Workshop Evaluation	11
CHAPTER 4. TASKS COMPLETED AFTER THE STANDARD-SETTING MEETING.....	12
4.1 ANALYSIS AND REVIEW OF PANELISTS' FEEDBACK	12
4.2 POLICY ADJUSTMENTS	12
4.3 PREPARATION OF STANDARD-SETTING REPORT	12
REFERENCES.....	13
APPENDICES.....	14
APPENDIX A—POWERPOINT PRESENTATION	
APPENDIX B—MEETING AGENDAS	
APPENDIX C—NONDISCLOSURE FORM	
APPENDIX D—PERFORMANCE LEVEL DESCRIPTORS	
APPENDIX E—COGNIA STANDARD-SETTING TOOLKIT	
APPENDIX F—READINESS SURVEYS	
APPENDIX G—WORKSHOP EVALUATION SURVEY	
APPENDIX H—PANELISTS	
APPENDIX I—STANDARD-SETTING ROUND RESULTS	
APPENDIX J—WORKSHOP EVALUATION RESULTS	
APPENDIX K—STANDARD-SETTING MEMO	
APPENDIX L—FINAL CUT POINTS	

Chapter 1. Overview of Standard-Setting Procedures

The purpose of this report is to summarize the activities involved in the standard-setting process for the Oklahoma College and Career Readiness Assessment (CCRA) in US History on behalf of the Oklahoma State Department of Education (SDE). The need for standard setting arises from the fact that this is a new assessment that was administered operationally for the first time in 2022. For such new assessments, performance standards must be set. The primary goal of the standard setting was to determine the knowledge, skills, and abilities (KSAs) that students must demonstrate to be classified into one of the performance levels (i.e., Advanced, Proficient, Basic, and Below Basic).

The standard-setting process used was the Item-Descriptor (ID) Matching method (Ferrara & Lewis, 2012; Cizek & Bunch, 2007). The ID Matching method was selected because it reduces cognitive burden on panelists as compared to other standard-setting methods that require probability judgments about hypothetical high- and low-performing students, and it most clearly translates content standards into performance categories as compared to other methods of standard setting (Cizek, Bunch, & Koons, 2004).

The standard-setting meeting was held from June 23rd through June 24th of 2022. In all, 11 panelists participated in the process and were organized into 3 tables of 3–4 panelists each plus a facilitator provided by Cognia.

This report is organized into three major sections, describing tasks completed prior to, during, and after the standard-setting meeting.

Chapter 2. Tasks Completed Prior to Standard Setting

2.1 Creation of Performance Level Descriptors

Oklahoma State Statute: Title 70. Schools, Chapter 22 – Testing and Assessment, Section 1210.541 – Student Performance Levels and Cut Scores – Accountability System mandates the adoption of “a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.” The law states that performance levels must be labeled and defined as follows:

1. Advanced, which shall indicate that students demonstrate superior performance on challenging subject matter;
2. Proficient, which shall indicate that students demonstrate mastery over appropriate grade-level subject matter and that students are ready for the next grade, course, or level of education, as applicable;
3. Basic, which shall indicate that students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level or course; and
4. Below Basic, which shall indicate that students have not performed at least at the limited knowledge level.

The PLDs were drafted by Cognia and approved by SDE in early 2020. SDE reviewed the PLDs electronically. The Borderline PLDs, used in the standard-setting process, were created jointly between Cognia team members and SDE team members through a virtual meeting in June 2022. Dr. Steve Ferrara gave a presentation at the start of the meeting on the importance of Borderline PLDs and how to draft them. During the meeting, the PLDs drafted in 2020 were used as a reference document in the creation of the Borderline PLDs.

2.2 Preparation of Materials

The following materials were assembled for presentation to the panelists at the standard-setting meeting in paper or digital form (as indicated):

- Opening session and workshop facilitator PowerPoint slides
- PLDs (paper)
- Meeting agendas (paper)
- Nondisclosure forms (paper)
- Test booklets (paper)
- Cognia Standard-Setting Toolkit (digital) which included the following: Practice item booklet, integrated item map and ordered item booklet, readiness surveys, and judgment forms.
- Evaluation forms (paper)

The PowerPoint presentation used in the opening session was prepared and approved by the SDE and TAC prior to the meeting. The same PowerPoint presentation slide deck also included the workshop facilitator slides used during the main portion of the standard-setting meeting. A copy of the presentation is included in Appendix A. Copies of the meeting agenda, nondisclosure forms, PLDs, the Cognia Standard-Setting Toolkit, the readiness surveys, and the workshop evaluation form are included in Appendices B through G.

2.3 Preparation of the Standard-Setting Toolkit for use during the Meeting

This section provides details about the Cognia Standard-Setting Toolkit that panelists used to complete standard-setting activities during the meeting. In addition, the setup of the digital ordered item booklet with integrated item map is discussed.

The Cognia Standard-Setting Toolkit was developed, tested, and set up by Cognia prior to the meeting and included a digital ordered item booklet with integrated item map, judgement forms, and readiness surveys. During traditional paper-based standard setting meetings, panelists would be provided with an ordered item book where each page in the book represented a different item, and the items were sorted by difficulty. In addition, panelists would also use an item map which consisted of a list of items that correspond to the pages in the ordered item booklet. Finally, panelists would have paper-based judgement forms which included space for panelists to write notes and make their judgments.

The Cognia Standard-Setting Toolkit consisted of a digital interface that first presented the ordered item map view (i.e., a list of items separated by rows with the easiest item at the bottom and the most difficult at the top). From the initial screen panelists could easily toggle to the corresponding ordered item booklet view (i.e., viewing each item as a single page with the option to use navigation arrows to move 'up' or 'down' in the booklet to a more difficult or easier item). The ordered item booklet was created by sorting the items according to their item response theory (IRT)-based difficulty values ($RP = .67$ was used). A three-parameter logistic IRT model was used to calculate the RP_{67} values for dichotomous items.

Integrated judgement forms were available within both the item map and booklet view. The judgment forms provided space for users to note (1) the relevant knowledge, skills, and abilities (KSAs) needed to answer the item, (2) why the item is more difficult than the previous item, (3) item descriptor matches, and (4) cut placements. Any notes entered by the user in the item map view screen would remain in place when the user switched to the booklet view screen and vice versa. In addition to the above, the toolkit included the round-specific readiness surveys that panelists completed before undertaking each judgment round.

Additional details and screenshots of the Cognia Standard-Setting Toolkit are available in Appendix E.

2.4 Selection of Panelists

As emphasized in Cizek and Bunch (2007), regardless of the method used, the selection of panelists is an important factor in determining standard-setting outcomes and maximizing the validity of the standard-setting process. The guidance provided by *Standards for Educational and Psychological Testing* (AERA et al., 1999) states that “a sufficiently large and representative group of judges should be involved to provide reasonable assurance that results would not vary greatly if the process were repeated.”

Consistent with the above guidance and respecting practical considerations regarding the maximum size of a group that can be successfully managed, the goal was to recruit a standard-setting panel of 10–12 members representing different stakeholder groups to set standards for US History. Targets for the size and composition of the panel were also consistent with federal guidelines as described in *Standards and Assessment Peer Review Guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001* (U.S. Department of Education, 2009).

The SDE selected panelists prior to the standard-setting meeting. The goal for panel selection was to include participants who are primarily teachers, but also to include school administrators, higher education personnel, and stakeholders from other interest groups. Moreover, to the extent possible, panelists were selected to reflect a balance of gender, race/ethnicity, and geographic location. Finally, panelists were selected who were familiar with the high school US History subject matter. A list of the panelists is included in Appendix H.

Chapter 3. Tasks Completed During the Standard-Setting Meeting

3.1 Overview of the ID Matching Method

The Item-Descriptor (ID) Matching method is appropriate for setting standards for standards-aligned assessments like the CCRA U.S. History assessment. Assessment programs around the world have used ID Matching (e.g., Delaware, Massachusetts, Maryland, Mississippi, New Mexico, New York, South Carolina, and West Virginia; the Chicago and Philadelphia Public Schools; and programs in Brazil, Germany, and Finland).

ID Matching has advantages over Bookmark, Angoff, and other standard-setting methods. Specifically, its cognitive-judgmental task requires that standard-setting panelists, who are typically classroom educators, undertake a judgmental task that they are well suited for—matching item knowledge and skill response demands with knowledge and skill expectations in performance level descriptors (PLDs). The Bookmark and other methods require panelists to make probability judgments—something that people in general do not do well (e.g., Murphy, 2002). In addition, panelists do not need to hold a hypothetical borderline student in mind when they match items to descriptors and recommend cut scores, so the cognitive load and complexity of ID Matching is more manageable.

During standard setting using ID Matching, panelists use borderline PLDs as their guide to match items to performance level descriptors. The structure of the PLDs provides a general characterization of expected student knowledge and skill at each level and examples of the knowledge and skills that students at each achievement level can be expected to demonstrate. The ordering of items by their empirical difficulty facilitates the matching process. By matching test items to specific claims from the borderline Proficient PLD, for example, panelists identify the evidence in test items that supports the claims in that descriptor. Supporting the claims represented in the borderline Proficient PLD contributes to the validity of interpretations of student achievement, based on the PLDs, and to the overall validity argument that a student who achieves that level on the assessment has demonstrated adequate understanding of essential concepts with respect to the standards being measured. This logic applies to all cut scores and performance levels.

3.2 General Orientation and Panelist Training

Concerning panelist training, the *Standards for Educational and Psychological Testing* (AERA et al., 2014) states the following:

Care must be taken to assure these persons understand what they are to do and that their judgments are as thoughtful and objective as possible. The process must be such that well-qualified participants can apply their knowledge and experience to reach meaningful and relevant judgments that accurately reflect their understandings and intentions. (p. 101)

The training of the panelists began with a general orientation session at the start of the standard-setting meeting. The purpose of the orientation was to ensure that all panelists received the same information about the need for and the goals of standard setting, and about their part in the process.

3.3 Becoming Familiar with the Test Items and Content

The first step after the opening session was for the panelists to take the US History test. The purpose of this step was to familiarize the panelists with the assessment and the test taking activities expected of students during administration. Once panelists completed the test, the answer key was distributed. At this point, panelists were encouraged to discuss any issues regarding items or scoring.

3.4 Use of the Standard-Setting Toolkit

Panelists were organized into tables such that each table included 3–4 panelists. Panelists used the provided laptop computers to securely access the Cognia Standard-Setting Toolkit. Within the digital tool, each panelist reviewed the domain-specific ordered item booklet item by item, considering the KSAs students needed to answer each one.

Panelists used the integrated ordered item booklet and judgment forms available within the Cognia Standard-Setting Toolkit to complete their judgments. The judgment form included space for the panelists to type in the KSAs required to answer each item correctly and to indicate why they believed each item was more difficult than the previous one. To ensure each panelist was comfortable using the provided laptop computers and understood the mechanics of data entry, Cognia Psychometricians Dr. Frank Padellaro, Dr. Robert Cook, and Dr. Robert Keller reviewed the technology the panelists would use to complete their judgment forms.

3.5 Review of Borderline Performance Level Descriptors

Before engaging in the judgment tasks, panelists reviewed the borderline PLDs. This important step was designed to ensure that panelists thoroughly understood the KSAs needed for students to be classified into performance levels (Below Basic, Basic, Proficient, and Advanced). The borderline PLDs are provided in Appendix D.

3.6 Judgment Rounds and Feedback

During the main portion of the standard-setting workshop, panelists completed a practice round followed by three consecutive rounds of judgments. After the completion of each judgment round, Cognia psychometricians calculated a variety of statistics which served various functions: feedback to panelists as part of the standard-setting process, reporting to Cognia and the SDE as intermediate evidence for the impact of panelists' judgments, and as quality control metrics. For each round, Cognia psychometricians calculated the median cut scores for the group based on their cut score recommendations, theta scale cut scores, the conditional standard error of measurement (SEM) for each of the cut scores, and impact data (i.e., the percentage of students in each performance level).

For each round, the overall cut points were determined by first calculating the median of the individual cut points obtained from each panelist, and then calculating the average of the RP67 thetas associated with the median OIB page number and the item just below it in the ordered item booklet. This calculation was repeated for each performance level cut point. The Mean Absolute Difference of the panelists' cut points indicates the extent to which judgments were consistent across panelists and reflects the level of

agreement among the ratings with each successive round of ratings. Conditional SEM characterizes the measurement precision for each of the scale cuts. Finally, impact data reflect the percentage of students across the state who would fall into each performance level category according to the total group median cut points. While these statistics were available, the only results revealed to panelists were those that were appropriate for the goals of the specific round. Results for panelist ratings across all rounds are displayed in Appendix I.

3.6.1 Modeling and Practice

To begin, the panelists completed a practice round of judgments. The purpose of the practice round was to familiarize the panelists with all the materials they would be using for the standard-setting process and become facile with the ID Matching judgments. Panelists used the provided laptop computers to access digital copies of the borderline PLDs and standards. In addition, panelists were provided with credentials to access the Cognia Standard-Setting Toolkit. Within the digital tool, panelists were presented with a practice ordered item book, which consisted of 6 items representing the range of difficulty on the test, as well as the integrated digital judgment forms.

The facilitator demonstrated how to navigate within the standard-setting tool and how to use the tool to make their judgments. Additionally, Cognia Psychometrician Dr. Frank Padellaro reviewed the technology panelists would use to complete their judgments, to ensure each panelist understood how to use the Cognia Standard-Setting tool. Then, beginning with the first ordered item and considering the skills and abilities needed to complete it, panelists were instructed to ask themselves two questions: (1) “What are the knowledge, skills, and abilities a student needs to respond to this item?” and (2) “Why is this item more difficult than the previous item?” Panelists considered each ordered item in turn, asking themselves the same two questions and assigning item descriptor matches (i.e., below basic, basic, proficient, advanced, or the threshold between two levels) to each item. The facilitator then led the panelists in a readiness discussion, asking panelists to share the reasoning behind their item descriptor matches with the group and assessing each panelist’s understanding of the judgment task and borderline PLDs.

At the end of the practice round, panelists completed the round one readiness survey (Appendix F). The readiness survey was designed to ascertain whether the panelists were comfortable moving ahead to the judgment task. Once all panelists completed the Round 1 Readiness Survey, Cognia psychometricians reviewed the responses to make sure panelists were ready to undertake the first round of judgments. In the event of any uncertainty (based on the survey responses), the specific information was relayed to the facilitator so that any questions or issues could be addressed before proceeding to the Round 1 judgments.

3.6.2 Round 1 Judgments and Results

In the first round, panelists worked individually with the borderline PLDs, the standard-setting tool, and the ordered item booklet (OIB). Beginning with the first ordered item and considering the skills and abilities needed to complete it, Panelists considered each ordered item in turn, asking themselves the same two questions and assigning item descriptor matches (i.e., below basic, basic, proficient, advanced, or threshold) to each item. They continued in this manner until they located a threshold region (a region in the item descriptor matches alternated between two performance levels), then placed their cut at the item

that marked the beginning of the region based on their judgments. Panelists then repeated the process for the other two cut points and used the integrated judgment forms to record their notes and judgments.

After the completion of round one, Cognia psychometricians calculated a variety of statistics as described previously. As a reminder, the Round 1 overall cut points were determined by first calculating the median of the individual cut points obtained from each panelist, and then calculating the average of the RP67 thetas associated with the median OIB page number and the item just below it in the ordered item booklet.

3.6.3 Round 2 Judgments and Results

The purpose of Round 2 was for panelists to discuss their Round 1 cut score recommendations and, if they determined it necessary, to revise their judgments. Prior to beginning their discussions, panelists were presented with the median cut scores based on their Round 1 judgments for each performance level cut score. The facilitator presented this information to the group using a projector and laptop and explained how to use it as they completed their discussions. The distribution of panelists' cut points was presented graphically, as histograms, in terms of location in the item map.

Panelists were then given the opportunity to share their individual rationales for their cut placements in terms of the necessary knowledge and skills for each classification. Panelists were asked to pay particular attention to how their individual judgments compared to those of other panelists in their room to assess whether they were unusually stringent or lenient within the group. They also were reminded to make their own independent judgments and that they did not have to agree with other panelist recommendations. Once the discussions were complete, panelists completed the round two readiness survey (Appendix F). The readiness survey was designed to ascertain whether the panelists were comfortable moving ahead to the second round of the judgment task. Once all panelists completed the Round 2 Readiness Survey, Cognia psychometricians reviewed the responses to make sure panelists were ready to undertake their second round of judgments. In the event of any uncertainty (based on the survey responses), the specific information was relayed to the facilitator so that any questions or issues could be addressed before proceeding to the Round 2 judgments.

Once all panelists indicated that they were ready to undertake the next round, they were given the opportunity to revise or retain their Round 1 judgments on the judgment forms within the digital tool. Panelists were told to place cut scores according to their individual best judgments; consensus among the panelists was not necessary. They were encouraged to listen to the points made by their colleagues but not to feel compelled to change their cut placements. When Round 2 judgments were complete, Cognia psychometricians calculated the statistics described previously and discussed the results with SDE staff. In addition, the results and associated impact data were presented to panelists at the conclusion of round 2.

3.6.4 Round 3 Judgments and Results

The purpose of Round 3 was for panelists to discuss their Round 2 cut score recommendations and, if necessary, to revise their judgments. Prior to beginning their discussions, panelists were presented with the median cut scores based on their Round 2 judgments as well as impact data for each performance level cut. The facilitator presented this information to the group using a projector and laptop and explained

how to use it as they completed their discussions. The distribution of panelists' cut points was presented graphically, as histograms, in terms of location in the ordered item booklet. The impact data was presented graphically in the form of a stacked bar chart.

Panelists were then given the opportunity to share their individual rationales for their cut score placements in terms of the necessary knowledge and skills for each classification. Panelists were asked to pay particular attention to how their individual judgments compared to those of other panelists in their room to assess whether they were unusually stringent or lenient within the group. Once the discussions were complete, panelists completed the round three readiness survey. The readiness survey was designed to ascertain whether the panelists were comfortable moving ahead to the second round of the judgment task. Once all panelists completed the Round 3 Readiness Survey, Cognia psychometricians reviewed the responses to make sure panelists were ready to undertake their second round of judgments. In the event of any uncertainty (based on the survey responses), the specific information was relayed to the facilitator so that any questions or issues could be addressed before proceeding to the Round 3 judgments.

Once all panelists indicated that they were ready to undertake the next round, they were given the opportunity to revise or retain their Round 2 judgments on the judgment forms within the digital tool. Panelists were told to place cuts according to their individual best judgments; consensus among the panelists was not necessary. They were encouraged to listen to the points made by their colleagues but not to feel compelled to change their cut placements. When Round 3 judgments were complete, Cognia psychometricians calculated the statistics described previously and discussed the results with SDE staff.

3.6.5 Workshop Evaluation

At the conclusion of the standard-setting meeting, panelists completed a final workshop evaluation form and gave their feedback on various aspects of the standard-setting meeting. Panelists indicated that they felt positive about how Cognia conducted the workshop and their final recommendations. Specifically, panelists expressed generally positive support for the workshop overall; workshop facilitation; training, practice, and the workshop process; the Cognia Standard-Setting tool; and other details in the standard-setting workshop process. When asked about panelists' perceptions in final cut scores, as shown in Table 1 of Appendix J, all panelists indicated that they were satisfied with final group cut scores. A copy of the evaluation survey is available in Appendix G; the workshop evaluation results are available in Appendix J.

Chapter 4. Tasks Completed After the Standard-Setting Meeting

Upon conclusion of the standard-setting meeting, several important tasks were completed. These tasks centered on the following: reviewing the standard-setting process and addressing issues presented by the outcomes; presenting the results to the SDE; and making any final revisions or adjustments based on policy considerations, under direction of the SDE. Shortly after the standard-setting meeting, Cognia provided SDE with a standard-setting memo that included an overview of the standard-setting process, as well as the final recommended cut scores. A copy of the memo is available in Appendix K.

4.1 Analysis and Review of Panelists' Feedback

The standard-setting literature considers evaluation of the workshop and its results to be another product of the standard-setting process (e.g., Reckase and Chen, 2012), as it provides important validity evidence supporting the cut scores that are obtained. To provide evidence of the participants' views of the standard-setting process, panelists were asked to complete a questionnaire at the end of the meeting.

After the evaluation forms were completed, panelists' responses were reviewed. This review did not reveal any anomalies in the standard-setting process or indicate any reason that a particular panelist's data should not be included when the final cut points were calculated. In general, participants felt that the recommended cut points were appropriate and that their judgments were based on appropriate information and decision making. The results of the evaluations are presented in Appendix J.

4.2 Policy Adjustments

After all standard-setting activities had been completed and all materials reviewed, the SDE recommended no adjustments to the Round 3 cuts as recommended by panelists at the standard-setting meeting. The full set of cuts are shown in Appendix L were presented to the CEQA and approved for use assigning students to performance levels in the 2022–2023 Oklahoma US History assessments.

4.3 Preparation of Standard-Setting Report

Following the final compilation of standard-setting results, Cognia prepared this report, which documents the procedures and results of the 2022 standard-setting meeting that was held to establish performance standards for the assessment.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: Establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press
- Reckase, M.D. (2001). Innovative methods for helping standard-setting participants to perform their task: The role of feedback regarding consistency, accuracy, and impact. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives* (pp. 159–173). Mahwah, NJ: Lawrence Erlbaum Associates.
- U.S. Department of Education. (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: U.S. Department of Education Office of Elementary and Secondary Education. Retrieved June 10, 2010, from the World Wide Web: www2.ed.gov/policy/elsec/guid/saaprguidance.pdf.

Appendices

APPENDIX—A
POWERPOINT PRESENTATIONS

Welcome!

- Introductions
- One minute each panelist
 - Your name, school district, what you teach
 - Experience in other standard-setting workshops
- Ask for show of hands
 - Who's been involved in SS before?
 - Which method(s)?
- Review the agenda

Overview

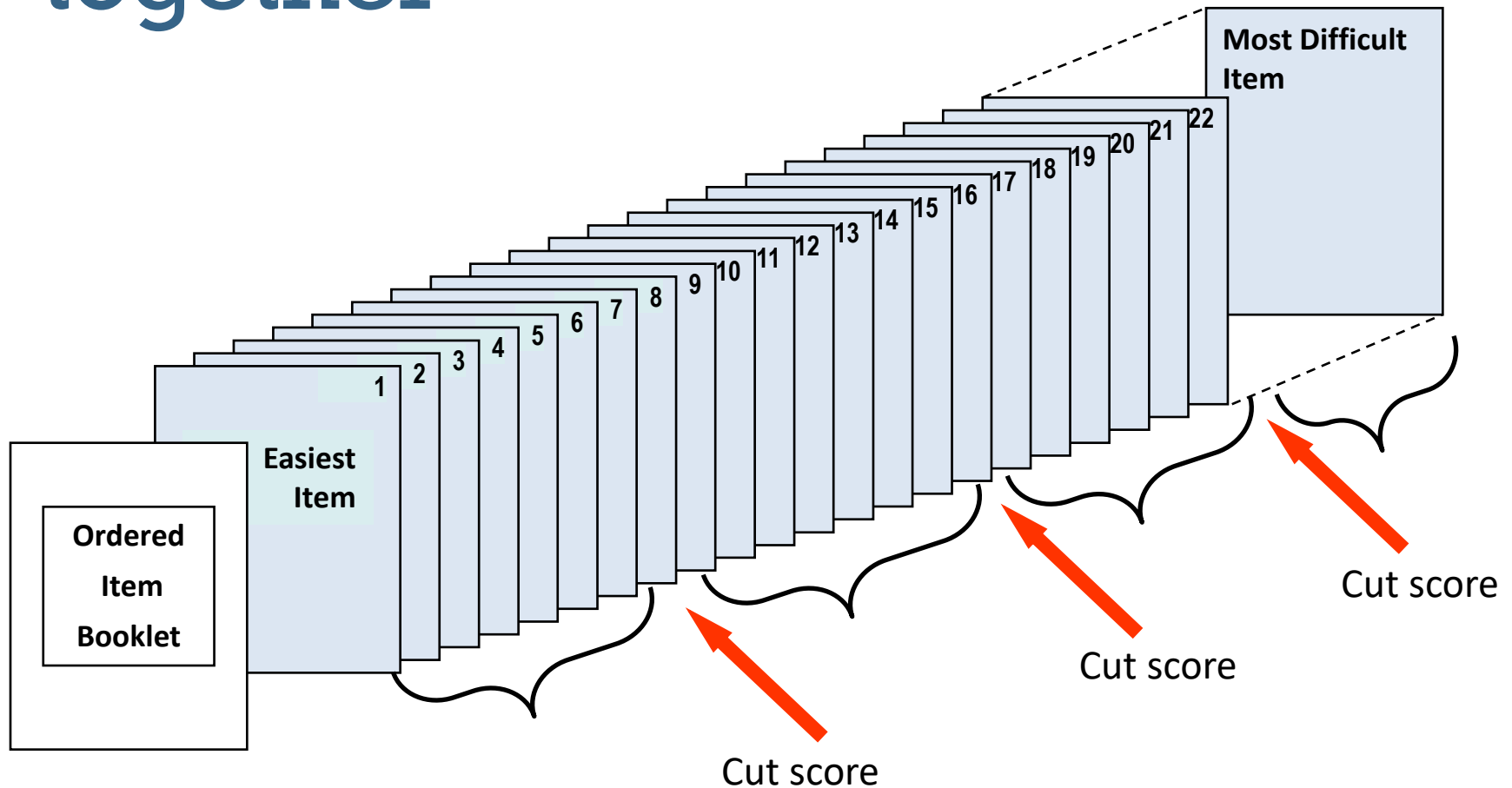
Rhythm

1. Become familiar with borderline PLDs, test items, training, and practice and using the standard-setting tool
2. Prepare for round 1
3. Complete round 1
4. Review feedback from round 1, prepare for round 2
5. Complete round 2
6. Etc.

Overview (cont.)

- Our shared goals
 - Get your recommendations for performance standards for CCRA US History assessment that provide meaningful and actionable information
- Your goals as panelists
 - Learn the concepts and procedures to recommend cut scores following Item-Descriptor (ID) matching
 - Follow the procedures we train you on
 - Recommend cut scores for Advanced, Proficient, Below Basic, and Basic
 - Rely on your expertise about the content standards, student learning, and students throughout the process

The outcome we're pursuing together



Note: Width of brackets irrelevant

At each table

- Introductions
- Pick a table leader
 - Facilitate discussion
 - Engage all panelists
 - Ask for help from facilitator, psychometricians for tool, Cognia and OSDE content experts
 - No need to act as spokesperson for your table; individuals can speak up for themselves

Key concepts and procedures

- Borderline PLDs
- ID Matching judgmental task
- Item map, OIB, online tool
- Threshold regions
- Become familiar with test items

- Preparation for the round 1
- Preparation for the round 2
 - Feedback interpretations and uses
- Preparation for the round 3

Performance level descriptors (PLDs)

- PLDs define knowledge and skills we can expect of students at each performance level
 - Advanced, Proficient, Basic, Below Basic
- **Range PLDs:** *solid* performance in a level
- **Borderline PLDs:** performance that is *just barely* in a level
- Review the borderline PLDs in the tool

Understanding the Borderline PLDs

Advanced:

Students at the borderline of the **Advanced** level can demonstrate superior performance on the challenging subject matter through the process of making connections more than 50% of the time on the assessment. While these students sometimes may only demonstrate the understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 50% of the time:

Understanding the Borderline PLDs continued

Proficient:

Students at the borderline of the Proficient level can demonstrate mastery over appropriate subject matter more than 50% of the time on the assessment. . . .

Basic:

Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills of the appropriate subject matter more than 50% of the time on the assessment. . . .

Understanding the Borderline PLDs continued

- Advanced:
 - Analyze the causes and effects of the United States developing in a world power in the late Nineteenth and early Twentieth centuries.
 - Evaluate how both the outbreak and events of World War II transformed the United States.
- Proficient:
 - Examine the causes and effects of the United States developing in a world power in the late Nineteenth and early Twentieth centuries.
 - Summarize how both the outbreak and events of World War II transformed the United States.
- Basic:
 - Ineffectively describe the causes and effects of the United States developing in a world power in the late Nineteenth and early Twentieth centuries.
 - Partially examine how both the outbreak and events of World War II transformed the United States.

**Connection
between OK
PLDs Verbiage
and Marzano's
Taxonomy**

Marzano Cognitive System Category	OK USH Standards Verbiage	OK USH PLD Verbiage
Comprehension <ul style="list-style-type: none"> • Synthesis • Representation 	Examine	Examine
Comprehension <ul style="list-style-type: none"> • Synthesis • Representation 	Summarize	Summarize
Analysis	Analyze	Analyze
Analysis	Evaluate	Evaluate

Modeling: The ID matching process

- Now I'll model the ID Matching process for item
- (a) Answer the two questions
 - What does a student need to know/be able to do to respond to this item/at this score level?
 - What makes this item more difficult than all previous items?
- (b) Match the items to a PLD
 - Explain how the item response demands align with PLD expectations
- I'll think out loud
- You'll see me do this again—then you'll practice doing it

Your answers identify the item's knowledge and skill demands

Standard-setting tools

Online Tool

- Item map
- Ordered item book
- Borderline PLDs
- Space for you to make notes to yourself (“item review”)
 - E.g., notes on your answers to the two questions
- Spaces for you to enter your item-PLD matches and to indicate your cut score recommendations (“judgment round”)

Standard-setting tool

- Demonstrate all other information and functionality
- Taking notes
 - Answering the two questions
 - Matches to PLDs
- Indicating your cut score recommendations

Item map and OIB

Item Map

- Each **line** contains one test item
- Items are ordered by difficulty: easiest to most difficult

OIB

- Each **page** contains one test item
- Items are ordered by difficulty: easiest to most difficult
- Passage(s) and other stimuli

The ID matching judgmental task

- Practice and guided feedback

ID matching judgmental task

- Step (a) Answer the two questions
- Step (b) Match items to PLDs
- Step (c) comes later
 - (Select your cut score in the threshold region)
- Work independently
- Take notes in the Tool

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

Which PLD most closely matches the knowledge and skill demands for each item?

Hint: Items are ordered by difficulty.

Panelist practice

- Facilitator models for one more item
 - Answer the two questions, match items to PLDs
 - Think out loud
 - Explain your thinking as a content-based rationale
- Panelists practice independently; enter answers to two Qs in the tool
- Table discussion: Share insights, look for shared understandings, no persuasion
- Room discussion: guided feedback

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

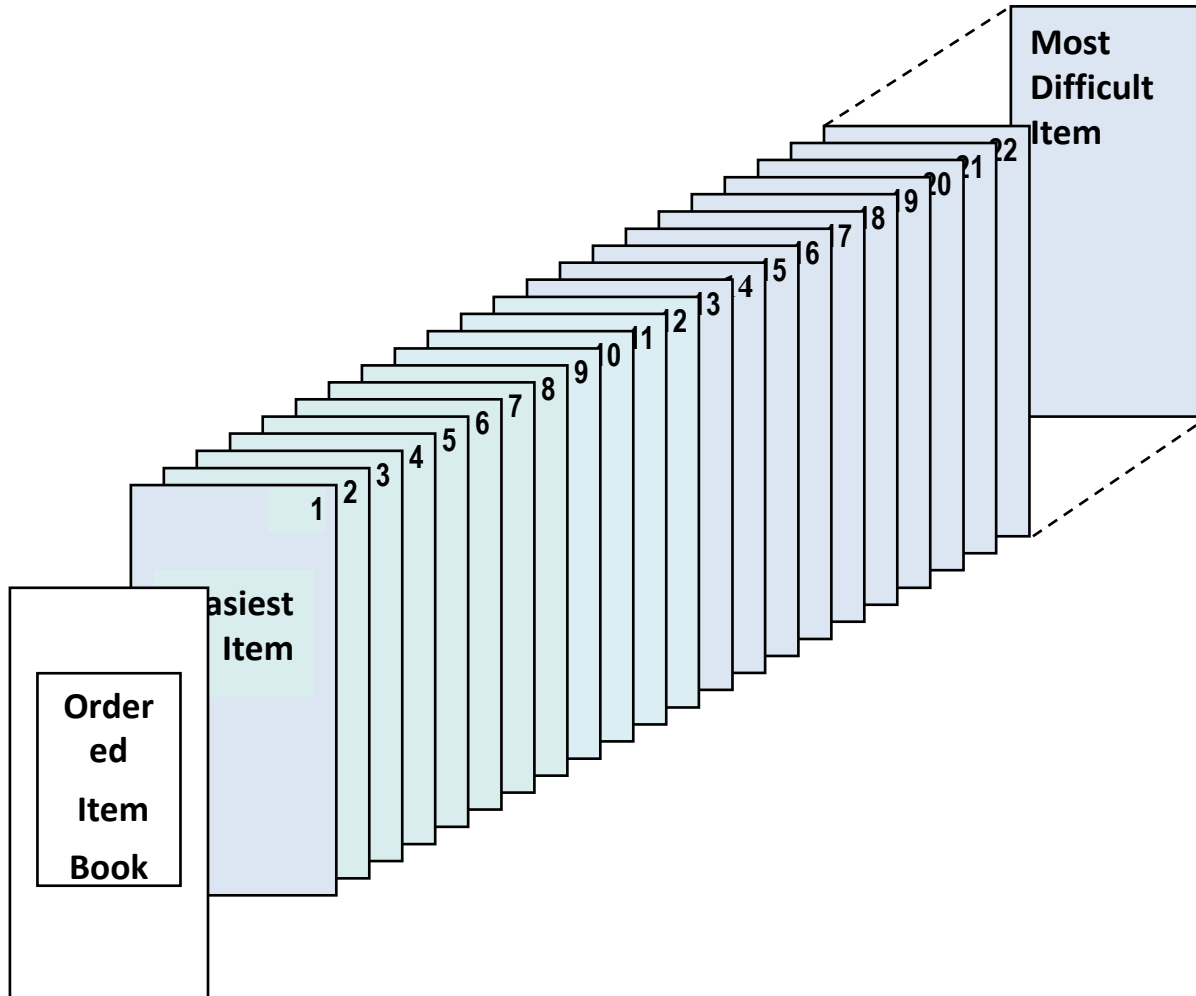
Which PLD most closely matches the knowledge and skill demands for each item?

Hint: Items are ordered by difficulty.

Considerations

- You may judge that an item seems out of order
- There are no right or wrong answers—only your best professional judgments

Threshold regions



OIB Page Number	PLD Match	Explanation
40	A	Item response demands clearly match the expectations in the Advanced PLD
39	A	
38	A	
37	A	
36	A	
35	P	Threshold region
34	A	
33	P	
32	P	
31	A	
30	A	
29	P	Item response demands clearly match the expectations in the Proficient PLD
28	P	
27	P	
26	P	
25	P	
24	B	Threshold region
23	B	
22	P	
21	BB	
20	P	
19	B	Item response demands clearly match the expectations in the Basic PLD
18	B	
17	B	
16	B	
15	B	
14	BB	Threshold region
13	B	
12	BB	
11	B	
10	BB	
9	B	
8	BB	Item response demands do not match the expectations in the Basic PLD
7	BB	
6	BB	
5	BB	
4	BB	
3	BB	
2	BB	
1	BB	

What is the threshold region?

- A sequence of items that match two adjacent PLDs in alternating and inconsistent sequence
- **Note:** If your threshold region is lengthy, go through the items at the top and bottom one more time—see if you can match some items to reduce the length
 - Don't force it; match item RDs to PLD expectations

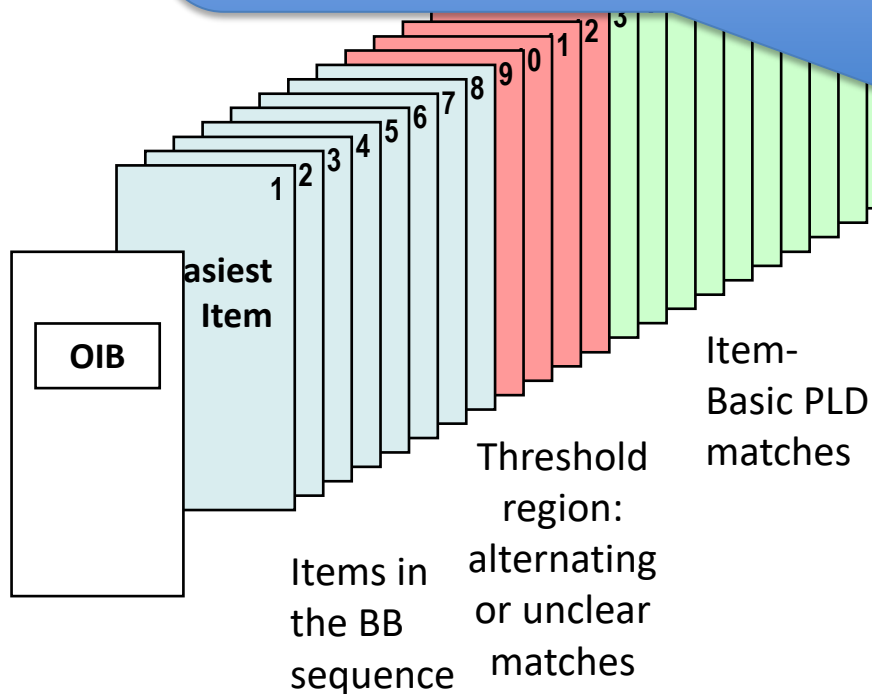
OIB Page Number	PLD Match	Explanation
40	A	Item response demands clearly match the expectations in the Advanced PLD
39	A	
38	A	
37	A	
36	A	
35	P	Threshold region
34	A	
33	P	
32	P	
31	A	
30	A	
29	P	Item response demands clearly match the expectations in the Proficient PLD
28	P	
27	P	
26	P	
25	P	
24	B	Threshold region
23	B	
22	P	
21	BB	
20	P	
19	B	Item response demands clearly match the expectations in the Basic PLD
18	B	
17	B	
16	B	
15	B	
14	BB	Threshold region
13	B	
12	BB	
11	B	
10	BB	
9	B	
8	BB	Item response demands do not match the expectations in the Basic PLD
7	BB	
6	BB	
5	BB	
4	BB	
3	BB	
2	BB	
1	BB	

ID matches and threshold

req.

Which PLD most closely matches the knowledge and skill demands for each item?

Hint: Items above the target cut score, items below the benchmarked cut score



19	B	Item response demands clearly match the expectations in the Basic PLD
18	B	
17	B	
16	B	
15	B	
14	B	
13	B	Threshold region
12	BB	
11	B	
10	BB	
9	B	Item response demands do not match the expectations in the Basic PLD
8	BB	
7	BB	
6	BB	
5	BB	
4	BB	
3	BB	
2	BB	
1	BB	

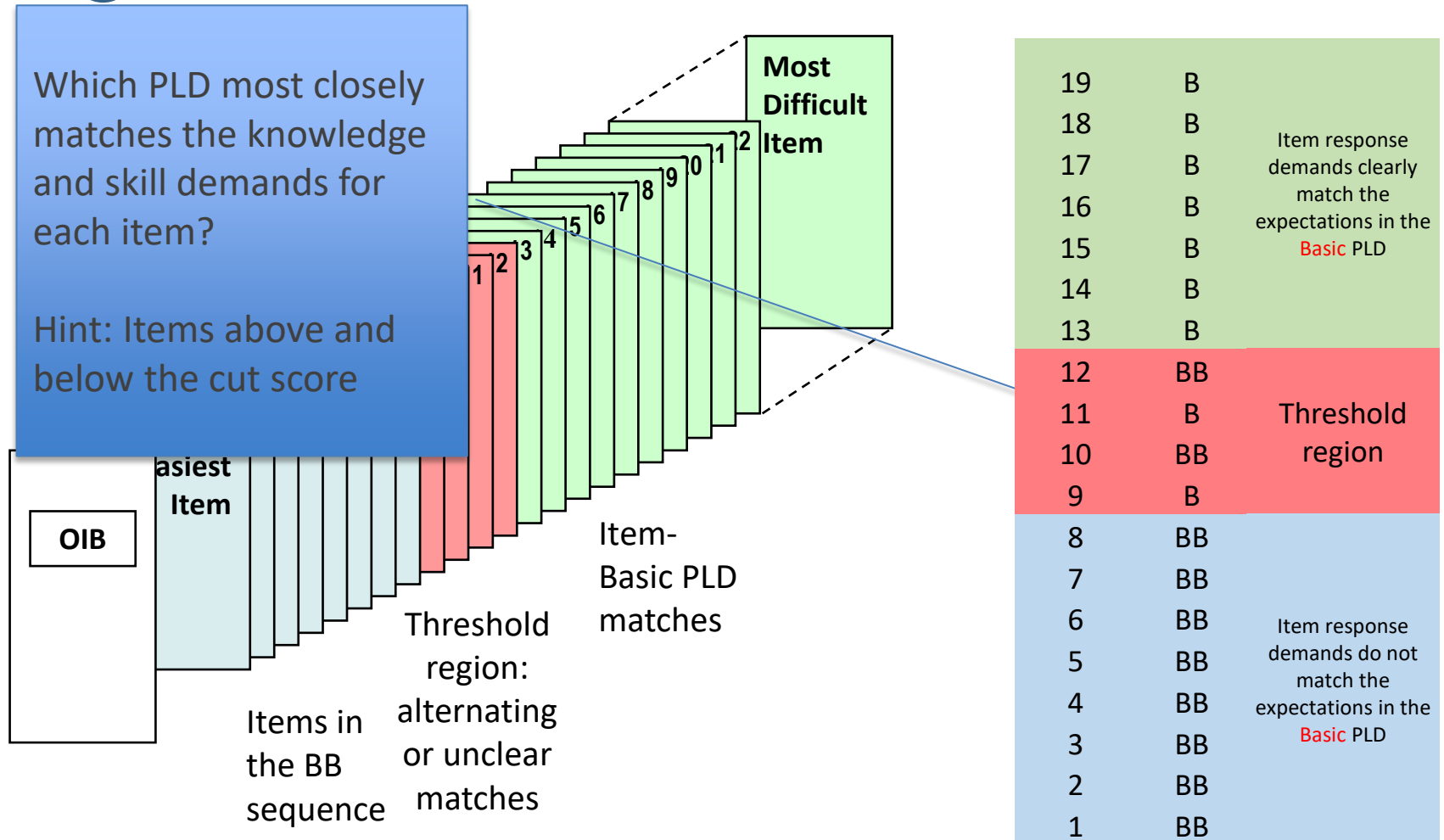
Why do you end up with threshold regions?

- Reasons why panelists put some items in threshold regions
 - The response demands of these items reflect some expectations in the Proficient PLD (for example), and some expectations in the Basic PLD
 - I can't make up my mind yet which PLD this item most closely matches
- **Note:** If your threshold region is lengthy, go through the items at the top and bottom one more time—see if you can match some items to reduce the length
 - Don't force it; match item RDs to PLD expectations

Placing cut scores in threshold regions

- In Round 1, not using the practice items

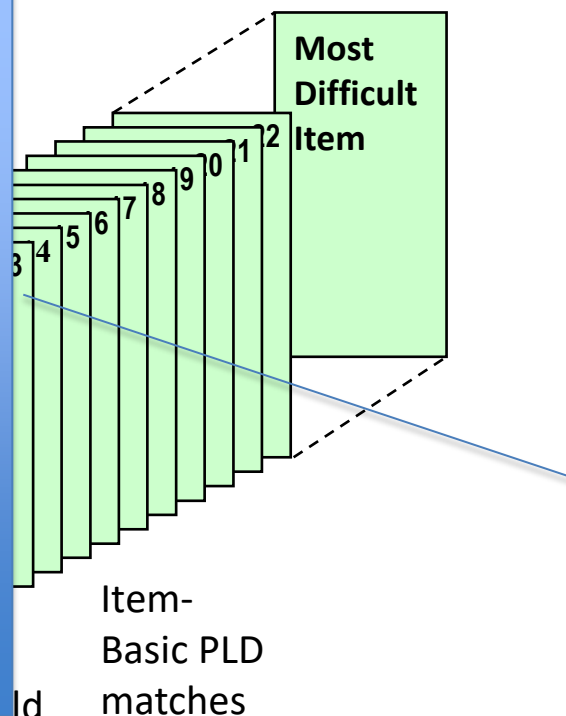
ID matches and threshold regions



ID matches and threshold regions

Do your best to identify the first item in the threshold region that most closely matches the Basic PLD. That is your recommendation for the basic cut score.

That means all of the items on that page and on the pages above are in the Basic region, and all the items below are in the Below Basic region.



19	B	Item response demands clearly match the expectations in the Basic PLD
18	B	
17	B	
16	B	
15	B	
14	B	
13	B	
12	BB	Threshold region
11	B	
10	BB	
9	B	Item response demands do not match the expectations in the Basic PLD
8	BB	
7	BB	
6	BB	
5	BB	
4	BB	
3	BB	
2	BB	
1	BB	

region:
Items in alternating the BB or unclear sequence matches

End of training and practice

- Do you feel ready to prepare for round 1?
- What questions, concerns, etc. remain?
- Table and room discussion

Prepare for round 1: review

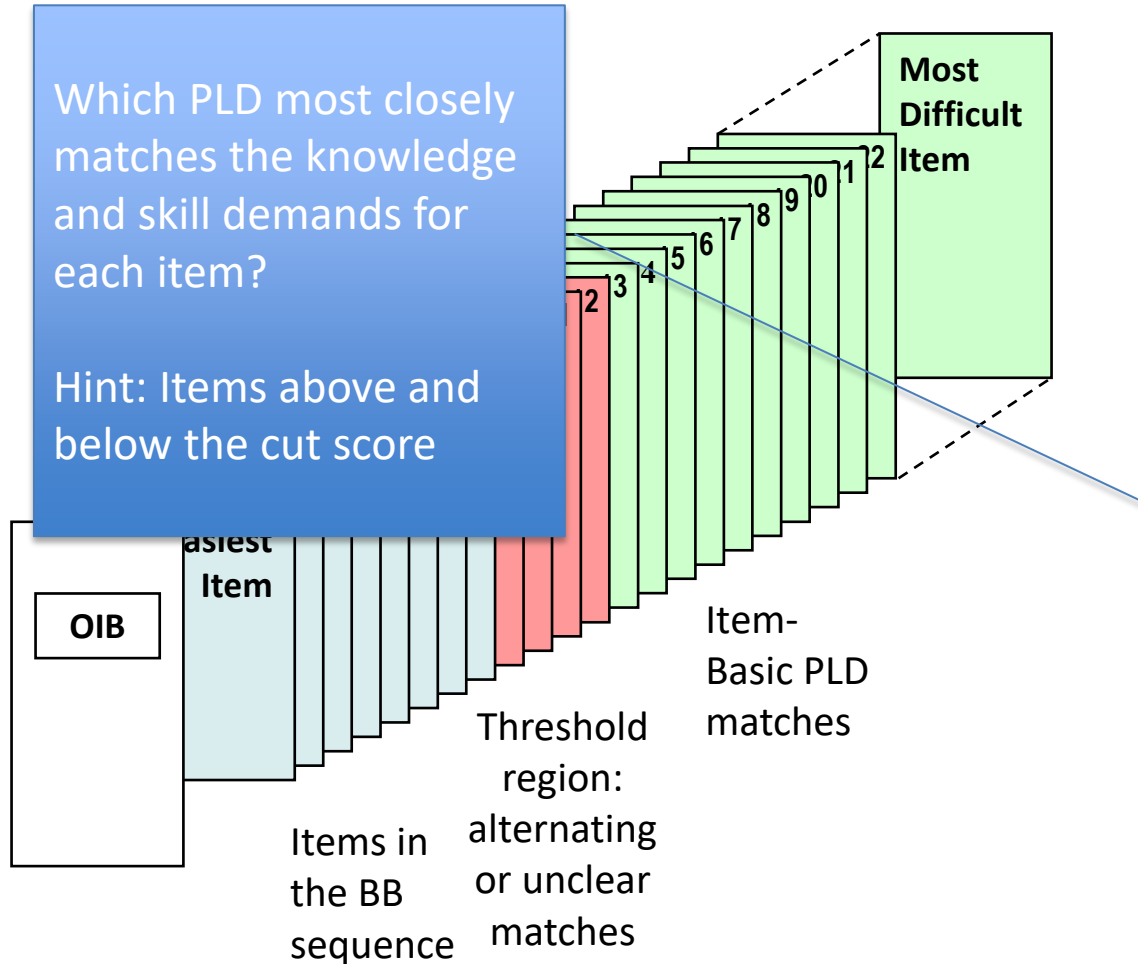
- The ID matching judgmental task
- Place cut scores in threshold regions

ID matching judgmental task

- Step (a) Answer the two questions
- Step (b) Match items to PLDs
- Work independently
- Trust your expertise
- Take notes in the tool

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

ID matches and threshold regions



19	B	Item response demands clearly match the expectations in the Basic PLD
18	B	
17	B	
16	B	
15	B	
14	B	
13	B	
12	BB	Threshold region
11	B	
10	BB	
9	B	Item response demands do not match the expectations in the Basic PLD
8	BB	
7	BB	
6	BB	
5	BB	
4	BB	
3	BB	
2	BB	
1	BB	

Are you ready to undertake round 1?

- Any final questions
- You can ask for more explanation, demonstration of steps, whatever you want
- Discuss with colleagues at your table or pose to the facilitator

Round 1 steps (cont.)

For Each Cut Score

- a) Answer the two questions
 - Start at page 1, finish when you have a clear sequence of items matched to the Advanced PLD
 - Notes on your item map in the tool
- b) Record item-PLD matches
 - Note clear matches and threshold region
 - Notes on your item map in the tool
- c) Place your cut score in the threshold regions
 - Proficient, Advanced, Basic
 - Record in tool

Are you ready to undertake round 1?

- Ask final questions
- Ask for more explanation, demonstration of steps
- Complete the Readiness Survey
 - Open the survey in the tool
- Work independently
- You have up to 120 minutes to complete Round 1

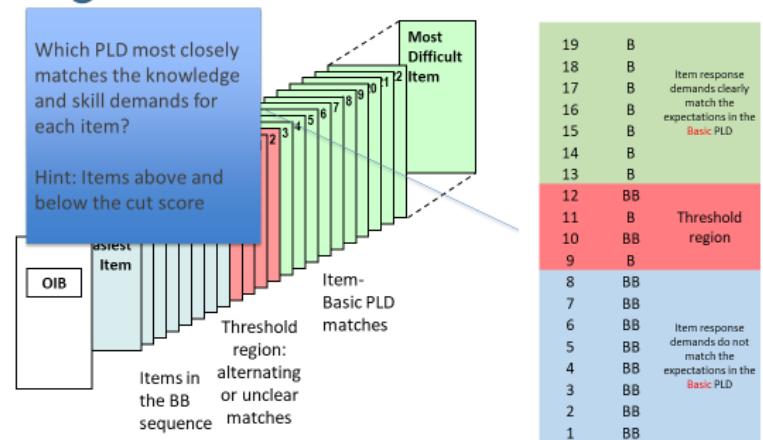
Display while panelists are working

ID matching judgmental task

- Step (a) Answer the two questions
- Step (b) Match items to PLDs
- Work independently
- Trust your expertise
- Take notes in the tool

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

ID matches and threshold regions



Preparation for round 2

Let's prepare for round 2

Review Together

- Cut score feedback from round 1
 - How to think about it as you make cut score decisions in round 2
 - For all cut scores
 - Bar charts in slides

Round 1 feedback

- For your table, for the room
 - Median for the room
 - Each anonymous panelist: highest and lowest OIB page
- Using the feedback
 - Demonstrate reasoning for OIB pages around the recommended cut score
 - Share insights
 - No right or wrong, no persuasion to change
 - Sharing the reasoning for each page is what matters
- Table and room discussion

Concepts to be clear on

- Items are ordered by difficulty
 - We know that panelists in other standard settings think they're ordered by cognitive complexity
- The group recommended cut score is the **average of all of your combined** recommended cut scores
 - Specifically, it's the median, which you can think of as something like the average we use in sports, etc.

Other concepts to be clear on

- Cut score feedback after round 1
 - These numbers are based your and your colleagues' recommended cut scores, from round 1
 - There is nothing about students or item difficulty in this feedback
 - Use this information to see where your recommendation is, compared to your colleagues
 - You do not have to change your recommendation to be closer to your colleagues—use content-based rationales to retain or adjust your own round 1 cut score recommendation when you get to rounds 2 and 3

Table and room discussions

- In all discussions with your colleagues
- Your goals
 - Share your insights
 - Listen to your colleagues' insights
 - Develop sharing understandings amap
 - Support independent decision making
 - Courtesy and respect
- **Not** your goals
 - Agree with your colleagues
 - Persuade your colleagues to agree with you
 - Reach consensus

Are you ready to undertake round 2?

- Ask final questions
- Ask for more explanation, demonstration of steps
- Complete the Readiness Survey
- Work independently
- 90 minutes

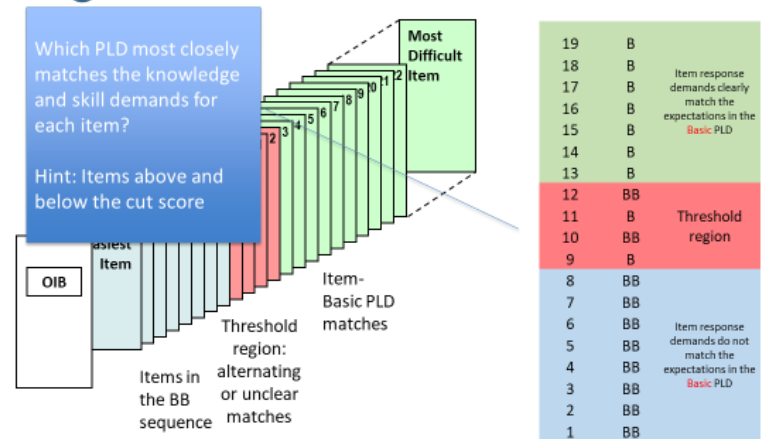
Display while panelists are working

ID matching judgmental task

- Step (a) Answer the two questions
- Step (b) Match items to PLDs
- Work independently
- Trust your expertise
- Take notes in the tool

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

ID matches and threshold regions



Preparation for round 3

Let's prepare for round 3

Two Types of Feedback Review Together

- Cut scores feedback from round 2
 - How to think about it as you make cut score decisions in round 3
 - For all cut scores
- Impact data based on round 2 cut scores

Round 2 cut score feedback

- For your table, for the room
 - Median for the room
 - Each anonymous panelist: highest and lowest OIB page
- Using the feedback
 - Demonstrate reasoning for OIB pages around the recommended cut score
 - Share insights
 - No right or wrong, no persuasion to change
 - Sharing the reasoning for each page is what matters
- Table and room discussion

Table and room discussions

- In all discussions with your colleagues
- Your goals
 - Share your insights
 - Listen to your colleagues' insights
 - Develop sharing understandings as possible
 - Support independent decision making
 - Courtesy and respect
- **Not** your goals
 - Agree with your colleagues
 - Persuade your colleagues to agree with you
 - Reach consensus

Round 2 impact data

- Based on room median recommended cut score
- Using the impact d
- Room discussion

Are you ready to undertake round 3?

- Ask final questions
- Ask for more explanation, demonstration of steps
- Complete the Readiness Survey
- Work independently
- 90 minutes

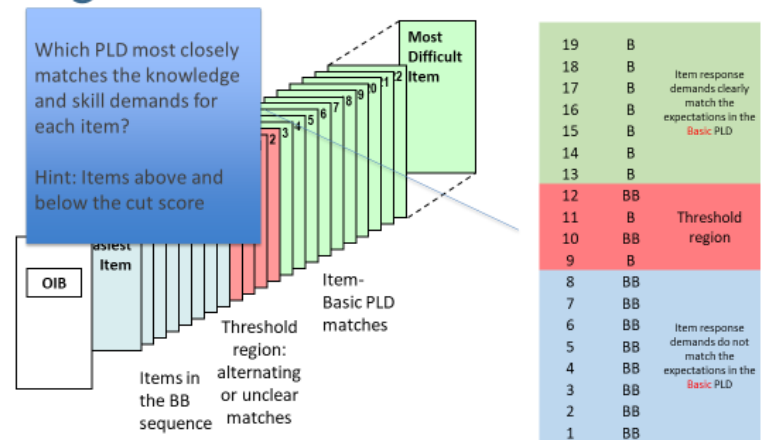
Display while panelists are working

ID matching judgmental task

- Step (a) Answer the two questions
- Step (b) Match items to PLDs
- Work independently
- Trust your expertise
- Take notes in the tool

- (1) What does a student need to know and be able to do in order to respond to this item?
- (2) What makes this item more difficult than the preceding items?

ID matches and threshold regions



Closing session

- Review final results; discussion
- Complete workshop evaluation
- Dismissal

APPENDIX—B
MEETING AGENDA

Oklahoma State Testing Program
CCRA US History Assessment
Standard-Setting Workshop Agenda

Day 1 June 23 Thursday		
8:00-8:30	Check-in and continental breakfast	All panelists
8:30-10:00	Introductions and overview: welcome (OK SDE), workshop goals (TBD), USH exam (OK SDE); standard setting and score reporting, the ID Matching method (Cognia)	All panelists
10:00-10:15	Break	All panelists
10:15-11:30	Training and practice on the ID Matching method: Facilitator models the cognitive-judgmental task, panelists practice, table and workshop discussion Select Table Leaders	All panelists
11:30-12:30 1:15-2:00	Familiarization with the US History assessment: Review range and borderline PLDs, content standards (brief); take the 50-item test; discuss the experience from the student pov	All panelists
12:30-1:15	Lunch	All panelists
2:00-3:00	Prepare for round 1: review IDM judgmental task and borderline PLDs; complete readiness survey	All panelists
2:00-4:00	Complete round 1	All panelists
4:00-4:30	Break and data analysis	All panelists
4:30-4:30	Prepare for round 2: Review round 1 cut score feedback: discuss agreements, disagreements, hypothetical rationales Complete readiness survey Begin round 2 (if time allows)	All panelists

Day 2
June 24 Friday

8:00-8:30	Continental breakfast	All panelists
8:30-9:00	Debrief day 1	All panelists
9:00-10:30	Complete round 2	All panelists
10:30-11:00	Break and data analysis	All panelists
11:00-12:00	Prepare for round 3: Review round 1 cut score feedback: discuss agreements, disagreements, hypothetical rationales; review impact data Complete readiness survey	All panelists
12:00-1:00	Lunch	All panelists
1:00-3:00	Complete round 3	All panelists
3:00-3:30	Break and data analysis	All panelists
3:30-4:30	Review final results; complete workshop evaluation; dismissal	All panelists

APPENDIX—C
NONDISCLOSURE FORM



OKLAHOMA
Education



Nondisclosure Agreement
Oklahoma State Test Program
College and Career Readiness Assessment
US History Standard Setting
June 23-24, 2022

The undersigned is an employee, contractor, assessment committee member, or person otherwise authorized to view secure state assessment materials. The undersigned hereby agrees to be bound to the terms of this agreement restricting the disclosure of said materials.

It is essential to the integrity of this item development project and testing program that all test items remain secure. To maintain this security, only authorized persons are permitted to view the test questions. With the exception of materials released by the Oklahoma State Department of Education for informational purposes, all test questions (draft or final) in hardcopy or electronic format and associated materials must be regarded as secure documents. As a result, such materials may not be reproduced, electronically transmitted, discussed, used in classroom instruction, or in any way released or distributed to unauthorized persons. All materials including items and item drafts must be returned at the end of the meeting.

I understand that I am responsible for test materials security. By breaching test materials security as described here, I am breaching professional testing ethics and may be subject to additional penalties under law.

Name: _____

Signature: _____

Date: _____

APPENDIX—D
PERFORMANCE LEVEL DESCRIPTORS

Oklahoma Performance Level Descriptors (PLDs)

U.S. History

Policy PLDs

Policy PLDs define the knowledge and skill level expectations for the Oklahoma Academic Standards U.S. History (USH).

Advanced

Students demonstrate superior performance on challenging subject matter.

Proficient

Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level.

Basic

Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level.

Below Basic

Students have not performed at least at the Basic level. Students in this range should be given comprehensive U.S. History instruction in order to achieve at the proficient level.

Borderline PLDs

Borderline PLDs describe the knowledge and skills that students within each proficiency level are just barely expected to be able to demonstrate. In line with Oklahoma Academic Standards, the statements combine the subject matter for U.S. History that students are expected to demonstrate.

Advanced

Students at the borderline of the **Advanced** level can demonstrate superior performance on the challenging subject matter through the process of making connections more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the **Advanced** level can do the following more than 50% of the time:

- | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Apply social studies content knowledge in order to make connections and thoroughly understand how the United States developed and changed over time.• Apply social studies content knowledge in order to make connections and thoroughly understand how eras and events throughout United States history have influenced subsequent eras.• Analyze how the “Civil War Amendments,” westward expansion, immigration, and industrialization impacted the development of the United States from 1865 to the 1920s.• Evaluate how the American Industrial Revolution, the growth of populism, and the Progressive Movement transformed the United States from the 1870s to the 1920s.• Analyze the causes and effects of the United States developing into a world power in the late Nineteenth and early Twentieth centuries.• Analyze the social, political, and economic factors that impacted the United States during the 1920s and 1930s. | <ul style="list-style-type: none">• Evaluate how both the outbreak and events of World War II transformed the United States.• Evaluate the social, political, and economic effects the expansion of communism and the Cold War had on the United States from 1945 to 1975.• Analyze how the events and effects of the Civil Rights Movement socially and politically transformed the United States from 1945 to 1975.• Evaluate the major events and presidential policies that affected the United States from 1977 to 2001.• Thoroughly comprehend, interpret, evaluate, and respond to primary sources, political cartoons, maps, photographs, and informational texts, applying critical thinking skills. |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Proficient

Students at the borderline of the **Proficient** level can demonstrate mastery over appropriate subject matter more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the **Proficient** level can do the following more than 50% of the time:

<ul style="list-style-type: none"> • Apply social studies content knowledge in order to make connections and sufficiently understand how the United States developed and changed over time. • Apply social studies content knowledge in order to make connections and sufficiently understand how eras and events throughout United States history have influenced subsequent eras. • Examine how the “Civil War Amendments,” westward expansion, immigration, and industrialization impacted the development of the United States from 1865 to the 1920s. • Examine how the American Industrial Revolution, the growth of populism, and the Progressive Movement transformed the United States from the 1870s to the 1920s. • Summarize the causes and effects of the United States developing into a world power in the late Nineteenth and early Twentieth centuries. • Examine the social, political, and economic factors that impacted the United States during the 1920s and 1930s. 	<ul style="list-style-type: none"> • Summarize how both the outbreak and events of World War II transformed the United States. • Examine the social, political, and economic effects the expansion of communism and the Cold War had on the United States from 1945 to 1975. • Examine how the events and effects of the Civil Rights Movement socially and politically transformed the United States from 1945 to 1975. • Summarize the major events and presidential policies that affected the United States from 1977 to 2001. • Sufficiently comprehend, interpret, evaluate, and respond to primary sources, political cartoons, maps, photographs, and informational texts, applying critical thinking skills.
<p>Basic</p> <p>Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills of the appropriate subject matter more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:</p>	

<ul style="list-style-type: none"> • Inconsistently apply social studies content knowledge in order to make connections and partially understand how the United States developed and changed over time. • Inconsistently apply social studies content knowledge in order to make connections and partially understand how eras and events throughout United States history have influenced subsequent eras. • Partially examine how the “Civil War Amendments,” westward expansion, immigration, and industrialization impacted the development of the United States from 1865 to the 1920s. Partially examine how the American Industrial Revolution, the growth of populism, and the Progressive Movement transformed the United States from the 1870s to the 1920s. • Ineffectively describe the causes and effects of the United States developing into a world power in the late Nineteenth and early Twentieth centuries. • Inconsistently identify the social, political, and economic factors that impacted the United States during the 1920s and 1930s. 	<ul style="list-style-type: none"> • Partially examine how both the outbreak and events of World War II transformed the United States. • Inconsistently identify the social, political, and economic effects the expansion of communism and the Cold War had on the United States from 1945 to 1975. • Partially examine how the events and effects of the Civil Rights Movement socially and politically transformed the United States from 1945 to 1975. • Inconsistently identify the major events and presidential policies that affected the United States from 1977 to 2001. • Partially comprehend, interpret, evaluate, and respond to primary sources, political cartoons, maps, photographs, and informational texts, applying critical thinking skills.
<p>Below Basic Students have not performed at least at the Basic level.</p>	

APPENDIX E
COGNIA STANDARD-SETTING TOOLKIT

Cognia Standard-Setting Toolkit

This appendix contains sample screenshots of the Cognia Standard-Setting Toolkit that panelists used for all standard-setting activities during the meeting. Images provided correspond to sample (1) login screen, (2) practice item booklet, (3) readiness survey screen, (4) ordered item booklet view, (5) item view, and (6) completion survey. A brief description accompanies each image.

Figure 1. Sample Login Screen


Panelists are provided with usernames and password to enable secure access to the toolkit



The screenshot displays the login interface for the Cognia Standard Setting Toolkit. At the top left, the Cognia logo is shown with the text "Standard Setting Toolkit" underneath. Below this, the word "Login" is prominently displayed. The login form consists of two input fields: "Name" and "Password", each with a corresponding text box. A blue "Log In" button is positioned below the password field. At the bottom left of the page, the copyright notice "© 2022 - Cognia" is visible.

Figure 2. Sample Practice Item Booklet

This image shows a list of sample practice items as a truncated item map view. Panelists use the practice item booklet during the practice round to become familiar with use of the tool and to practice the ID matching process.




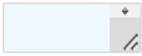
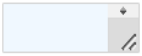
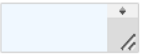
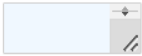
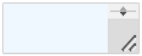
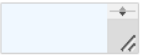
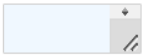
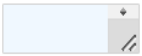
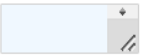



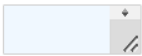
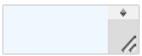
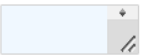


User 1

Program Name Subject, Grade

Practice Round


[Advance To Next Step When Prompted](#)

Position	Description	Relevant KSAs	Why More Difficult Than Previous	Rationale	Item Descriptor Match	Cut Placement	
34	Item 34 Description				-- ▾	-- ▾	View Item
26	Item 26 Description				-- ▾	-- ▾	View Item
22	Item 22 Description				-- ▾	-- ▾	View Item
15	Item 15 Description				-- ▾	-- ▾	View Item
12	Item 12 Description				-- ▾	-- ▾	View Item
3	Item 3 Description				-- ▾	-- ▾	View Item

[Advance To Next Step When Prompted](#)

Figure 3. Sample Readiness Survey

Before each round of judgements, panelists complete a readiness survey to indicate whether they are ready to undertake the associated judgement round.



cognia™
Standard Setting Toolkit

User 1

Program Name Subject, Grade

Round 1 Readiness

[Advance To Next Step When Prompted](#)


Question	Response
I understand how to use my expert judgment to answer the two questions about each item	--
I understand how to use my expert judgment to match each item to a PLD	-- Yes No
I understand how and why items appear in threshold regions	--
I understand how to use my expert judgment to place Proficient, Advanced, and Nearing Proficiency cut scores	--
I may not feel completely comfortable, but I am ready to undertake round 1	--

[Advance To Next Step When Prompted](#)

© 2022 - Cognia

Figure 4. Sample Ordered Item Map View (truncated)

This image shows a sample view of the item map as displayed on panelists' screens. As a reminder, the item list is ordered from easiest (at the bottom) to most difficult (at the top).


Standard Setting Toolkit

User 1

Program Name Subject, Grade


Grade X Judgments

[Advance To Next Step When Prompted](#)

Position	Description	Relevant KSAs	Why More Difficult Than Previous	Rationale	Item Descriptor Match	Cut Placement	
44	Item 44 Description	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	-- <div></div>	-- <div></div>	View Item
43	Item 43 Description	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	-- <div></div>	<div>--<div></div><div>Performance Level 2</div><div>Performance Level 3</div><div>Performance Level 4</div></div>	View Item
42	Item 42 Description	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div></div>	-- <div></div>	-- <div></div>	View Item

Figure 5. Sample Ordered Item Booklet Page View


The ordered item booklet view displays each item as a digital page in the booklet along with links to any associated stimuli and/or rubrics. In addition, notes below the item provide the item description, the associated standard, and (when relevant) notes about possible score points for the item. Panelists used the navigational arrows to move 'up' and 'down' pages in the booklet.



Standard Setting Toolkit

Program Name Grade Subject - Grade X Judgments - User 1

[Return to Item Booklet](#)

Item 44  [View Stimulus](#)

PDF OF ITEM
DISPLAYED HERE

- **Description:** Item 44 Description
- **Standard:** Sample L.OX.04.b: Text describing the associated standard appears in this section for easy reference.

Relevant KSAs

Figure 6. Sample Judgement Form

The judgement form provides space for panelists to write notes about (1) the relevant knowledge, skills, and abilities (KSAs) needed to respond to the item, (2) why the item is more difficult than the previous item, and (3) content-based rationales. In addition, dropdown menus are provided for the item descriptor matches and the cut placements. Note that the judgement form can be accessed through both the booklet view and the item map view.

Relevant KSAs

Why More Difficult Than Previous

Rationale

Item Descriptor Match

--


Cut Placement

--

[Return to Item Booklet](#)

Figure 7. Sample Completion Survey (truncated)

This image provides a truncated view of the completion survey provided to panelists at the end of the standard-setting meeting to collect their final evaluations and feedback on various aspects of the meeting.


Standard Setting Toolkit

User 1

Program Name Subject, Grade

Completion Survey

[Advance To Next Step When Prompted](#)

Question	Response
I understood the goals of the standard setting workshop.	--
I understood the procedures we followed to recommend standards.	--
I understood that my role was to recommend cut scores to the State Department of Education.	Strongly Disagree Disagree Undecided Agree Strongly Agree Not Applicable
The workshop procedures made sense to me, and I learned how to apply them efficiently.	
I am confident about my understanding of this standard setting process.	--
The workshop facilitator explained things clearly to us.	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--
The workshop facilitator provided clear and helpful responses to my	

APPENDIX—F
READINESS SURVEYS

Readiness Surveys

Round 1 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I understand how to use my expert judgment to answer the two questions about each item		
I understand how to use my expert judgment to match each item to a PLD		
I understand how and why items appear in threshold regions		
I understand how to use my expert judgment to place Basic, Proficient, and Advanced cut scores		
I may not feel completely comfortable, but I am ready to undertake round 1		

Round 2 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I know that feedback and discussion in preparation for round 2 will help me feel even more comfortable		
I understand the round 1 feedback about (a) our group cut scores for Basic, Proficient, and Advanced, and (b) the highest and lowest panelist cut scores for each level		
I understand the ground rules for discussing feedback in preparation for round 2: sharing information, avoiding persuasion		
I understand that I should use the round 1 feedback as information, not persuasion, for me to consider as I place my cut scores in round 2		
I'm ready to undertake round 2		

Round 3 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I know that feedback and discussion in preparation for round 3 will help me feel even more comfortable		
I understand the round 2 feedback about (a) our group cut scores for Basic, Proficient, and Advanced, and (b) the highest and lowest panelist cut scores for each level		
I understand the ground rules for discussing feedback in preparation for round 3: sharing information, avoiding persuasion		
I understand that I should use the round 2 feedback as information, not persuasion, for me to consider as I place my cut scores in round 3		
I'm ready to undertake round 3		

APPENDIX—G
WORKSHOP EVALUATION SURVEY

OK CCRA US History Standard Setting

Final Workshop Evaluation

Please respond to the items below to provide your feedback on the training we provided so that you could recommend cut scores following the ID Matching process.

Your feedback is anonymous. We will summarize feedback for all panelists and use the summary for the standard-setting final technical report.

The Workshop Overall

1. I understood the goals of the standard-setting workshop.
2. I understood the procedures we followed to recommend standards.
3. I understood that my role was to recommend cut scores to the Oklahoma State Department of Education.
4. The workshop procedures made sense to me, and I learned how to apply them efficiently.
5. I am confident about my understanding of this standard-setting process.

Workshop Facilitation

6. The workshop facilitator explained things clearly to us.
7. The workshop facilitator encouraged us to raise questions and put our understandings into our own words.
8. The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.
9. The workshop facilitator took steps to help the standard setting process run smoothly.

Training, Practice, and the Standard Setting Workshop Process

10. Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.
11. I understood the progressions in expectations across the Basic, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.
12. I became sufficiently familiar with the CCRA US History assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.
13. I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores in threshold regions.

The Standard Setting Tool

I understood how to use the standard-setting tool to

14. Record my responses to the two questions about each item I reviewed
15. Record my recommended cut scores.
16. Record other notes

Threshold Regions

17. I understood why I had threshold regions and how to place a cut score in those regions in round 1 of the workshop

Feedback After Round 1, Preparation for Round 2

18. I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores. I understood how to use the feedback after round 1 on the group

recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.

Final Cut Scores

19. I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.
20. If yes, would you recommend moving a cut score up or down in the OIB, and by how many pages?

Optional Open-ended Comments

21. Please indicate any parts of the standard-setting training and process that we should improve.
22. Please indicate any concerns you may have about the workshop process and the final recommended cut scores.

Main Sections of the Standard-Setting Workshop

Please rate the usefulness of each section:

Usefulness	1 Not at all useful	2	3	4	5 Extremely useful
The opening session	-	-	-	-	-
Working together at my table	-	-	-	-	-
Parsing the ALDs	-	-	-	-	-
Answering the two questions about each item	-	-	-	-	-
Table-level discussions	-	-	-	-	-
Cross-table discussions	-	-	-	-	-

Please rate the influence of the following when setting standards:

Usefulness	1 Not at all useful	2	3	4	5 Extremely useful
ALDs: Overall descriptors	-	-	-	-	-
ALDs: overall bulleted descriptors	-	-	-	-	-
My answers to the two questions about each item	-	-	-	-	-
My judgements about match of items to ALDs	-	-	-	-	-
My experience working with students	-	-	-	-	-

What materials, information, or procedures were most influential in your placement of the cut scores? In what ways?

Finally

Please provide any additional comments you would like us to consider.

Please provide any other recommendations that could help us improve future standard setting workshops.

Thanks for participating in this workshop and completing the evaluation.

Safe and easy travels!

APPENDIX—H
PANELISTS

2022 Oklahoma U.S. History Standard-Setting Participant List

Last Name	First Name	Current Position/Title	School/Site Name	District
Butler	Jennifer	Teacher	Edmond North High School	Edmond Public Schools
Dormiani	Angela	High School History Teacher	ASTEC Charter School	ASTEC Charter Schools
Doudican	Kevin “Mike”	Teacher	Glenpool High school	Glenpool
Frazier	Stephen	District Social Studies Department Chair, HS History Teacher	Dove Science Academy Tulsa HS	Dove Schools of Tulsa
Lamkin	Jennifer	U.S. History Teacher	Tulsa School of Arts and Sciences	TPS-Public Charter
Purcell	Jane	Social Studies Coordinator	Curriculum Center	Norman Public Schools
Walden	Stephen	AP/Standard US History Teacher	Tahlequah High School	Tahlequah Public Schools
Mosqueda	Stephanie	US History, APUSH Teacher	Elk City HS	Elk City
Tillotson	Heather	Teacher	Wagoner HS	Wagoner Public Schools
Thom	David	Teacher	Memorial HS	Tulsa Public Schools
Stewart	Arletta			Cache Public Schools

APPENDIX—I
STANDARD-SETTING ROUND RESULTS

OK CCRA USH Standard-Setting Round Results

Table 1. OK CCRA USH Standard-Setting Round 1 Results

Performance Level	OIB Page #	Raw Score	Theta (Median)	Median Abs. Diff.	Percent Students
Below Basic	--	--	--	--	39.9
Basic	6	22	-0.26	0.18	20.0
Proficient	17	28	0.33	0.59	30.2
Advanced	41	42	1.30	0.11	9.9
Proficient + Advanced	--	--	--	--	40.1

Table 2. OK CCRA USH Standard-Setting Round 2 Results

Performance Level	OIB Page #	Raw Score	Theta (Median)	Median Abs. Diff.	Percent Students
Below Basic	--	--	--	--	39.9
Basic	6	22	-0.26	0.18	23.0
Proficient	18	28	0.37	0.62	27.2
Advanced	41	42	1.30	0.00	9.9
Proficient + Advanced	--	--	--	--	37.1

Table 3. OK CCRA USH Standard-Setting Round 3 Results

Performance Level	OIB Page #	Raw Score	Theta (Median)	Median Abs. Diff.	Percent Students
Below Basic	--	--	--	--	39.9
Basic	6	22	-0.26	0.16	13.9
Proficient	14	25	0.14	0.40	36.3
Advanced	41	42	1.30	0.00	9.9
Proficient + Advanced	--	--	--	--	46.2

APPENDIX—J
WORKSHOP EVALUATION RESULTS

OK CCRA USH Standard-Setting Workshop Evaluation Results

Table 1. Frequency of Evaluation Responses (N = 11)

	Yes	No	No Response
I understood the goals of the standard-setting workshop.	11	--	--
I understood the procedures we followed to recommend standards.	11	--	--
I understood that my role was to recommend cut scores to the State Department of Education.	11	--	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	10	--	1
I am confident about my understanding of this standard-setting process.	11	--	--
The workshop facilitator explained things clearly to us.	11	--	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	11	--	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	11	--	--
The workshop facilitator took steps to help the standard-setting process run smoothly.	11	--	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	10	1	--
I understood the progressions in expectations across the Basic, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	10	1	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	11	--	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	11	--	--
I understood how to use the standard-setting tool to record my responses regarding skills and item difficulties as instructed.	11	--	--
I understood how to use the standard-setting tool to record my recommended cut scores.	11	--	--
I understood how to use the standard-setting tool to record other notes.	11	--	--
I understood why I had threshold regions and how to place a cut score in those regions in round 1 of the workshop	11	--	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores. I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	11	--	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	11	--	--

Table 2. Open-ended responses

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	"3 days instead of 2. More static schedule"
	"Half days over more time. The work was heavy big brain thinking. Otherwise, it was fine."
	"A little more time for practice would've been nice. More practice on sample items for notetaking on KSAs"
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	"Scaffolding was great, but it took forever to get started. Too much seeking confirmation of understanding. Everyone was kind and patient! The hotel was nice. Still not sure the operating assumptions on which questions are automatically "basic" region v. "advanced" region (bottom/top list) are correct or helpful"
	"I think there was too much opportunities for questions rather than giving us some time to practice with sample questions, then being able to ask questions. Its hard to ask questions when you haven't interacted with material prior to round1. thank you. you were super nice and pleasant to work with"
	"Excellent team facilitating this workshop!"
What materials, information, or procedures were most influential in your placement of the cut scores? In what ways?	"The discussion helped tremendously. When others explain their reasoning, it helps me see what I missed"
	"PLDs gave guidance"
	"Listing all of my answers and getting an overview"
	"PLDs, seeing data following each round"
	"All of it was very helpful in keeping with the process of placement of the cut scores. It help knowing what was expected for each level"
	"The PLDs were moderately influential; however, discussion with colleagues was most"
	"2 questions. Discussion"
	"the graph slides. Visualizing the data helped me narrow it down"
	"Q&A based on the shared experiences of teachers in the room helped me contextualize my decisions"
	"The questions and ALDs"
	"PLDs & bulleted descriptors helped me to define what student performance should look lik. Discussions with table + whole group helped clarify transition points between levels"

Table 3. Frequency of Responses to Rating Scale Questions

	1 Not at all useful	2	3	4	5 Extremely useful
Please rate the usefulness of each section:					
The opening session	--	--	--	3	8
Working together at my table	--	--	2	1	8
Parsing the ALDs	--	--	--	3	8
Answering the two questions about each item	--	--	--	2	9
Table-level discussions	--	--	1	2	8
Cross-table discussions	--	--	--	1	10
Please rate the influence of the following when setting standards:					
ALDs: Overall descriptors	--	--	1	4	6
ALDs: overall bulleted descriptors	--	--	1	4	6
My answers to the two questions about each item	--	--	--	3	8
My judgements about match of items to ALDs	--	--	1	3	7
My experience working with students	--	--	--	2	9

Table 4. Frequency of Evaluation Responses for MSSA ELA Grade 6–8 (N = 1--)

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	--	--	--	2	8	--
I understood the procedures we followed to recommend standards.	--	--	--	4	6	--
I understood that my role was to recommend cut scores to the State Department of Education.	--	--	--	2	8	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	--	--	1	4	5	--
I am confident about my understanding of this standard-setting process.	--	--	--	3	7	--
The workshop facilitator explained things clearly to us.	--	--	--	3	7	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	1	9	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	--	1	9	--
The workshop facilitator took steps to help the standard-setting process run smoothly.	--	--	--	3	7	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	--	1	4	3	2	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	--	5	5	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	--	--	5	5	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	--	--	--	3	7	--
I understood how to use the standard-setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	2	8	--
I understood how to use the standard-setting tool to record my recommended cut scores.	--	--	--	2	8	--
I understood that I could recommend retaining or adjusting the target cut scores.	--	--	--	2	8	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	--	7	3	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	5	4	1
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	--	4	6	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	--	--	--	3	2	2
	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	1	1	--	4

Table 5. Open-ended responses for MSSA Grade 6–8 ELA

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	I feel that there should have been one more day.
	"It might have been beneficial to have one more day to work on the panel. Completing Round 1 is going to take more time because it takes a little bit to get into the swing of things. I feel honored to be on the panel and enjoy the process; I just felt a little rushed."
	Introduce the PLD's in more detail. Reduce or eliminate the repetitive discussion about processes. Increase work time by reducing breakfast and lunch. Add half a day.
	Overall the process was exciting and interesting. I feel like we needed more time to read the content prior to making cut scores--an hour for 44 questions just isn't enough. I feel like the presentation during breakfast the first day wasn't necessary--that information was given in content meetings.
	technology!
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	More time needed to complete this panel discussion
	Taking the 8th grade test was extremely beneficial because it gave us time to read the passages. We ran out of time and the decision was made to not take the 7th and 6th grade tests. I did not feel as familiar with the 7th and 6th grade tests and passages as I did with 8th grade.
	I have no concerns at this point. It was a fantastic experience and I appreciate the opportunity to participate.

Table 6. Frequency of Evaluation Responses for MSSA Mathematics Grade 3–5 (N = 9)

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	--	--	--	3	6	--
I understood the procedures we followed to recommend standards.	--	--	--	2	7	--
I understood that my role was to recommend cut scores to the State Department of Education.		--	--	2	7	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.		--	--	3	6	--
I am confident about my understanding of this standard-setting process.	--	--	--	2	7	--
The workshop facilitator explained things clearly to us.	--	--	--	2	7	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	2	7	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	--	3	5	1
The workshop facilitator took steps to help the standard-setting process run smoothly.	--	--	--	3	6	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	--	--	--	3	6	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	--	3	6	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	--	--	3	6	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	--	--	--	3	6	--
I understood how to use the standard-setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	3	6	--
I understood how to use the standard-setting tool to record my recommended cut scores.	--	--	--	3	6	--
continued						

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood that I could recommend retaining or adjusting the target cut scores.	--	--	--	1	8	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	--	4	5	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	2	7	--
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	--	2	7	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	--	--	--	5	4	--
	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	--	--	--	7

Table 7. Open-ended responses for MSSA Mathematics Grade 3–5

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	Better ventilated and cooler room. It was bit hot to work in.
	The process would be more practical as a 3-day meeting, rather than 2.5 days.
	I felt well trained.
	None, the process was smooth and clearly understandable
	The training process was ideal.
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	none difficult process but became easier as I went along
	--

Table 8. Frequency of Evaluation Responses for MSSA Mathematics Grade 6–8 (N = 1--)

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	1	--	--	2	6	--
I understood the procedures we followed to recommend standards.	1	--	--	3	5	--
I understood that my role was to recommend cut scores to the State Department of Education.	--	--	--	1	8	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	1	--	1	2	4	1
I am confident about my understanding of this standard-setting process.	--	--	--	4	5	--
The workshop facilitator explained things clearly to us.	1	--	--	2	6	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	1	8	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	1	--	1	7	--
The workshop facilitator took steps to help the standard-setting process run smoothly.	--	1	--	2	6	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	--	--	2	3	4	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	--	2	7	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	1	--	2	6	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	1	--	--	3	5	--
I understood how to use the standard-setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	3	6	--
I understood how to use the standard-setting tool to record my recommended cut scores.	--	--	--	2	7	--
I understood that I could recommend retaining or adjusting the target cut scores.	1	--	--	3	5	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	1	4	4	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	5	3	1
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	1	3	5	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	1	--	--	4	3	--
	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	1	--	--	5

Table 9. Open-ended responses for MSSA Mathematics Grade 6–8

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	It's good
	"We never once talked about allowing us as participants to recommend retaining or adjusting the target cut scores. (See 6th understand question above this one). With regard to the question just above this -- there were some cut scores that should have been moved up at least 2 pages and others that should have been moved down at least 2 pages -- the answer is dependent on the grade level AND on the PDL cut."
	let get started with work sooner
	I would recommend more collaboration between the panelist during the round 1 process. The working independently was a good process but being able to collaborate would allow for a good experience between educators and allow for more experiences to allow better understanding of content that may not be understood fully.
	It started off a little bit slow and then I felt rushed at the end. Maybe time management or hands on practice with individual help as needed rather than just explaining the process.
	Providing the Answers and a calculator to move through the process
	"Provide calculator sheets and calculators (or asks educators to bring) More time to complete tasks-felt rushed sometimes Every task should be completed for accuracy with cut scores Allow for 5 -1-- minute breaks every two ours- mentally draining Copy of math practices Very hot in rooms made it very difficult to stay focused"
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	"Concerned because (1) at least 2 participants complained in the hallway numerous times that ""i don't care what the pdl's say as my students can't do this"", (2) at least 1 person has never taught math and has no background in math, (3) several people believed the proficient HAD to be in the green area and move not move from there, (4) at least one person at my table never understood the threshold portion of rating. An additional tech person is highly needed as there was quite a bit of wasted time waiting for tech issues. People constantly coming in and out was very distracting, and actually seems like a breach in security (especially when ""outside"" people came into the room)."
	let us get started with work sooner
	I do not have any concerns. I thought that the workshop was conducted well and that the gentlemen and ladies who were in charge of the workshop did a great job of handling the workshop.
	Need to shorten the initial training to 1/2 day, so committee can start working on day one not day two

Table 10. Frequency of Evaluation Responses for ASR Science Grade 5 (N = 1--)

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	--	--	--	--	1--	--
I understood the procedures we followed to recommend standards.	--	--	--	1	9	--
I understood that my role was to recommend cut scores to the State Department of Education.	--	--	--	--	1--	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	--	--	--	3	7	--
I am confident about my understanding of this standard-setting process.	--	--	1	2	7	--
The workshop facilitator explained things clearly to us.	--	--	--	1	9	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	--	1--	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	--	1	9	--
The workshop facilitator took steps to help the standard setting process run smoothly.	--	--	--	--	1--	--
Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	1	--	--	--	9	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	--	2	8	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	--	--	2	8	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	--	--	--	3	7	--
I understood how to use the standard-setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	2	8	--
I understood how to use the standard-setting tool to record my recommended cut scores.	--	--	--	2	8	--
I understood that I could recommend retaining or adjusting the target cut scores.	--	--	--	1	9	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	1	3	6	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	2	8	--
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	--	3	7	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	--	--	1	1	8	--
	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	4	--	1	4

Table 11. Open-ended responses for ASR Science Grade 5

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	"Facilitator was clear and thorough. More in-depth screening/application process for the panelists - I feel some people were here just for a vacation at the Sheraton rather than being passionate about the assessment process and success of our children."
	I felt that the training and process would have been solid with 2 days instead of 3
	I feel it was well organized and presented. No improvements are recommended.
	I think you all were clear and careful about your procedures.
	I thought that everything was explained well, and plenty of time to ask questions, and re explained.
	Everything was, hands on training is more understandable for me, but by asking questions I was able to accomplish
	A hands-on example or two about the process would be helpful.
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	The workshop was very well organized and structured - this is my second event and pleased overall with the professionalism of Cognia and the presence of the PED in this process.
	No concerns.
	You did a great job with some challenging material.

Table 12. Frequency of Evaluation Responses for ASR Science Grade 8 (N = 11)

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	--	--	--	2	9	--
I understood the procedures we followed to recommend standards.	--	--	--	3	8	--
I understood that my role was to recommend cut scores to the State Department of Education.	--	--	--	3	8	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	--	--	--	4	7	--
I am confident about my understanding of this standard-setting process.	--	--	--	3	8	--
The workshop facilitator explained things clearly to us.	--	--	--	2	9	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	--	11	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	--	1	1--	--
The workshop facilitator took steps to help the standard-setting process run smoothly.	--	--	--	--	11	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	--	--	1	2	8	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	1	4	6	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	--	--	5	6	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	--	--	--	6	5	--
I understood how to use the standard setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	4	7	--
I understood how to use the standard-setting tool to record my recommended cut scores.	--	--	--	4	7	--
continued						

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood that I could recommend retaining or adjusting the target cut scores.	--	--	--	3	8	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	--	6	5	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	4	7	--
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	--	5	5	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	--	--	--	5	5	--
	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	1	--	--	3

Table 13. Open-ended responses for ASR Science Grade 8

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	<p>"Better organized PLD sheet we use as a guide ex. Life Science LS4 have Advanced, Proficient and Nearing Proficiency standards on one page. When returning to the Booklet have the page that you return to be where you left off and not back to the top of the booklet page. This will decrease the amount of scrolling needed. "</p> <p>Everything worked!!!!</p> <p>I would like to see the standards that move from NP to P to A be separated (maybe bulleted) by the topics covered.</p> <p>The room temperature. perhaps practicing too with novice questions</p> <p>The PLD's should be organized by standard.</p> <p>"More user-friendly standards pbls more examples to familiarize with the tool sample rationales"</p> <p>having PLDs separated by standard (PS, LS, ESS, etc)</p> <p>all of it was really good</p> <p>air conditioning</p>
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	--

Table 14. Frequency of Evaluation Responses for ASR Science Grade 11 (N = 9)

Response	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I understood the goals of the standard-setting workshop.	--	--	--	2	7	--
I understood the procedures we followed to recommend standards.	--	--	--	2	7	--
I understood that my role was to recommend cut scores to the State Department of Education.	--	--	--	1	8	--
The workshop procedures made sense to me, and I learned how to apply them efficiently.	--	--	--	4	5	--
I am confident about my understanding of this standard-setting process.	--	--	--	4	5	--
The workshop facilitator explained things clearly to us.	--	--	--	2	7	--
The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	2	7	--
The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	--	2	7	--
The workshop facilitator took steps to help the standard-setting process run smoothly.	--	--	--	1	8	--
Sufficient time was allotted for training and practice on the standard-setting concepts, tasks, and procedures.	--	--	--	2	7	--
I understood the progressions in expectations across the Nearing Proficiency, Proficient, and Advanced levels as defined by the borderline Performance Level Descriptors.	--	--	--	5	4	--
I became sufficiently familiar with the assessment to recommend cut scores, based on responding to items on the test and answering the two questions about items.	--	--	--	3	6	--
I understood the ID Matching task, including answering the two questions about each item, matching those item response demands to PLDs, and how to place cut scores.	--	--	1	2	6	--
I understood how to use the standard setting tool to record my responses regarding skills and item difficulties as instructed.	--	--	--	2	7	--
I understood how to use the standard setting tool to record my recommended cut scores.	--	--	--	2	7	--
I understood that I could recommend retaining or adjusting the target cut scores.	--	--	--	1	8	--
I understood how to write content-based rationales for my cut score recommendations.	--	--	--	4	5	--
I understood that the group recommended cut score was the average (i.e., the median) of all 13 recommended cut scores.	--	--	--	2	7	--
I understood how to use the feedback after round 1 on the group recommended cut score and the individual panelist highest and lowest cut scores, in preparation for round 2.	--	--	--	2	7	--
I am satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.	--	--	--	1	8	--
Response	Up 2 Pages	Up 1 Page	Do Not Move	Down 1 Page	Down 2 Pages	Not Applicable
If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?	--	--	1	--	--	2

Table 15. Open-ended responses for ASR Science Grade 11

Questions	Responses
Please indicate any parts of the standard-setting training and process that we should improve.	There were three or four questions for which, after the discussion, for which I would have liked to have confirmed the answers. It may have been a case of knowing too much complexity about the subject.
	Randi was an excellent facilitator.
	It was all very helpful and the steps to learn the process well conveyed and reinforced. Maybe provide so general tools, prior reading for general understanding But the process was challenging but very very great learning experience.
	It would be helpful to have an actual mouse to use with the computers.
	Maybe a broad preview on the first day to show how this process fits in to the development of the ASR test.
	"The process is difficult, but well worth it. I am not sure if there would be a way to improve it. ** AC would be great ;)"
Please indicate any concerns you may have about the workshop process and the final recommended cut scores.	None, everything was explained as we worked through the material. I was a little slow on the uptake, but the facilitators were very responsive and patient with me. Their demeanor made the process more successful for me.
	I hope that we have set scores that allow for growth across the years.

APPENDIX—K
STANDARD-SETTING MEMO

Oklahoma Standard Setting

CCRA US History Assessment

June 23-24, 2022

Cognia and the Oklahoma Department of Education convened a panel of high school US History teachers during June 23-24, 2022 to recommend Basic, Proficient, and Advanced cut scores to enable reporting of student performance on the CCRA US History assessment. Eleven educators from around the state participated in two days of training and decision-making with Cognia standard-setting specialists. The standard-setting panelists reviewed test content and performance level descriptors and followed the Item-Descriptor (ID) Matching standard setting method to recommend these cut scores.

In the ID Matching method, the high school US History teachers reviewed the knowledge and skill response demands of CCRA US History assessment items placed in ordered item books (i.e., ordered from the easiest to the most difficult item). They matched those item response demands to knowledge and skill expectations in borderline performance level descriptors for the Basic, Proficient, and Advanced levels. (Borderline performance level descriptors define knowledge and skills that students who are *just barely* in a performance level are expected to know and be able to demonstrate.) Working independently, the standard-setting panelists conducted the ID matching process in three rounds and recommended cut scores for each of the three levels in each of the three rounds. After rounds 1 and 2, the Cognia workshop facilitator led panelists through a discussion of agreements and disagreements among the panelists and rationales for the various cut scores they recommended. The ensuing discussion enabled panelists to consider their colleagues' insights about item response demands and rationales for matching items to descriptors, and to consider adjusting their cut score recommendations in rounds 2 and 3. After the round 2 recommendations, and in preparation for making final cut score recommendations in round 3, panelists also reviewed impact data. (Impact data are the percentages of students who would be sorted into the Below Basic, Basic, Proficient, and Advanced performance levels, using their scores from the 2022 administration of the US History assessment, and based on the cut scores recommended in round 2.) The impact data gave the panelists one final opportunity to consider whether to adjust their cut scores in round 3.

In the final workshop evaluation, panelists expressed generally positive support for the workshop overall; workshop facilitation; training, practice, and the workshop process; the online standard setting tool; and other details in the standard setting workshop process. They responded this way to a final evaluation statement:

I'm satisfied with the final group cut scores. I would not recommend changing any of the group cut scores.

Table 1. Frequency of Responses for Final Evaluation Statement

N	Yes	No
11	10	1

If no, would you recommend moving a cut score up or down in the OIB, and by how many pages?

"Proficient – 4 pages higher" (panelist response).

Final recommended cut scores are calculated as the average recommended cut score (specifically the median cut score) across the 11 panelists. The final recommended cut scores and corresponding impact data appear in the table below.

The Oklahoma State Department of Education can accept these recommended cut scores and adopt them as is. Or the department may choose to make “policy adjustments” to the cut scores, using standard errors of the cut scores, for example, to account for the newness of the US History assessment and curriculum, overall test difficulty, resource limitations to support students who need more instruction in US History before retesting, and other considerations. Cognia can advise the department on psychometrically defensible ways to make policy adjustments.

Table 2. Final Recommended Cut Scores

Performance Level	Ordered Item Book Page	Percentage of Students
Advanced	41	9.9
Proficient	14	36.3
Basic	6	13.9
Below Basic	-	39.9
Proficient + Advanced	-	46.2

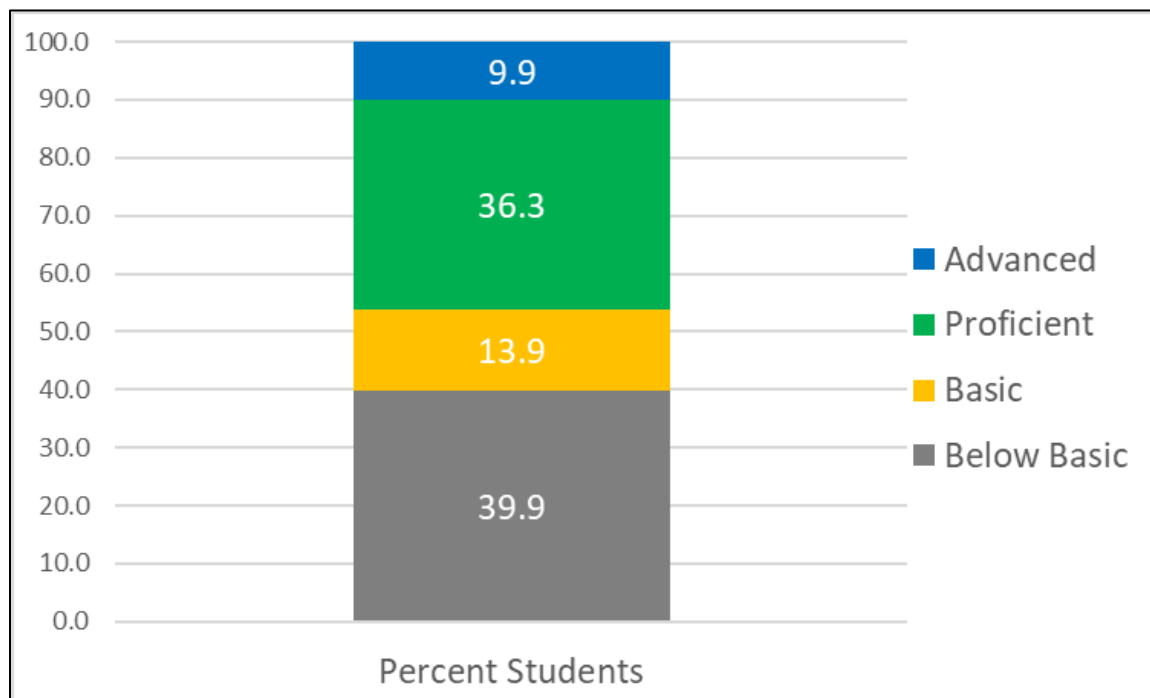
APPENDIX—L
FINAL CUT POINTS

OK CCRA USH Standard-Setting Final Cut Points

Table 1. OK CCRA USH Grade 11 Standard-Setting Final Cut Points

Performance Level	OIB Page #	Raw Score	Theta (Median)	Percent Students
Below Basic	--	--	--	39.9
Basic	6	22	-0.26	13.9
Proficient	14	25	0.14	36.3
Advanced	41	42	1.30	9.9
Proficient + Advanced	--	--	--	46.2

Figure 1. OK CCRA USH Grade 11 Impact Data based on Final Cut Points



APPENDIX P
2023 OSTP SCIENCE GRADE 8
STANDARD SETTING REPORT



2023 Oklahoma Standard Setting Report

OSTP Science—Grade 8

June 22–23, 2023—Oklahoma City, Oklahoma

Prepared by Cognia for the Oklahoma Department of Education



TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF STANDARD SETTING PROCEDURES.....	4
CHAPTER 2. TASKS COMPLETED PRIOR TO STANDARD SETTING	5
2.1 CREATION OF PERFORMANCE LEVEL DESCRIPTORS.....	5
2.2 PREPARATION OF MATERIALS	6
2.2.1 Ordered Item Booklet (OIB).....	6
2.2.2 Content-Based Benchmarks	7
2.2.3 Cognia Standard Setting Toolkit.....	7
2.2.4 Panelist Materials	8
2.2.5 Presentation Materials.....	8
2.2.6 Data, Information and Analysis Materials.....	9
2.3 SELECTION OF PANELISTS	10
CHAPTER 3. DURING THE STANDARD SETTING MEETING.....	11
3.1 OVERVIEW OF THE ID MATCHING METHOD	11
3.2 MEETING LOGISTICS.....	11
3.2.1 Standard Setting Panelists and Workshop Staff	11
3.2.2 Standard Setting Meeting Schedule.....	15
3.2.3 Standard Setting Meeting Security.....	15
3.2.4 ID Matching Standard Setting Procedure.....	15
3.3 CUT SCORE CALCULATION.....	15
3.4 GENERAL ORIENTATION AND PANELIST TRAINING	16
3.5 BECOMING FAMILIAR WITH THE TEST ITEMS AND CONTENT.....	16
3.6 USE OF THE COGNIA STANDARD SETTING TOOLKIT.....	17
3.7 REVIEW OF THE STANDARDS AND PERFORMANCE LEVEL DESCRIPTORS	17
3.8 TRAINING ON THE ID-MATCHING JUDGMENTAL TASK.....	17
3.8.1 Modeling and Practice	18
3.9 JUDGMENT ROUNDS AND FEEDBACK	18
3.9.1 Round 1 Judgments	20
3.9.2 Round 2 Judgments	20
3.9.3 Round 3 Judgments and Results	21
3.9.4 Validation Step	21
3.9.5 Workshop Evaluation	21
CHAPTER 4. TASKS COMPLETED AFTER THE STANDARD SETTING MEETING	22
4.1 FINAL ANALYSIS AND REVIEW	22
4.1.1 Review and Analysis of Standard Setting Results	22
4.1.2 Analysis and Review of Panelists' Feedback.....	23
4.2 POLICY ADJUSTMENTS	23
4.3 PREPARATION OF STANDARD SETTING REPORT.....	23
REFERENCES.....	24
APPENDICES	25

APPENDIX A—LOGISTIC REGRESSION CALCULATION
APPENDIX B—COGNIA STANDARD SETTING TOOLKIT
APPENDIX C—POWERPOINT PRESENTATIONS
APPENDIX D—MEETING AGENDA
APPENDIX E—NON-DISCLOSURE AGREEMENT
APPENDIX F—PERFORMANCE LEVEL DESCRIPTORS
APPENDIX G—READINESS SURVEYS



APPENDIX H—ROUND RESULTS
APPENDIX I—WORKSHOP EVALUATION SURVEY
APPENDIX J—WORKSHOP EVALUATION RESULTS
APPENDIX K—STANDARD SETTING MEMO
APPENDIX L—FINAL CUT POINTS



Chapter 1. Overview of Standard Setting Procedures

The purpose of this report is to summarize the activities involved in the Standard Setting process for the Oklahoma School Testing Program (OSTP) in grade 8 science on behalf of the Oklahoma State Department of Education (SDE). Changes in the Oklahoma Academic Standards for Science grade 8 were implemented in Fall 2022, necessitating the need to reset standards. The primary goal of the standard setting was to determine the knowledge, skills, and abilities (KSAs) that students must demonstrate to be classified into one of the performance levels (i.e., Advanced, Proficient, Basic, and Below Basic).

The Standard Setting process used was a modified version of the Item-Descriptor (ID) Matching method (Ferrara & Lewis, 2012; Cizek & Bunch, 2007). The ID Matching method was selected because it reduces cognitive burden on panelists as compared to other Standard Setting methods that require probability judgments about hypothetical high- and low-performing students, and it most clearly translates content standards into performance categories as compared to other methods of standard setting (Cizek, Bunch, & Koons, 2004).

The Standard Setting meeting was held from June 22nd through June 23rd of 2023. In all, 11 panelists participated in the process and were organized into three tables of 3–4 panelists each plus a facilitator provided by Cognia.

This report is organized into three major sections, describing tasks completed prior to, during, and after the Standard Setting meeting.

Chapter 2. Tasks Completed Prior to Standard Setting

2.1 Creation of Performance Level Descriptors

Oklahoma State Statute: Title 70. Schools, Chapter 22 – Testing and Assessment, Section 1210.541 – Student Performance Levels and Cut Scores – Accountability System mandates the adoption of “a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.” The law states that performance levels must be labeled and defined as follows:

1. Advanced, which shall indicate that students demonstrate superior performance on challenging subject matter;
2. Proficient, which shall indicate that students demonstrate mastery over appropriate grade-level subject matter and that students are ready for the next grade, course, or level of education, as applicable;
3. Basic, which shall indicate that students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level or course; and
4. Below Basic, which shall indicate that students have not performed at least at the limited knowledge level.

Cognia collaborated with the Oklahoma State Department of Education (OSDE) to develop Range performance level descriptors (PLDs) for OSTP Science Grade 8. Prior to this collaboration, Policy PLDs were established by the OSDE to define the knowledge and skill level expectations for the Oklahoma Academic Standards for Science (OAS-S).

In developing the draft Range PLDs, Cognia worked collaboratively with OSDE and took into consideration the content standards and the achievement construct the PLDs represent, and used statements developed for the OSTP Science Grade 8 assessment to organize Range PLDs for each assessable OSTP Science Grade 8 performance expectation (PE) by Science and Engineering Practice (SEP). Cognia reviewed the content standards to select (a) verbs that define science skills and thinking processes, (b) nouns to identify knowledge and understanding of science facts and concepts, and (c) modifiers (i.e., adverbs, adjectives) that indicate levels of frequency, consistency, or quality of student performance. Following the framework described in Egan et al. (2012), Cognia collaborated with the OSDE and other stakeholders to review the draft Range PLDs (i.e., knowledge and skill expectations for all students who have achieved the range of scores in a performance level). Lastly, Cognia and OSDE worked together to approve final Range PLDs in 2023. The final Range PLDs were approved by SDE in April of 2023.

Following approval of the final Range PLDs, Cognia developed the Borderline PLDs. The Borderline PLDs were developed with specific nouns, verbs, adjectives, and adverbs to describe the knowledge and skills that students within each proficiency level are just barely expected to be able to demonstrate. In line with the OAS-S, the statements combine the subject matter for science that students are expected to demonstrate at the borderline of each proficiency level.

2.2 Preparation of Materials

Preparing for the Standard Setting meeting involved analyzing operational test data and organizing key materials. The materials that were prepared prior to the Standard Setting meeting included the following:

- Ordered Item Booklet (OIB)
- Content-based benchmarks
- The Cognia Standard Setting Toolkit
- Panelist materials
- Presentation materials
- Data, information, and analysis materials

Details related to the materials preparation for each of the above categories are provided below.

2.2.1 Ordered Item Booklet (OIB)

The standard setting was conducted using test items from the Spring 2023 administration. The initial OIB comprised the test items, which were ordered in terms of difficulty. Item difficulty, as defined by its scale location given a response probability (RP) value, was calculated based on data from OSTP Science Grade 8 students during the Spring 2023 administration. Items ascended in terms of difficulty throughout the OIB. Easier items appeared earlier in the OIB, and more difficult items appeared later.

Response probability (RP) criterion. The RP 67 criterion, defined by the Item Response Theory (IRT) scale value associated with a 67% chance of answering the item correctly, was used to order items in the OIB for the OSTP Science 8 standard setting meeting.

Collection of items for the OIB. To ensure that the items included in the OIB spanned the difficulty continuum—from easy to difficult—and that items were found around the points on the test scale where cut scores were likely to appear, the following procedure was used for building the OIB.

- Start with an operational test form: Cognia ordered the items from the Spring 2023 operational test form. Operational items that fell below the statistical thresholds for psychometric adequacy were replaced with Spring 2023 field test (FT) items from the same domain that did meet the thresholds.
- Augment the OIB with additional field test items: As needed, Cognia chose additional items for the OIB from previously field-tested items. For example, if the OIB did not have many items near the point in the test scale where the Proficient benchmark was expected, then items were added to the OIB that had locations around this point based on availability of such items in the pool.
- Review the balance of content against the blueprint: Since additional items were substituted in or added to the OIB, Cognia confirmed that the items had a balance of content consistent with the test blueprint to ensure that individual content strands were less likely to be overrepresented in the OIB through the augmentation process.

2.2.2 Content-Based Benchmarks

In standard setting, benchmarks refer to any content- or policy-based information that comes from an external source and is presented to panelists. The exact way that the benchmarks are used in the standard setting depends upon the methodology used. However, the general use is the same: Standard Setting panelists see and consider information from these external measures as they engage in the Standard Setting meeting activities.

Content-based benchmarks were used for the OSTP Science Grade 8 standard setting. The procedure for determining the content-based benchmarks was as follows:

- Prior to the Standard Setting meeting, Cognia content teams reviewed each item in the OIB and matched the items to one of three PLD levels (Basic, Proficient, or Advanced). Note that the Cognia content specialists did not assign any items to the Below Basic PLD. This is because all OSTP Science items are written according to level Basic and above, and the Below Basic performance level is described simply as the inability to perform at the Basic level.
- Cognia psychometricians then compiled the content specialists' item-PLD alignments and calculated threshold regions through logistic regression. Specifically, the regions were calculated by combining the item-PLD judgments to derive a set of cut scores with two standard errors added below and above each cut score. See Appendix A for calculation details.
- The above process resulted in content-based benchmark regions for the Proficient and Advanced levels.

Special Considerations for the Basic Benchmark Region. As mentioned previously, the Below Basic performance level is described as the inability to perform at the Basic level; therefore, items were not written to the Below Basic level and, by extension, it was not feasible to align items to the Below Basic level. Since there were no Below Basic item-PLD alignments, the above logistic regression method could not be employed to calculate a cut and corresponding region for the Basic level.

Thus, to facilitate the Basic level cut score identification, Cognia psychometricians empirically derived the cut score by constructing a mini-Test Characteristic Curve (TCC) based on items that were aligned to the Basic PLD. Cognia interpreted the borderline PLD of 50% to mean that a student placed in the Basic performance level should be answering items aligned to the Basic PLD correctly 50% of the time when chance is considered. Thus, Cognia calculated a theta value that was associated with 50% beyond chance of the expected score of the mini TCC. The '50% beyond chance' criterion is reflected in the performance level descriptor and takes guessing into account. Two OIB pages were added below and above the empirical cut score to create an empirical threshold region for the Basic level.

2.2.3 Cognia Standard Setting Toolkit

This section provides details about the Cognia Standard Setting Toolkit that panelists used to complete the main Standard Setting activities during the meeting. The Cognia Standard Setting Toolkit was

developed, tested, and set up by Cognia prior to the meeting and included a digital ordered item booklet with integrated item list, judgment forms, readiness surveys, and the final workshop evaluation survey.

The Cognia Standard Setting Toolkit consisted of a digital interface that first presented the ordered item list view (i.e., a list of items separated by rows with the easiest item at the top and the most difficult at the bottom). From the initial screen, panelists could toggle to the corresponding item detail view and use navigation arrows to move 'up' or 'down' in the booklet. The item detail view showed a PDF of the full item with the response options, as well as any stimuli or rubrics associated with the item. The ordered item booklet was created as discussed in a previous section of this document. Integrated judgment forms were available within both the item list and detail views. The judgment forms provided space for users to note (1) the relevant knowledge, skills, and abilities (KSAs) needed to answer the item, (2) any additional information that came to mind as panelists undertook the judgment task for each item, and (3) item descriptor matches. Any notes entered by the user in the item list view screen persisted when the user switched to the detail view screen and vice versa. In addition to the above, the Cognia Toolkit included the round-specific readiness surveys that panelists completed before undertaking each judgment round. Finally, the toolkit included the final workshop evaluation survey that panelists completed at the conclusion of the Standard Setting meeting.

Additional details and screenshots of the Cognia Standard Setting Toolkit are available in Appendix B.

2.2.4 Panelist Materials

Cognia developed specific and relevant materials that were used by panelists during the meeting. Because panelists utilized the Cognia Standard Setting Toolkit for most of the Standard Setting activities, some of the materials were presented digitally within the Toolkit. Table 1 includes a list of the materials developed for the panelists and their mode of presentation.

Table 1. Panelist Materials Prepared Prior to the Standard Setting Meeting

Panelist Material	Paper	Digital Online	Digital Within the Toolkit
Meeting Agenda	✓	✓	
Non-disclosure Agreement	✓		
OSTP Science 8 Test		✓	
Performance Level Descriptors (PLDs)	✓		✓
Science Standards			✓
Practice Items and Judgment Forms			✓
Round Readiness Surveys			✓
Ordered Item Booklet (OIB)			✓
Integrated Item Map and Judgment Forms			✓
Workshop Evaluation Survey			✓

2.2.5 Presentation Materials

PowerPoint presentations guided the facilitator through the distribution of information and materials during the Standard Setting meeting. Cognia developed the initial presentations and OSDE reviewed the presentations prior to the standard setting meeting.

Notes and scripts that coincided with the PowerPoint slides were added within the presentation to guide facilitators. The notes and scripts for the meeting provided information, including procedural steps, talking points, definitions to explain concepts to panelists, answers to commonly asked questions, and specific materials to distribute to panelists. Copies of the PowerPoint presentations are available in Appendix C.

2.2.6 Data, Information and Analysis Materials

Prior to the Standard Setting meeting, data, information, and other relevant analysis materials were generated for use during the meeting. Table 2 shows a list of materials that were generated, as well as the purpose of each.

Table 2. Data, Information, and Analysis Materials Generated Before the Standard Setting Meeting

Data, Information, and Analysis Materials	Description/Purpose
Ordered Item Booklet (OIB)	The OIB was a set of items ordered by item difficulty and was generated according to the procedures outlined in section 2.2.1 of this report. Panelists worked within the OIB to review items and follow the ID Matching process.
Content-based benchmark regions	Benchmark regions were calculated according to the procedures outlined section 2.2.2 of this document. Panelists viewed and considered information from these benchmark regions as they engaged in the Standard Setting meeting activities.
Cognia Standard Setting Toolkit	A digital platform that was setup and tested prior to the meeting and included all necessary item data and information, as well as information related to the standards and PLDs.
Student Test Data	Student test data from the Spring 2023 administration of the OSTP Science grade 8 test were prepared to enable the calculation impact data during and after the meeting.
Programming	Cognia created and tested programming for computing the following: <ul style="list-style-type: none"> - Theta cut scores: Cut scores on the theta scale based on panelists' judgments after each judgement round. - Various statistics: Standard errors, percent exact and adjacent (based on differences between judgments from panelists and content specialists). - Panelist judgment frequency distributions: Computed for all panelists after each round. The code also produced presentation artifacts for use during the discussion session after each round. - Impact data: Code that used the theta cut scores and student test data to calculate the percentage of students in each performance level category.

2.3 Selection of Panelists

As emphasized in Cizek and Bunch (2007), regardless of the method used, the selection of panelists is an important factor in determining Standard Setting outcomes and maximizing the validity of the standard-setting process. The guidance provided by *Standards for Educational and Psychological Testing* (AERA et al., 2014) states that “a sufficiently large and representative group of judges should be involved to provide reasonable assurance that results would not vary greatly if the process were repeated.”

Consistent with the above guidance and respecting practical considerations regarding the maximum size of a group that can be successfully managed, the goal was to recruit a Standard Setting panel of 10–12 members representing different stakeholder groups to set standards for science. Targets for the size and composition of the panel were also consistent with federal guidelines as described in *Standards and Assessment Peer Review Guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001* (U.S. Department of Education, 2009).

Two goals were proposed for recruiting Standard Setting panelists: (a) diverse experience and points of view regarding students, student learning, and Oklahoma content standards and (b) diverse representation among panelists in years of teaching, geographic regions in the state, school system sizes, school system urbanicity, and the racial/ethnic make-up of the student and teacher populations.

Chapter 3. During the Standard Setting Meeting

3.1 Overview of the ID Matching Method

The Item-Descriptor (ID) Matching method is appropriate for setting standards for standards-aligned assessments like the OSTP Grade 8 Science assessment. Assessment programs around the world have used ID Matching (e.g., Delaware, Massachusetts, Maryland, Mississippi, New Mexico, New York, South Carolina, and West Virginia; the Chicago and Philadelphia Public Schools; and programs in Brazil, Germany, and Finland).

ID Matching has advantages over Bookmark, Angoff, and other Standard Setting methods. Specifically, its cognitive-judgmental task requires that Standard Setting panelists, who are typically classroom educators, undertake a judgmental task that they are well suited for—matching item knowledge and skill response demands with knowledge and skill expectations in performance level descriptors (PLDs). The Bookmark and other methods require panelists to make probability judgments—something that people in general do not do well (e.g., Murphy, 2002). In addition, panelists do not need to hold a hypothetical borderline student in mind when they match items to descriptors and recommend cut scores, so the cognitive load and complexity of ID Matching is more manageable.

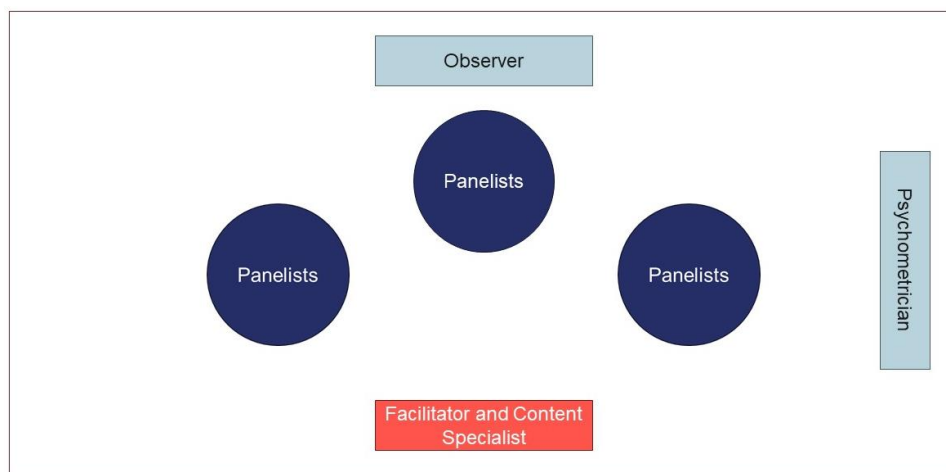
During standard setting using ID Matching, panelists use PLDs as their guide to match items to performance level descriptors. The structure of the PLDs provides a general characterization of expected student knowledge and skill at each level and examples of the knowledge and skills that students at each achievement level can be expected to demonstrate. By matching test items to specific claims from the Proficient PLD, for example, panelists identify the evidence in test items that supports the claims in that descriptor. Supporting the claims represented in the Proficient PLD contributes to the validity of interpretations of student achievement, based on the PLDs, and to the overall validity argument that a student who achieves that level on the assessment has demonstrated adequate understanding of essential concepts with respect to the standards being measured. This logic applies to all cut scores and performance levels.

3.2 Meeting Logistics

3.2.1 Standard Setting Panelists and Workshop Staff

Participants of the OSTP Science Grade 8 Standard Setting meeting included meeting facilitators, panelists, observers, and psychometricians. Figure 1 illustrates the room setup for the Standard Setting meeting.

Figure 1. Standard Setting Room Setup



Facilitators

The Standard Setting meeting was led by a facilitator with support from a science content specialist. The facilitator was a member of Cognia’s staff who has experience facilitating Standard Setting meetings and was responsible for leading the panelists through the Standard Setting process. The content specialist was a Cognia science test development specialist and was responsible for leading the panelists through the development of the test, procedures for scoring the items, and the review of PLDs.

The facilitator, with support from a Cognia psychometrician, ensured that appropriate Standard Setting processes were followed throughout all phases of the meeting and verified that panelists had a solid understanding of the tasks they were being asked to complete. The facilitator, along with the content specialist, underwent preparatory training to lead the Standard Setting meeting. Psychometric staff from Cognia conducted the training, which included:

- OSTP Science Grade 8 assessment overview: The facilitators were provided with an overview of the OSTP Science 8 test, including the different item types, scoring rules, and performance levels.
- Cognia Standard Setting Toolkit: Both the facilitator and content specialist became familiar with the Cognia Toolkit to lead the Standard Setting process.
- Standard setting process: Facilitators participated in a walkthrough of the Standard Setting meeting, with a focus on specific issues for these meetings, such as time management, the use of the Cognia Standard Setting Toolkit, and communicating feedback information.
- Training slides and presentation script/notes: As part of the walkthrough of the standard setting process, facilitators reviewed the Standard Setting training slides. Notes in the Standard Setting training slides and a presentation script provided the facilitators with guidance, including when specific language was to be used.

Panelists

The SDE selected panelists prior to the Standard Setting meeting. The goal for panel selection was to include participants who are primarily teachers, but also to include school administrators, higher education personnel, and stakeholders from other interest groups. Moreover, to the extent possible, panelists were selected to reflect a balance of gender, race/ethnicity, and geographic location. Finally, panelists were selected who were familiar with the eighth-grade science subject matter. Table 3-1 provides information about the panelists that participated in the OSTP science grade 8 standard setting.

Table 3-1. OK OSTP Science Grade 8 Standard Setting Committee Participant List

Panelist #	District	Grade Level Currently Teaching Science	Science Content Experience	Gender	Suburban, Urban, Rural	District Ethnicity Breakdown	Site
1	Mustang	8th	Life Science	F	Suburb	52% White; 19% Hispanic; 8% Asian; 6% AA; 3% AI	Mustang North Middle School
2	Woodward	8th	Physical Science	F	Rural	67% White; 24% Hispanic; 5% AI; 1% AA; 1% Asian	Woodward Middle School South
3	Alex	8th	Life Science	F	Rural	74% White; 9% Hispanic; 6% AI	Alex Jr/Sr High
4	Owasso	8th	Earth and Space Science	F	Suburb	54% White; 12% Hispanic; 8% AI; 6% AA; 6% Asian	Owasso 8th Grade Center
5	Tahlequah	8th	Life Science	F	Rural	36% AI; 20% White; 20% Hispanic; 1% Asian; 1% AA	Tahlequah Middle School
6	Edmond	8th	Life Science	F	Urban	70% White; 11% Hispanic; 5% AA; 3% AI; 3% Asian	Edmond - Central Middle School
7	Enid	8th	Physical Science	F	Rural	47% White; 29% Hispanic; 12% Asian; 4% AA; 3% AI	Enid - Emerson Middle School
8	Woodward	8th	Earth and Space Science	F	Rural	67% White; 24% Hispanic; 5% AI; 1% AA; 1% Asian	Woodward Middle School
9	Skiatook	8th	Physical Science	F	Suburb	56% White; 19% AI; 7% Hispanic	Newman Middle School
10	Putnam City	8th	Earth and Space Science	F	Urban	39% Hispanic; 23% White; 21% AA; 4% Asian; 2% AI (Capps)	Capps Middle School/Hefner MS
11	Woodward	8th	Physical Science	F	Rural	67% White; 24% Hispanic; 5% AI; 1% AA; 1% Asian	--

3.2.2 Standard Setting Meeting Schedule

The Standard Setting meeting consisted of two days of activities. The meeting started with an opening session on the morning of day 1 before continuing with training, practice, and round 1. On day 2, panelists completed rounds 2 and 3, and concluded the meeting with the final workshop evaluation survey. A detailed meeting agenda can be found in Appendix D.

3.2.3 Standard Setting Meeting Security

During the meeting, panelists reviewed operational test items, preliminary cut score recommendations, and associated impact data. Due to the nature of this information, security was a critical component of the meeting. Specific procedures were established to ensure the security of all materials was maintained.

As part of the meeting, the facilitator reviewed the process for maintaining the security of materials, discussions, and preliminary results from the meeting. Panelists were not permitted to share or discuss secure materials and information outside of meeting rooms. To confirm that the panelists understood and agreed to the security conditions, they signed security and non-disclosure agreements (an example is provided in Appendix E).

To preserve the security of the materials and activities within the Cognia Standard Setting Toolkit, each panelist was provided a Chromebook and unique login credentials. The supporting Cognia psychometrician-controlled panelist access to each section of the Toolkit throughout the meeting. Access to the Toolkit was disabled at the conclusion of the Standard Setting meeting and the Chromebooks were wiped clean of all data.

Additional materials were provided to panelists in their meeting folders after signing the non-disclosure agreement. All printed materials were collected at the end of each day to maintain test security.

3.2.4 ID Matching Standard Setting Procedure

Over the course of two days, panelists engaged in Standard Setting activities, starting with an opening session on day one. The opening session was followed by the main Standard Setting session during which panelists received training and engaged in a practice round. Next, panelists engaged in three consecutive judgment rounds, with preparation and discussion between rounds. The Standard Setting meeting will conclude after the third round, at which point a final workshop evaluation will be administered.

3.3 Cut Score Calculation

To calculate the Proficient and Advanced cut scores during the Standard Setting meeting, all item-PLD alignment judgments from each panelist were gathered and used as input in a logistic regression calculation (see Appendix A for details).

To facilitate the Basic level cut score identification, Cognia psychometricians empirically derived the cut score by constructing a miniature Test Characteristic Curve (TCC) based on items that were aligned to the Basic PLD. Cognia interpreted the borderline PLD of 50% to mean a student placed in the Basic

performance level should be answering items aligned to the Basic PLD correctly 50% of the time when chance is considered. Thus, Cognia calculated a theta value that was associated with 50% beyond chance of the expected score of the mini TCC. The '50% beyond chance' criterion reflected in the performance level descriptor also takes guessing into account.

Note that during the first round of standard setting, panelists made item-PLD alignments for each item. During rounds 2 and 3, they had the opportunity to change their item-PLD alignments as they saw fit. Thus, the above process was used to calculate cuts during each round of the standard setting by using the complete set of panelists' judgments for that specific round.

3.4 General Orientation and Panelist Training

Concerning panelist training, the *Standards for Educational and Psychological Testing* (AERA et al., 2014) states the following:

Care must be taken to assure these persons understand what they are to do and that their judgments are as thoughtful and objective as possible. The process must be such that well-qualified participants can apply their knowledge and experience to reach meaningful and relevant judgments that accurately reflect their understandings and intentions. (p. 101)

The training of the panelists began with a general orientation session at the start of the meeting. During the main Standard Setting session, panelists were organized such that three to four panelists were assigned to each table. Chromebooks, supplied by Cognia and set up for the standard setting, were distributed to all panelists. Facilitators guided panelists through the following activities:

- Overview and introductions
- Taking the test
- Use of the Cognia Standard Setting Toolkit
- Review of the standards and PLDs
- Training on the ID Matching process
- Modeling and practice
- Judgment rounds and feedback
- Final workshop evaluation survey

To begin the main Standard Setting session, the individuals in the room introduced themselves. After introductions, the facilitator reviewed the security and non-disclosure information. The facilitator then provided a high-level overview of the process. The panelists were given opportunities to ask questions before proceeding.

3.5 Becoming Familiar with the Test Items and Content

After the overview and introductions, panelists experienced the OSTP Science Grade 8 test. The purpose of this step was to familiarize the panelists with the assessment and the test taking activities expected of students during administration.

Using individual Chromebooks provided by Cognia, panelists were instructed on how to log into their Chromebooks and navigate to the testing platform site. Cognia staff provided panelists with unique login credentials and once they successfully accessed the testing platform, panelists experienced the test the same way students do, to become familiar with the test from the student's perspective.

In the interest of time and efficiency, panelists were presented with session 1 of the OSTP Science Grade 8 test. Session 1 represented half of the full test. Cognia's science content specialist confirmed that the set of items in the first session included all the item types that would be encountered on the full test. In addition, the range of content standards and item difficulties in session 1 were representative of the full test and the test blueprint.

3.6 Use of the Cognia Standard Setting Toolkit

The facilitator (with support from a Cognia psychometrician) guided panelists through the steps needed to log in and access the Cognia Standard Setting Toolkit. Each panelist used their email and an initial assigned password to access the site. After their initial log in, panelists were directed to change their passwords, and then prompted to log back into the system with their new passwords. Their emails and individual passwords were used to access the Toolkit for the duration of the Standard Setting meeting. Once everyone completed the log in procedure, they viewed an initial screen with tabs that linked to the standards and PLDs.

3.7 Review of the Standards and Performance Level Descriptors

Before engaging in the judgment tasks, panelists reviewed the standards and the performance level descriptors (PLDs). This important step was designed to ensure that panelists thoroughly understood the KSAs needed for students to be classified into performance levels (Below Basic, Basic, Proficient, and Advanced).

Panelists studied the standards and range PLDs associated with the OSTP Science Grade 8 assessment. Panelists were asked to consider the knowledge, skills, and abilities (KSAs) detailed in the standards, and how they were reflected in the PLDs. Next, panelists focused on the borderline PLDs that describe the KSAs expected of students who just barely meet each performance level. Panelists then reflected on their conversations about the standards and the PLDs. The PLDs are provided in Appendix F.

3.8 Training on the ID-Matching Judgmental Task

Once panelists reviewed and discussed the standards, range and borderline PLDs, the facilitator guided them through more detailed training on the ID-Matching judgmental task. The facilitator used a customized PowerPoint slide deck to explain the following concepts: the ordered item booklet (OIB), how to review items and what information to consider while doing so, and how to make item-descriptor matches. The facilitator emphasized the importance of considering the knowledge, skills, and abilities (KSAs) required by an item, as well as the information in the PLDs to make their item-descriptor matches.

After explaining the main concepts and the process for making item-descriptor matches, the facilitator provided a high-level description of the round-by-round judgement procedures and what to expect before (i.e., readiness survey), during (i.e., judgmental tasks and, when relevant, consideration of benchmarks), and after (i.e., presentation of results and discussion) each round.

During the training, the facilitator provided clear explanations and directions while ensuring that the panelists had all the information and support needed to undertake the Standard Setting process. To that end, the facilitator used a customized script alongside the PowerPoint slide deck to guide panelists through the training.

The facilitator encouraged panelists to ask questions during the training but also reminded panelists that they would have the opportunity to practice before beginning the first round. In addition, the facilitator reminded panelists that they would review concepts as needed throughout the Standard Setting process.

3.8.1 Modeling and Practice

After training on the ID-Matching process, the facilitator provided a brief demonstration of the Cognia Standard Setting Toolkit. A Cognia psychometrician, with dedicated access to a management screen within the Cognia Toolkit, was responsible for managing aspects related to the system.

After the initial demonstration of the Cognia Toolkit, the facilitator proceeded with the practice round, which consisted of three sample items. The facilitator used the three sample items to model the judgmental task and guide panelists through making their own item-descriptor matches. During this practice round, the facilitator reinforced the training concepts.

The three sample items were chosen such that (1) none of the items were part of the OIB, (2) the first two items were relatively easy to identify in terms of item-PLD alignment, and (3) the last item was more challenging to identify in terms of item-PLD alignment (i.e., the item was expected to fall in a borderline region). Using sample items that were not part of the OIB allowed the facilitator to avoid undue influence over panelists' judgmental tasks. In addition, the mix of items gave panelists the opportunity to experience different levels of cognitive load while making their judgments, as would be the case once they considered the full set of items contained in the OIB. During the modeling and practice session, panelists also had the opportunity for discussion with each other, to ask questions, and become more familiar with the Toolkit.

3.9 Judgment Rounds and Feedback

During the main portion of the Standard Setting workshop, panelists completed three consecutive rounds of judgments. Each judgment round consisted of three distinct sessions: Readiness, Judgment, and Feedback and Discussion. This was an iterative process during which the outcomes of each judgment round were considered during the next judgment round. Table 3 provides a crosswalk of the activities, analyses, and outcomes for each session within each judgment round.

Table 3-2. Crosswalk of Activities, Analyses, and Outcomes by Judgment Round

Round	Session	Panelist Activities	Analyses	Outcomes
1	Readiness	Complete Round 1 readiness survey.	Determine if all panelists are ready to proceed.	
	Judgment	Review all items, identify KSAs, and align each item to a PLD.	1. Calculate threshold regions for Proficient and Advanced levels (cut score with 2 standard errors) 2. Calculate % exact agreement on OIB items 3. Create presentation artifacts	1. Initial threshold regions 2. Presentation artifacts
	Feedback & Discussion	Discuss round 1 results: items with the most disagreement		
2	Readiness	Introduce content-based benchmark regions. Complete Round 2 readiness survey.	Determine if all panelists are ready to proceed.	
	Judgment	Review items (with special attention to items discussed in round 1 feedback) and make changes to item-PLD alignments as desired.	1. Calculate threshold regions for Proficient and Advanced levels (cut score with 1 standard error) 2. Calculate % exact agreement on OIB items 3. Create presentation artifacts	1. Narrowed threshold regions 2. Presentation artifacts
	Feedback & Discussion	Discuss round 2 results: items with the most disagreement and benchmarks		
3	Readiness	Complete Round 3 readiness survey.	Determine if all panelists are ready to proceed.	
	Judgment	Review items (with special attention to items discussed in round 2 feedback) and make changes to item-PLD alignments as desired.	1. Calculate cut scores 2. Calculate associated impact data 3. Create presentation artifacts	1. Cut scores and impact data 2. Presentation artifacts
	Feedback & Discussion	Additional validation step to address any remaining differences between panel results and content-based benchmarks. Present final cut scores and impact data to panelists		1. Group-level content-based rationale for final cut score recommendations.

Readiness Surveys: Before each judgment round, panelists completed a readiness survey that consisted of questions about whether they felt prepared to undertake the upcoming round of judgements. All questions had yes/no response options, and all “yes” responses indicated that panelists were ready to proceed. See Appendix G for the readiness surveys for all three rounds. If one or more panelists answered “no” to one or more questions, the facilitator reviewed the concepts associated with those questions, and panelists were then asked to complete the readiness survey again. Panelists moved on to the judgement round only when everyone indicated that they were ready to do so.

Feedback and Discussion: After each judgement round, Cognia psychometricians calculated a variety of statistics as described previously. In addition, the psychometricians created a presentation artifact in the form of a frequency chart. During the feedback and discussion portion that followed each judgement round, the facilitator presented the frequency chart to the panelists and used it to facilitate table and room discussions. The discussion focused on items that showed the most disagreement between panelists, and panelists were encouraged to share their thoughts and viewpoints. Panelists were encouraged to refer to training materials (e.g., OIB, item information, PLDs, and standards) as well as their own notes (taken within the Toolkit) throughout this discussion. Panelists were also reminded that the goal of the discussion was not to persuade or influence others. Instead, the discussion centered around sharing their

own reasoning for their PLD matches and listening to other panelists' reasons as additional information to consider.

3.9.1 Round 1 Judgments

During the first round, panelists worked individually with the PLDs, the standards, and the ordered item booklet (OIB). For each item in the OIB, panelists considered the knowledge, skills, and abilities (KSAs) needed to respond to the item (i.e., asking themselves 'what does a student need to know and be able to do to respond to this item?'). After identifying the KSAs required by the item, panelists then assigned an item descriptor match (i.e., basic, proficient, or advanced) to the item. They continued in this manner until they reviewed all items in the OIB.

At the conclusion of round 1 judgments, Cognia psychometricians compiled all judgments from all panelists to calculate cut scores and associated standard errors. The cuts with two standard errors above and below represented the threshold regions for round 1. In addition, Cognia psychometricians calculated an item-level % exact PLD agreement to facilitate round 1 discussion. Finally, the psychometricians created the presentation artifact (i.e., a graphical representation of results) that was handed off to the facilitator.

3.9.2 Round 2 Judgments

Before starting the second round of judgements, the panelists were introduced to the content-based benchmarks. The facilitator, with support from a psychometrician, described how the benchmarks were calculated, demonstrated how they would be presented within the Cognia Toolkit, and explained how panelists should consider the information represented by the benchmarks as they engaged in round 2 of the Standard Setting activities. Panelists were reminded that benchmarks were provided for their consideration, and not to influence their judgments.

Next, panelists completed the round 2 readiness survey and once all panelists indicated that they were ready to proceed, they continued to round 2 of the judgement task.

During the second round, panelists once again worked individually with the PLDs, the standards, and the ordered item booklet (OIB). Taking into consideration the feedback and discussion after round 1, as well as the additional information represented by the content-based benchmarks, panelists reviewed their work from round 1. Panelists could keep their judgment from round 1 or revise it. All panelists made their round 2 judgments individually and without discussion.

At the conclusion of round 2 judgments, Cognia psychometricians again compiled all judgments from all panelists to calculate cut scores and associated standard errors. The cuts with one standard error above and below represented the narrowed threshold regions for round 2. In addition, Cognia psychometricians calculated an item-level % exact PLD agreement to facilitate round 2 discussion. Finally, the psychometricians created the presentation artifact (i.e., a graphical representation of results) that was handed off to the facilitator.

3.9.3 Round 3 Judgments and Results

After round 2 feedback and discussion portion, but before round 3, panelists once again completed a readiness survey. Once all panelists indicated that they were ready to proceed, they continued to round 3 of the judgment task.

During the third round, panelists once again worked individually with the PLDs, the standards, and the ordered item booklet (OIB). Taking into consideration the feedback and discussion after round 2, panelists reviewed their work from round 2. Panelists could keep their judgment from round 2 or revise it. All panelists made their round 3 judgments individually and without discussion.

At the conclusion of the round 3 judgments, Cognia psychometricians again compiled all judgments from all panelists and, using the same procedures already detailed in previous sections, used the panelists' item-PLD judgements to calculate the final cut scores, as well as associated impact data. In addition, the results were reviewed and compared to the content-based benchmarks.

The frequency of panelists item-PLD judgments across the basic, proficient, and advanced levels for each of the three rounds are available in Appendix H. Note that these frequency results are the same graphical displays that were presented to panelists after each round.

The round 3 results were not congruent with the content-based benchmarks. As per the Standard Setting plan and discussion with SDE, the results (including impact data) were shared with panelists and panelists were asked to complete a validation step.

3.9.4 Validation Step

During the validation step, the facilitator guided the panelists to write content-based rationales for their judgments associated with items that were still very much discrepant from the content-based benchmarks. Panelists captured these content-based rationales in the Cognia Toolkit.

3.9.5 Workshop Evaluation

At the conclusion of the Standard Setting meeting, panelists completed a final workshop evaluation form and gave their feedback on various aspects of the Standard Setting meeting. Panelists indicated that they felt positive about how Cognia conducted the workshop and their final recommendations. Specifically, panelists expressed generally positive support for the workshop overall; workshop facilitation; training, practice, and the workshop process; the Cognia Standard Setting Toolkit; and other details in the workshop process. A copy of the evaluation survey is available in Appendix I; the workshop evaluation results are available in Appendix J.

Chapter 4. Tasks Completed After the Standard Setting Meeting

Upon conclusion of the Standard Setting meeting, several important tasks were completed. These tasks centered on the following: reviewing the Standard Setting process and addressing issues presented by the outcomes; presenting the results to the SDE; and making any final revisions or adjustments based on policy considerations, under direction of the SDE. Shortly after the Standard Setting meeting, Cognia provided SDE with a Standard Setting memo that included an overview of the Standard Setting process, as well as the final recommended cut scores. A copy of the memo is available in Appendix K.

4.1 Final Analysis and Review

The Standard Setting literature considers evaluation of the workshop and its results to be another product of the Standard Setting process (e.g., Reckase and Chen, 2012), as it provides important validity evidence supporting the cut scores that are obtained. To that end, a final review and analysis of the Standard Setting results was conducted. In addition, to provide evidence of the participants' views of the Standard Setting process, a review and analysis of panelists' feedback on the workshop evaluation survey was also conducted.

4.1.1 Review and Analysis of Standard Setting Results

First, Cognia conducted statistical analyses of panelists' item-PLD alignment data by calculating the percent exact, adjacent, and discrepant for each panelist on each performance level. Panelists with the least percentage exact were identified as showing statistically aberrant behavior. Next, an independent subject matter expert (SME) reviewed the qualitative data for all panelists identified as statistically aberrant. The SME reviewed panelists' notes on the knowledge, skills, and abilities required by the items, as well as their content-based rationales to determine if the panelists were on task.

After the statistical analyses and qualitative review of panelist data, one panelist was determined to be statistically and qualitatively aberrant. Consequently, their data were removed from the final analyses.

The next phase of the analyses included conducting logistical regression to calculate cut scores. Since the logistical regression method is sensitive to statistical outliers and the presence of such outliers violates the assumptions of the model, an outlier analysis was performed in the form of visual inspection of the initial logistic regression curves for any statistical outliers. A total of 430 data points (10 panelists made judgements on each of 43 items) were included in the logistic regression calculation. Visual inspection of the initial logistic regression curves revealed seven statistical outlier data points.

After the seven data points were removed, the final logistic regression analyses were conducted to calculate the proficient and advanced cut scores. Next, the TCC method was used to calculate the Basic cut score.

Finally, the resulting cut scores were applied to student data from the Spring 2023 administration of the OSTP Science Grade 8 assessment to calculate the impact data (i.e., the percentage of students that would be classified into each performance level based on the Standard Setting cut scores).

4.1.2 Analysis and Review of Panelists' Feedback

After the evaluation forms were completed, panelists' responses were reviewed. This review did not reveal any anomalies in the Standard Setting process. In general, participants felt that the recommended cut points were appropriate and that their judgments were based on appropriate information and decision making. The results of the evaluations are presented in Appendix J.

4.2 Policy Adjustments

After all Standard Setting activities had been completed and all materials reviewed, the SDE recommended an adjustment to the Standard Setting results. Specifically, all three cut scores were adjusted by one standard error. The full set of cuts, shown in Appendix L, were presented to the Commission for Educational Quality and Accountability (CEQA), and approved for use assigning students to performance levels in the 2022–2023 Oklahoma Science Grade 8 assessment.

4.3 Preparation of Standard Setting Report

Following the final compilation of Standard Setting results, Cognia prepared this report, which documents the procedures and results of the 2023 Standard Setting meeting that was held to establish performance standards for the OSTP Science Grade 8 assessment.

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological Testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: Establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press
- Reckase, M.D. (2001). Innovative methods for helping Standard Setting participants to perform their task: The role of feedback regarding consistency, accuracy, and impact. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives* (pp. 159–173). Mahwah, NJ: Lawrence Erlbaum Associates.
- U.S. Department of Education. (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: U.S. Department of Education Office of Elementary and Secondary Education. Retrieved June 10, 2010, from the World Wide Web: www2.ed.gov/policy/elsec/guid/saaprguidance.pdf.

Appendices

APPENDIX A
LOGISTIC REGRESSION CALCULATION

Logistic Regression Calculation

The proficient and advanced cut scores were computed using the logistic regression as follows:

$$\log \frac{P}{1-P} = \beta_0 + \beta_1 \theta$$

which is equivalent to:

$$P = \frac{\exp(\beta_0 + \beta_1 \theta)}{1 + \exp(\beta_0 + \beta_1 \theta)}$$

Where β_0 (intercept) and β_1 (slope) are two regression coefficients that need to be computed, theta (θ) is the RP67 value associated with each OIB page, and P is the probability of observing a performance level (level X or above) given theta. After fitting the model with data, the theta cut score is obtained by finding which score corresponds to a probability of 0.5 for being rated above the cut as follows:

$$\log \frac{0.5}{1-0.5} = 0 = \beta_0 + \beta_1 \theta$$

Solving the equation, the following is obtained:

$$\theta = -\frac{\beta_0}{\beta_1}$$

Additionally, the variance of the theta estimate will be computed as:

$$VAR(\theta) = \frac{\mu_{\beta_0}^2}{\mu_{\beta_1}^2} \left[\frac{\sigma_{\beta_0}^2}{\mu_{\beta_0}^2} - 2 \frac{Cov(\beta_0, \beta_1)}{\beta_0 \beta_1} + \frac{\sigma_{\beta_1}^2}{\mu_{\beta_1}^2} \right]$$

Therefore, the standard error of the estimate is given by:

$$SE(\theta) = \sqrt{VAR(\theta)}.$$

APPENDIX B


COGNIA STANDARD-SETTING TOOLKIT

Appendix B: Cognition Standard Setting Toolkit

This appendix contains sample screenshots of the Cognition Standard Setting Toolkit that panelists used for all standard setting activities during the meeting. Images provided correspond to samples (1) login screen, (2) readiness survey screen, (3) ordered item booklet view, and (4) item detail view.

Figure 1. Sample Login Screen

Panelists are provided with usernames and passwords to enable secure access to the toolkit.

 **Standard Setting Toolkit** [Home](#) [Register](#) [Login](#)

Log in


Email

Password

☐ Remember me?

[Log in](#)

Figure 2. Sample Readiness Survey


 **Standard Setting Toolkit** [Home](#) [Admin](#) sstksetup@cognition.org [Logout](#)

Questionnaire

Demo Subject Grade, step 1 - Readiness Survey

Position	Question	Response
1	I understand the goals of this meeting	-- ▾
2	I understand the task at hand	-- ▾
3	I am ready to proceed with the meeting activities	-- ▾

Figure 3. Sample Ordered Item Booklet View


Standard Setting Toolkit
[Home](#)
[Admin](#)

sttksetup@cognia.org
[Logout](#)

Item Review

Demo Subject Grade Step 2 Item Review

[Hide Documents](#)
[Standards](#)
[Range PLDs](#)
[Borderline PLDs](#)


























































Position	Asset ID	Description	Point Value	Relevant KSAs	Rationale or Notes	Item Descriptor Match Level	
1	ItemID01	Short Item Description 1	1			-- 	Detail
2	ItemID02	Short Item Description 2	1			-- 	Detail
3	ItemID03	Short Item Description 3	1			-- 	Detail
4	ItemID04	Short Item Description 4	1			-- 	Detail
5	ItemID05	Short Item Description 5	1			-- 	Detail
6	ItemID06	Short Item Description 6	1			-- 	Detail
7	ItemID07	Short Item Description 7	1			-- 	Detail
8	ItemID08	Short Item Description 8	1			-- 	Detail
9	ItemID09	Short Item Description 9	1			-- 	Detail
10	ItemID10	Short Item Description 10	1			-- 	Detail
11	ItemID11	Short Item Description 11	1			-- 	Detail
12	ItemID12	Short Item Description 12	1			-- 	Detail
13	ItemID13	Short Item Description 13	1			-- 	Detail
14	ItemID14	Short Item Description 14	1			-- 	Detail
15	ItemID15	Short Item Description 15	1			-- 	Detail
16	ItemID16	Short Item Description 16	1			-- 	Detail
17	ItemID17	Short Item Description 17	1			-- 	Detail
18	ItemID18	Short Item Description 18	1			-- 	Detail
19	ItemID19	Short Item Description 19	1			-- 	Detail

Figure 4. Sample Item Detail View

cognia Standard Setting Toolkit Home Admin ▼sstksetup@cognia.org Logout

Item Detail

Demo Subject Grade step 2 Item Review

[Return to item review](#)

Position AssetId Point Value [Prev Item](#) [Next Item](#)

[Item](#) [Standards](#) [Range PLDs](#) [Borderline PLDs](#)

Item

ItemImage

1 / 1 58% +

Which of the following values belong in the \square to make this statement true?

$11 \times \frac{\square}{10}$ is greater than 11.

Select the **three** correct answers.

☐ A. 3

☐ B. 7

☐ C. 8

☐ D. 9

Standard

Standard code here

Standard description detail here

Relevant KSAs

Rationale or Notes

Item Descriptor Match

--

APPENDIX—C
POWERPOINT PRESENTATIONS



OSTP Science Grade 8

Standard Setting Orientation

June 22 – 23, 2023

Welcome!

Thank you for taking time out of your summer to help us.

Assessment History

- In 2016, the Oklahoma Legislature directed the State Board of Education to evaluate Oklahoma's current state assessment system and make recommendations for its future.
- As a result, the Oklahoma State Department of Education
 - Held regional meetings across the state to determine stakeholder concerns
 - Convened the Oklahoma Assessment & Accountability Task Force to develop recommendations
 - Followed federal requirements and rules as described in ESSA.

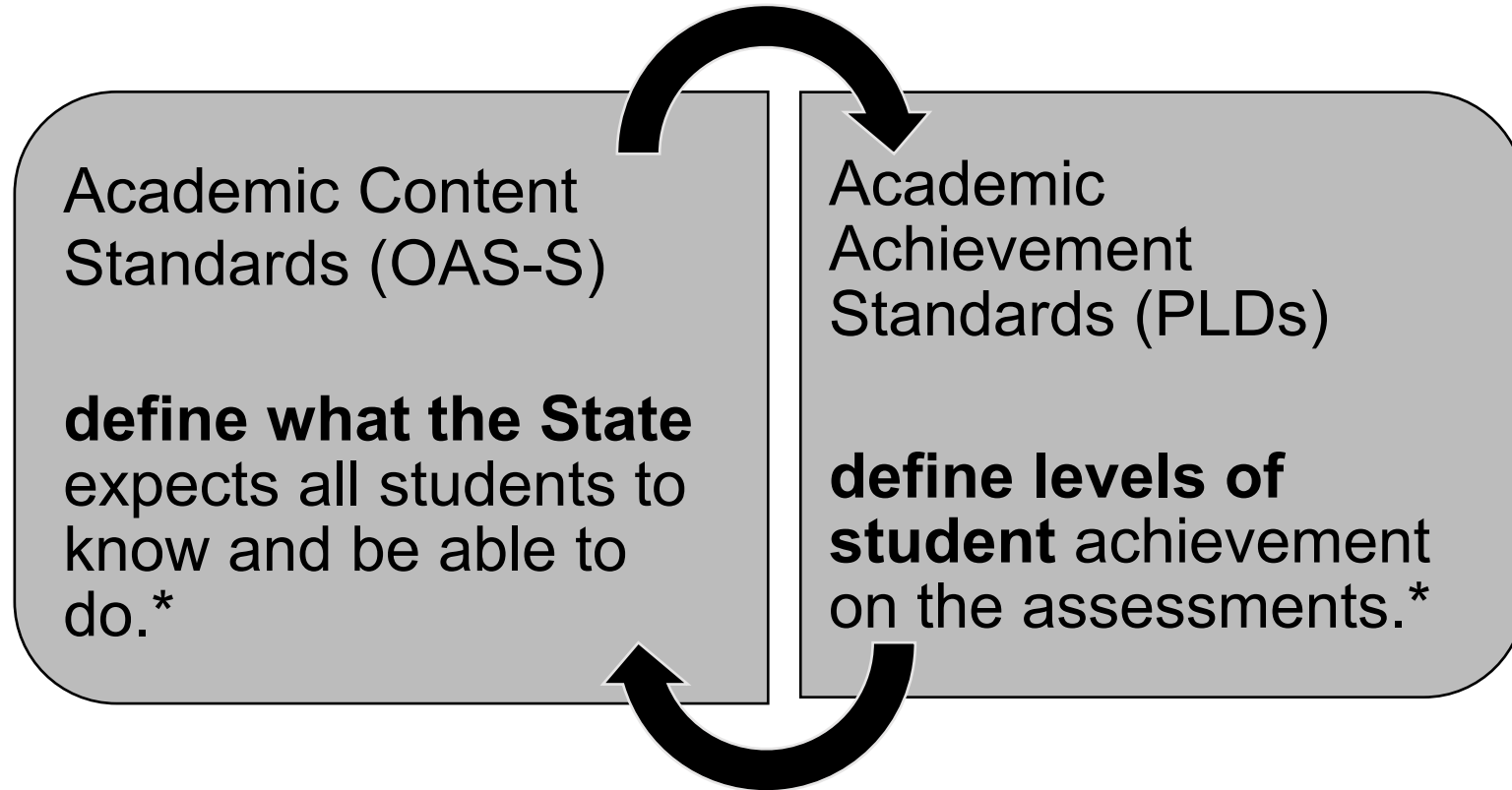
Goals for Oklahoma Schools

- Focus on college- and career-readiness:
 - *College and career ready means that students graduate from high school prepared to enter and succeed in postsecondary opportunities whether college or career.*
- Students should graduate high school ready for postsecondary success and should be able to demonstrate that they are on track toward that goal.

Commission for Educational Quality

- The Commission for Educational Quality and Accountability shall determine and adopt a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.

Content Standards and PLDs



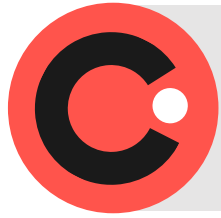
**U.S. Department of Education Peer Review of State Assessment Systems Non-Regulatory Guidance for States, September 25, 2015*



OSTP Science Grade 8

Standard Setting Orientation

June 22 – 23, 2023



Orientation Session - Agenda

- Introduction of the Standard Setting Team
- Standard Setting Goals and Outcomes
- Overview of the OSTP Science Grade 8 Assessment
- Overview of Standard Setting
- Overview of Key Concepts and Procedures
- Overview of Performance Level Descriptors

A large, stylized graphic on the left side of the slide. It consists of a white semi-circle with a thick black outline. Inside the white area, there is a solid red circle and a thick black curved line, together forming a stylized smiley face.

Welcome!

Thank you for taking time out of your summer to help us.

Standard Setting Team

- Oklahoma SDE Members
 - Catherine Boomer—Program Director of State Assessment
 - Eric Jones—Program Manager of State Assessment
 - Samantha Sheppard—Project Manager of Science Assessment
 - Heather Johnston—Project Manager of Secondary Science and Engineering
 - Caroline Misner—Project Manager of OAAP
- OSTP Technical Advisory Committee Member
 - Juan D'Brot (observer)
- Cognia
 - David Harrison (facilitator)
 - Mary-Alice Corliss (content SME)
 - Liz Garcia (lead program manager)
 - Frank Padellaro (VP Psychometrics and Reporting)



Standard Setting Goals

Our shared goals

- Use your judgments to help provide performance standards recommendations for the OSTP science grade 8 assessment that provide meaningful and actionable information

Your goals as panelists

- Learn concepts and procedures following the Item-Descriptor (ID) Matching method
- Follow the procedures to complete the standard setting activities
- Rely on your expertise about the content standards, student learning, and students throughout the process

Expectations of all Panelists



- Security is of the utmost importance
 - You can discuss the process in general terms
- You may NOT
 - Share details about the items or specific details about the process (e.g., cuts that were recommended)
 - Use your phones or personal devices while in the room
 - Use the Chromebooks for anything other than standard setting activities



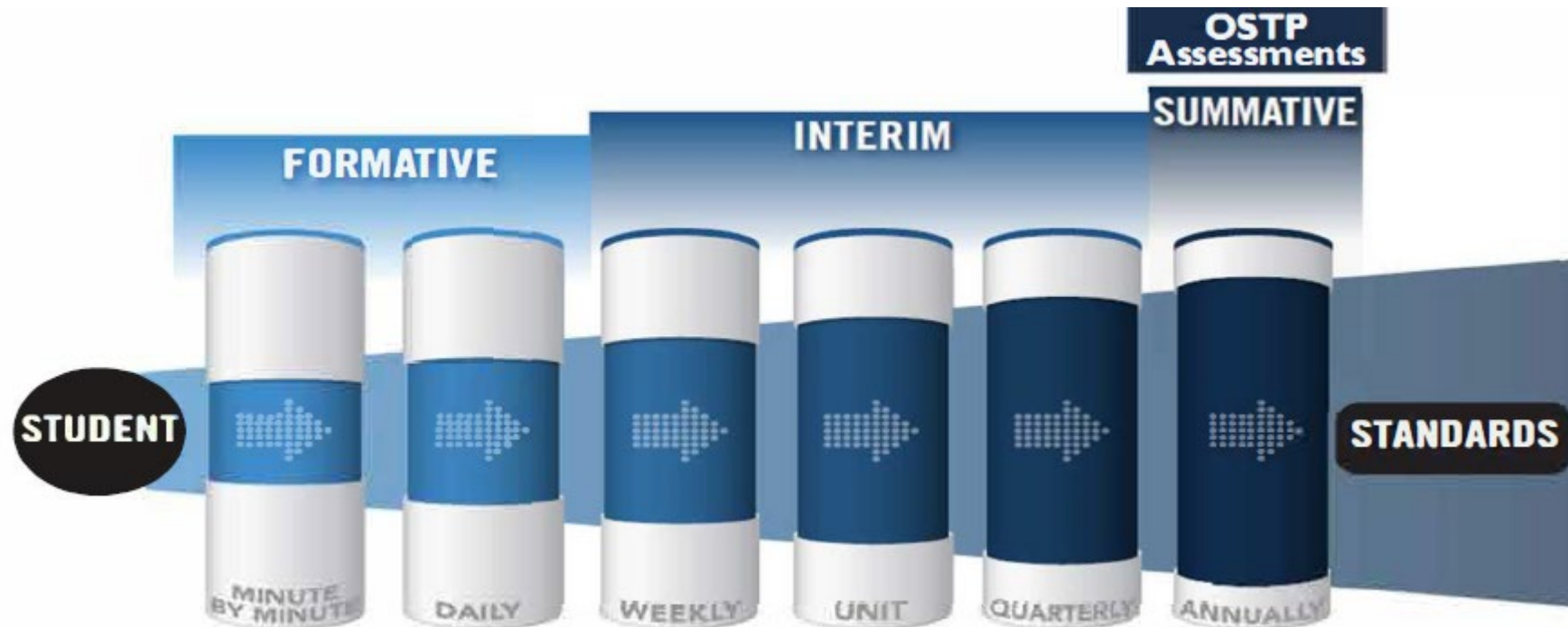
OSTP Science Grade 8 Assessment Overview



Assessment Overview Topics

- c OSTP Science Test Purpose
- c Development and Administration Process
- c The three dimensions of science learning
- c OSTP Science Test Design overview

OSTP Science Test Purpose



“A robust assessment system is predicated upon the knowledge that no one assessment is able to provide answers to all questions affecting instructional decisions. An assessment system utilizes different types of assessment to gather multiple pieces of evidence to provide timely, relevant, actionable, and reliable information about what students know and can do relative to a set of standards.”

OSTP Science Development and Administration Process

- The items developed for the OSTP Science Grade 8 Test are aligned to the Oklahoma Academic Standards-Science (OAS-S).
- The OSTP Science Grade 8 test is administered online. Paper/pencil testing is only provided as a testing accommodation
- The OSTP Science Grade 8 test is separated into two sessions. Districts may exercise flexibility in determining how to administer the sessions. The Grade 8 Science test is meant to be administered in two sessions within one day or on consecutive instructional days. When testing a session, test administrators may give students additional time if they need it, but the additional time is to be given as an extension of that specific testing session.

The OAS-S Standards are 3-dimensional

- Science and Engineering Practices (SEPs)
 - What students are expected to do
- Disciplinary Core Ideas (DCIs)
 - What students are expected to know
- Crosscutting Concepts (CCCs)
 - How students think and connect ideas

Integrating dimensions

- The standards integrate all three dimensions.



Overview – Clusters and Test Design

- Individual items are organized as part of a **cluster**, which consists of 1 stimulus or passage and the 3 items associated with that stimulus. All items in a cluster are aligned to a single OAS-S standard

Grade 8

- The Grade 8 Science Test consists of 15 OP clusters, or a total of 45 OP items.
- The test blueprint targets a specific % of clusters that cover three domains of science: Physical Science, Earth and Space Science, and Life Science

Reporting Category	Percentage	Number of Items	Number of Points
Physical Sciences (8.PS2.1, 8.PS2.2, 8.PS2.3, 8.PS2.4, 8.PS2.5, 8.PS4.1, 8.PS4.3)	33-40%	15-18	16-19
Life Sciences (8.LS1.4, 8.LS1.5, 8.LS3.1, 8.LS3.2, 8.LS4.1, 8.LS4.2, 8.LS4.3, 8.LS4.4, 8.LS4.5, 8.LS4.6)	40-46%	18-21	19-22
Earth and Space Sciences (8.ESS1.1, 8.ESS1.2, 8.ESS1.3)	21-27%	9-12	10-13
Total Operational Test	100%	45	48

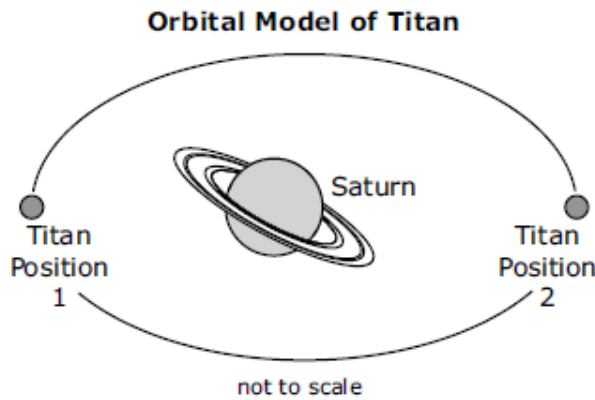
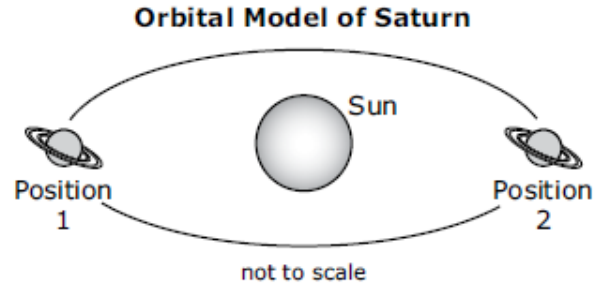
Overview - Item Types

Items developed for the Grade 8 Science test are either multiple choice (MC) items and technology enhanced items (TEIs). A cluster is either a set of three MC items linked to a common stimulus or a set of two MC items and a TEI linked to a common stimulus:

- MC
 - 4 options and 1 key, item is worth 1 point
- TEIs—worth 2 pts, depending on the TEI students may be able to receive 1 pt credit based on scoring notes
 - Drag and Drop (dragging an option into a chart or graphic)
 - Hotspot (clicking on a relevant option in a graphic)
 - Ordering (arranging options in the correct sequence)
 - Inline Choice (select the words to complete a sentence)

A class visits a planetarium where students watch a presentation that models movements in the solar system. During the presentation, the students see the planet Saturn and one of its moons, Titan.

The students want to learn more about Saturn and Titan. They find models of Saturn's orbit and Titan's orbit. The models are shown.



The students also find a table that shows the mass of each object, as shown.

Mass of Three Objects in the Solar System

Object	Mass	Average Distance from the Sun	Average Distance from Saturn
Saturn	5.7×10^{26} kg	1.4×10^9 km	N/A
Sun	2.0×10^{30} kg	N/A	1.4×10^9 km
Titan	1.3×10^{23} kg	1.4×10^9 km	1.2×10^6 km

Example G8 Science Stimulus and Item

A student claims that according to the models, Titan only orbits Saturn and Saturn only orbits the Sun.

Which statement **best** evaluates the student's claim?

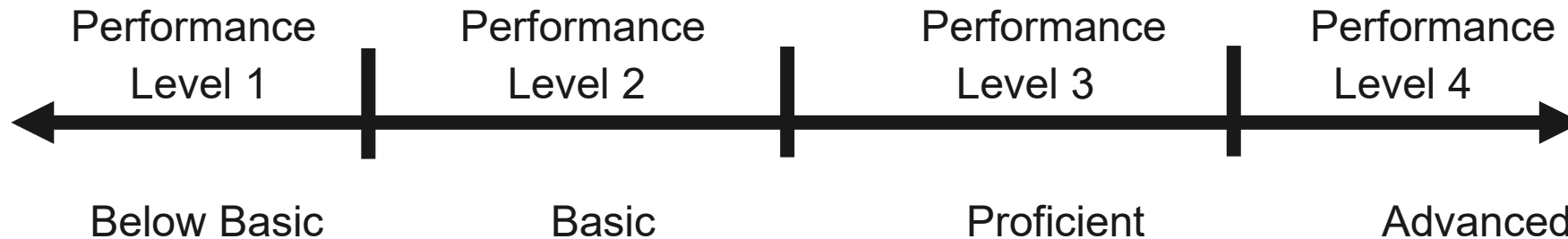
- A** The student is correct because only Saturn is shown orbiting the Sun in the model.
- B** The student is correct because all moons orbit planets and Titan is classified as a moon.
- C** The student is incorrect because all objects in the solar system orbit the Sun because it has the largest mass.
- D** The student is incorrect because Saturn has less mass than the Sun which causes Titan to only orbit the Sun.

Purpose of standard setting

- Allows Oklahoma State Department of Education (OSDE) to have educator expertise inform performance standards for the OSTP Science Grade 8 assessment:
 - Teachers, administrators, higher-ed and vocational specialists.
- Opportunity for educator input on cut scores used to define performance levels
- To ensure recommendations are consistent with expectations stated in the Performance Level Descriptors

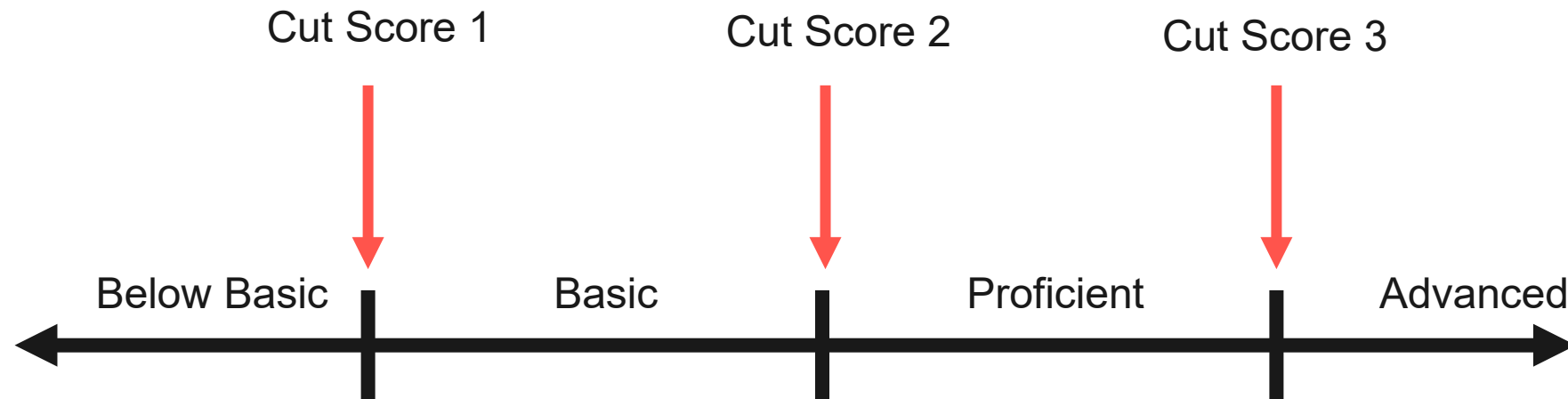
What are performance levels?

- Performance Levels reflect the specific knowledge and skills that a student should be able to demonstrate based on their performance on the test.



What exactly are we doing here?

- What is a Cut Score?
 - A cut score is the minimum test score a student must earn to be considered at a specific performance level.
 - Three cut scores result in four levels of performance.

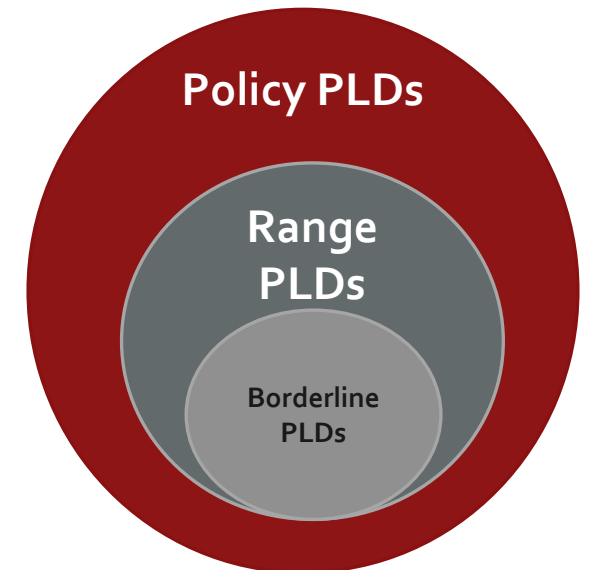


What exactly are we doing here?

- How do we consider cut scores?
 - We don't rely on percentages.
 - They are arbitrary and don't consider the content.
 - We use content-based judgment.
 - Content links assessment items, PLDs, and Performance Standards.
 - Content lets you consider OSDE's objective for students.

Performance Level Descriptors (PLDs)

- PLDs describe the specific knowledge and skills that a student at a given performance level should be able to demonstrate.
 - **Policy PLDs**
 - High-level descriptors that define the knowledge and skill level expectations.
 - **Range PLDs**
 - Content-specific descriptors that link back to the standards.
 - **Borderline PLDs**
 - Specifically define what it takes for a student to attain each performance level, just barely.



Language for Science PLDs

Below Basic	Basic	Proficient	Advanced
<p><i>Students have not demonstrated they can perform at the Basic level.*</i></p> <p>Students scoring at the Below Basic level should be given comprehensive science instruction.</p>	<p><i>Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level.*</i></p> <p>Students scoring at the Basic level typically...</p>	<p><i>Students demonstrate mastery over appropriate grade-level subject matter, and students are ready for the next grade level.*</i></p> <p>Students scoring at the Proficient level typically...</p>	<p><i>Students demonstrate superior performance on challenging subject matter.*</i></p> <p>In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically...</p>

*Policy PLDs



Range PLD Organization

- PLDs are arranged by the Science and Engineering Practices (SEPs).
- Each PLD incorporates the knowledge, skills, and abilities from each PE containing the SEP.
- PLDs are three dimensional and therefore include language from the SEP, DCI (Disciplinary Core Ideas), and CCC (Cross-cutting Concepts).

PLD Example

LS3.1 LS3.2 ESS1.2 ESS1.1	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level.	Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level.	Advanced: Students demonstrate superior performance on challenging subject matter.
Develop and Use Models DCI <ul style="list-style-type: none"> LS3.A Inheritance of Traits LS1.B Growth and Development of Organisms ESS1.A The Universe and Its Stars CCC <ul style="list-style-type: none"> Structure and Function Cause and Effect Systems and System Models 	<div>Standards that include the SEP</div> <div>SEP Grouping</div> <div>DCIs in the above standards</div> <div>CCCs in the above standards</div>	Students scoring at the Basic level typically identify or describe basic components or concept(s) of a model involving the relationship between protein function and gene structure, the effect of reproduction on genetic variation, cyclic patterns in relation to the position of the Earth, Sun, and Moon, or the role of gravity within the solar system.	Students scoring at the Proficient level typically make evaluations about, describe, develop, or use a given model involving the relationship between protein function and gene structure, the effect of reproduction on genetic variation, cyclic patterns in relation to the position of the Earth, Sun, and Moon, or the role of gravity within the solar system.	Students scoring at the Advanced level typically predict, revise, or develop a model from evidence, or apply models involving the relationship between protein function and gene structure, the effect of reproduction on genetic variation, cyclic patterns in relation to the position of the Earth, Sun, and Moon, or the role of gravity within the solar system.

Borderline PLDs

LS3.1 LS3.2 ESS1.1 ESS1.2	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 50% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 50% of the time:
<p>Develop and Use Models DCI</p> <ul style="list-style-type: none"> LS3.A Inheritance of Traits LS1.B Growth and Development of Organisms LS3.B Variation of Traits ESS1.A The Universe and Its Stars <p>CCC</p> <ul style="list-style-type: none"> Structure and Function Cause and Effect Patterns Systems and System Models 		<p>identify or describe basic components or concept(s) of a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.</p>	<p>develop or use a model to describe: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.</p>	<p>evaluate, revise, or predict a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.</p>

Overview of ID Matching



The diagram consists of three identical rectangular boxes arranged horizontally. Each box has a thick red top bar and a light pink body. The text is centered within the pink body. The first box contains 'Item-centered Method', the second 'Content-based Judgment', and the third 'Iterative Process'.

Item-
centered
Method

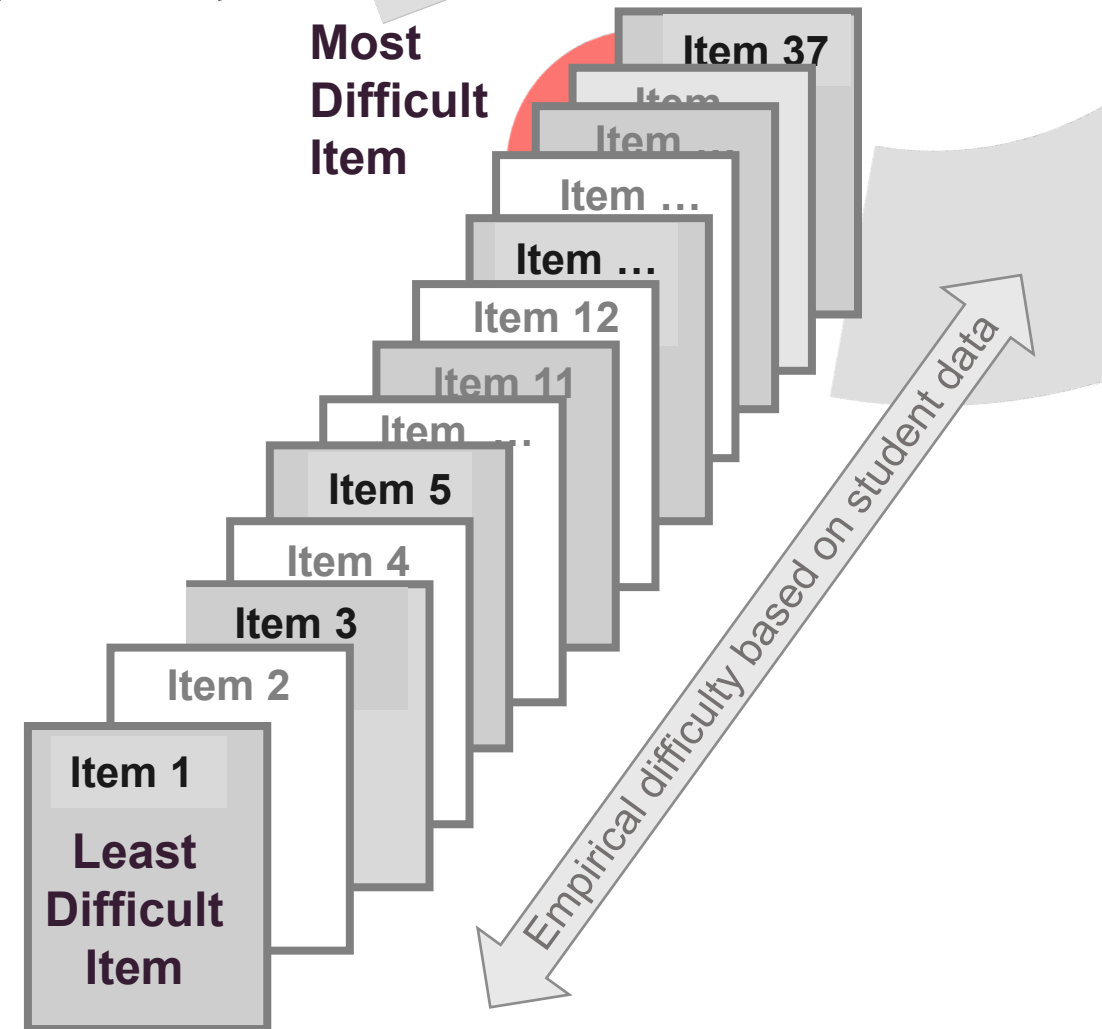
Content-
based
Judgment

Iterative
Process

Ordered Item Booklet* (OIB)

- One item per page
- Easiest item first
- Items ascend by difficulty
- Hardest item last

*The order of the OIB items is based on their empirical difficulties and not the order in which they appear for students during the test.



Overview of ID Matching Method

- Panelists review each item in the OIB.
 - Identify the knowledge, skills, and abilities (KSAs) required to answer the item correctly.
- For each item, make the following judgment:
 - Match the knowledge, skills, and abilities (KSAs) required by the item with the expectations described in either the Basic, Proficient, or Advanced performance level descriptor (PLD).
- Judgements are made independently

Your Judgmental Task

- For each item: Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?
 - Use range and borderline PLDs to make a judgment about items in the ordered item booklet (OIB).
 - Consider the Standards described in the PLDs.
 - Consider the knowledge and skills demand of an item.
- You will engage in three rounds of the ID-Matching judgmental task.

Panelists Judgments: 3 rounds

- Panelists will complete three rounds of the ID-Matching judgments.
 - Round 1: Align items to PLDs, discuss items with panelist disagreement
 - Round 2: Introduce benchmark, align items to PLDs, discuss items with panelist disagreement
 - Round 3: Align items to PLDs and write group-level content-based rationales for cut scores if necessary.

Content-based Judgment - Overview



Good

- Based on Content
- Links items to PLDs
- Refers to specific knowledge, skills, and abilities (KSAs)



Bad

- Based on something other than the content
- Too general
- Based on a specific student or class

Content-Based Benchmarks - Overview

- Benchmarks based on Cognia content team judgements
 - Benchmarks will be presented to you at the beginning of Round 2.
 - The benchmark region represents a likely transition between two levels based on those judgments.
 - Benchmarks serve as a guide.
 - You may consider the benchmark as you engage in Round 2 judgment.
- More detailed information/training to come later today

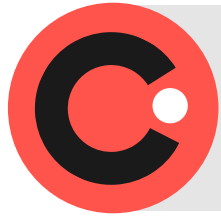


OSTP Science Standard Setting – Day 1

Standard Setting – Day 1

Facilitator: David Harrison

Content specialist: Mary-Alice Corliss



Day 1 - Agenda

- Welcome and Introductions – Panelists
- Access to the Cognia Standard Setting Toolkit
- Familiarization with PLDs and Content Standards
- Experience the OSTP Science Grade 8 Test
- Lunch
- Training on the Item-Descriptor (ID) Matching Method
- Modeling and Practice
- Begin Round 1

Welcome & Introductions - Panelists



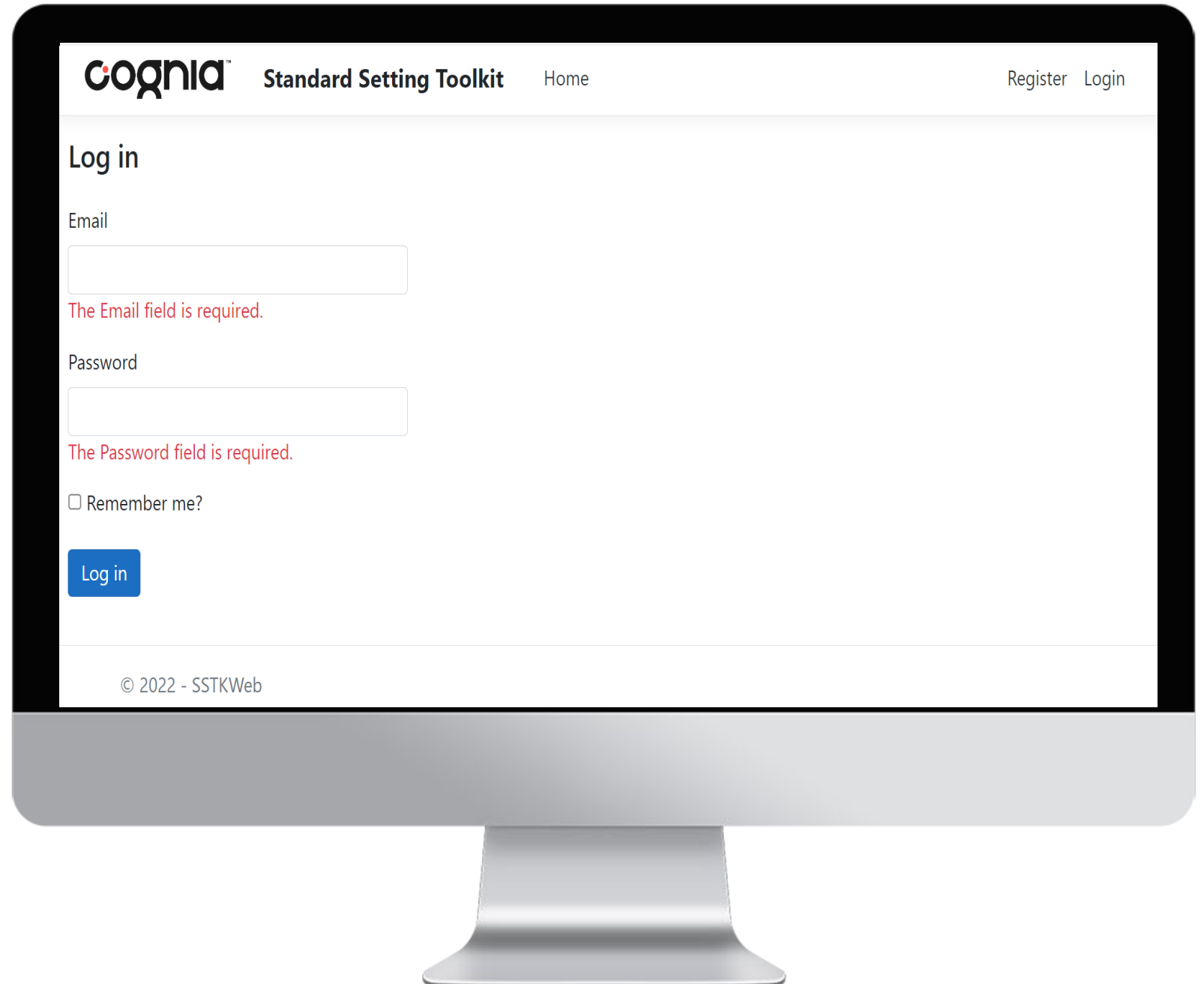
- Introduce yourself
 - Your name, school district, what you teach
 - A little selected background information
- Show of hands
 - Who's been involved in standard setting before?
 - Which method(s)?

Meeting Norms

- All conversations are confidential
 - What happens here stays here.
 - When you return to your state, please do talk about the process we undertake, but do not disclose the specifics.
 - Do not discuss item specifics outside of the panels or after standard setting.
- Please DO NOT
 - Use personal devices in the room - you may step out at any time if needed.
 - Use the Chromebooks for anything other than the standard setting activities

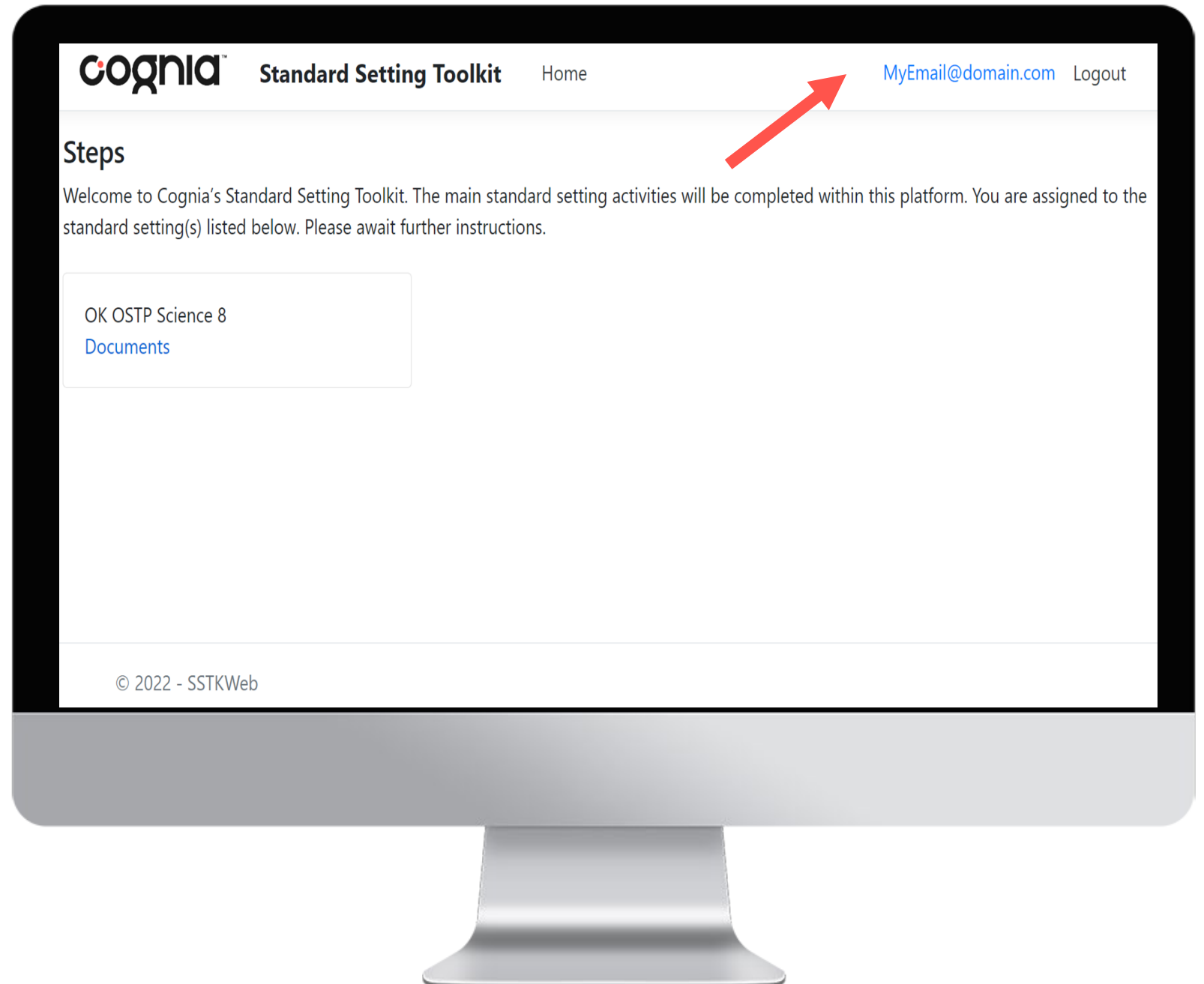
Cognia Standard Setting Toolkit

- Use **your email** and **initial password** to log on to the platform
- Email: Your own email that was used to register for this meeting
- Password: Everyone has the same initial password



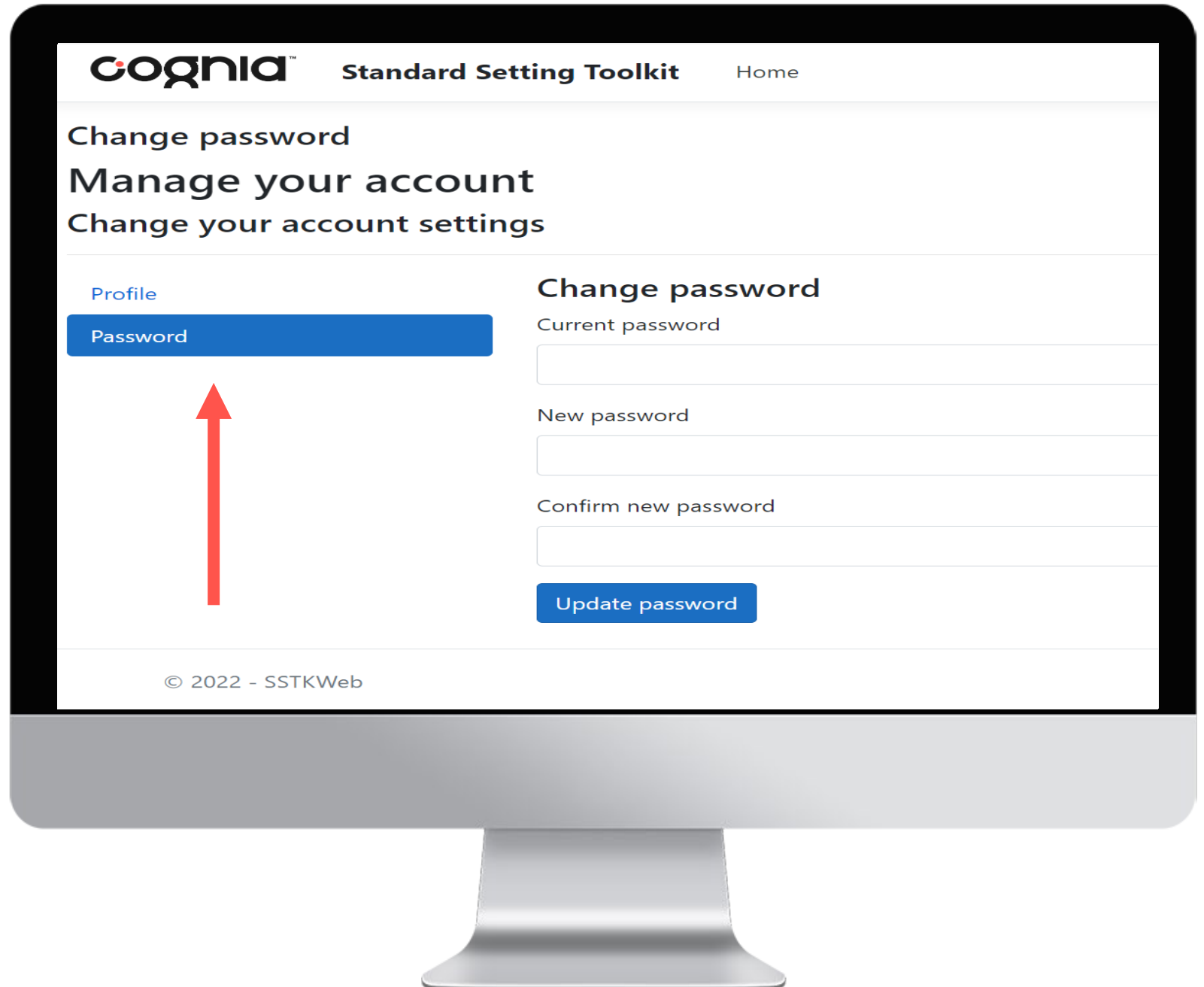
Cognia Toolkit: Change Your Password

- Click on your email in the top right corner
- This will bring you to a profile page
- Click on the “Password” tab shown to the left



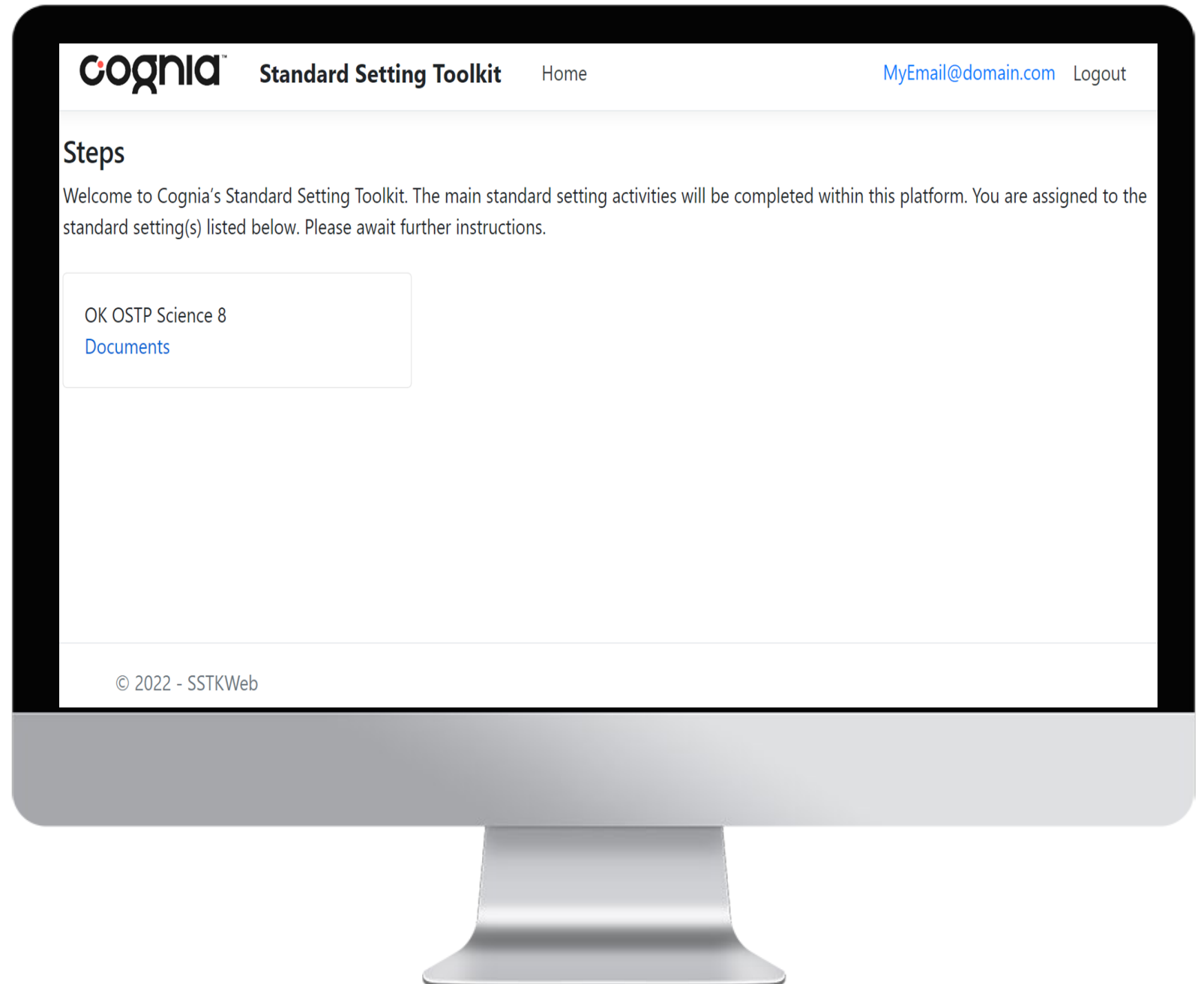
Cognia Toolkit: Change Your Password

- Click on the “Password” tab shown to the left
- Enter the initial password
- Enter New Password:
 - Upper Case Letter
 - Lower Case Letter
 - Number
 - At least 6 Characters
- Click “Update Password”
- Log out and Log back in with your new password.



The screenshot displays the Cognia Standard Setting Toolkit interface. At the top, the Cognia logo is on the left, and 'Standard Setting Toolkit' and 'Home' are on the right. The main heading is 'Change password', followed by 'Manage your account' and 'Change your account settings'. On the left sidebar, there are two tabs: 'Profile' and 'Password'. The 'Password' tab is selected and highlighted with a blue background. A red arrow points upwards to this tab. The main content area on the right is titled 'Change password' and contains three input fields: 'Current password', 'New password', and 'Confirm new password'. Below these fields is a blue button labeled 'Update password'. At the bottom of the page, the copyright notice '© 2022 - SSTKWeb' is visible.

You Should
Now Be Back on
the Following
Screen:





Familiarization OSTP Science

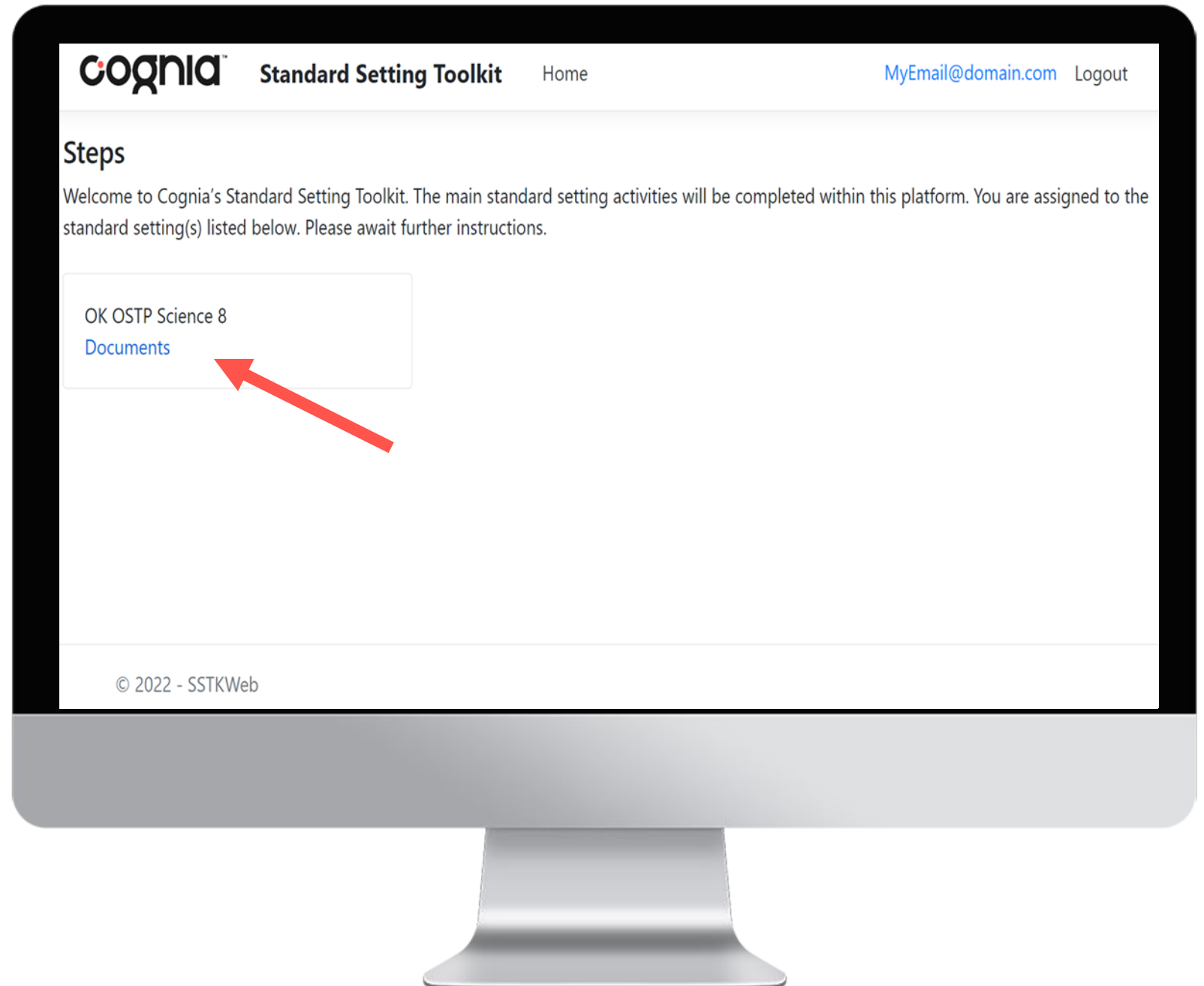
Content Standards and Performance Level
Descriptors (PLDs)

Review PLDs

- Brief Background on PLD development
- Obtain an understanding of PLDs in relation to Content Standards.
 - This activity is critical because you will make judgements based on your understanding of PLDs.
 - The PLD documents will be used throughout the workshop to make item-PLD alignment.
- Individually review PLDs within the Cognia Standard Setting Toolkit

Where to find the PLDs and Standards

- In the Toolkit: Blue “documents” link
- Document links for the Standards, Range, and Borderline PLDs appear
- Paper copies of the PLDs also distributed for easy reference

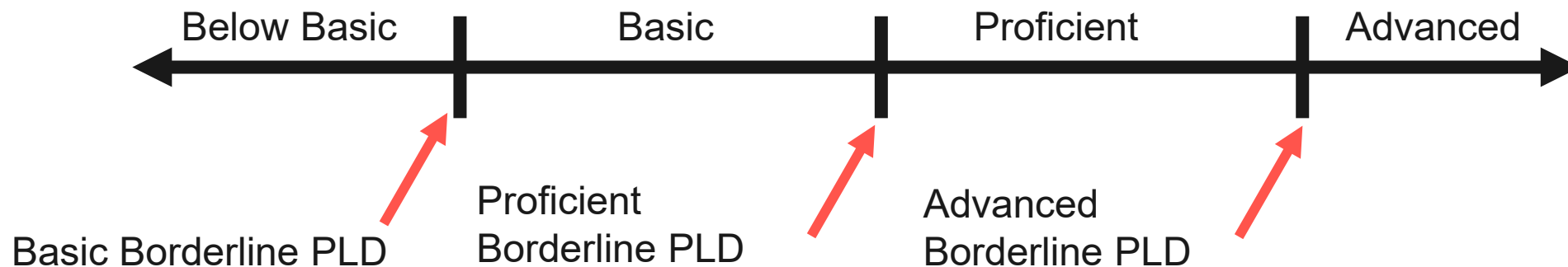


Discuss Range PLDs

- Collegial discussion to clarify questions
- Reach common understanding of what it means to be in each performance level.
 - Start with Basic PLD, then Proficient and Advanced.
 - Focus on how the levels differ in content, cognitive complexity.
 - Discuss Below Basic PLD as an extension of Basic PLD.

Discuss Borderline PLDs

- Borderline PLDs describe the knowledge, skills and abilities (KSAs) expected of students who just barely meet each performance level.
 - Draw similar connection between performance expectations and borderline students who can demonstrate a level of KSAs that is barely pass the entry point for a given performance level





Experience the Test

Experience the Test

- You will experience the OSTP Science Grade 8 test in a format that is similar to student experience.
 - Briefly examine the test items in the testing platform.
 - Try not to linger on any one item; this session is scheduled for a duration of **one hour**.
- Purpose:
 - Get familiar with the items as they appeared to students.
 - Science items sets appear together in the testing platform but do not appear together in the OIB.
 - You will see most of the items from the testing platform in the OIB.

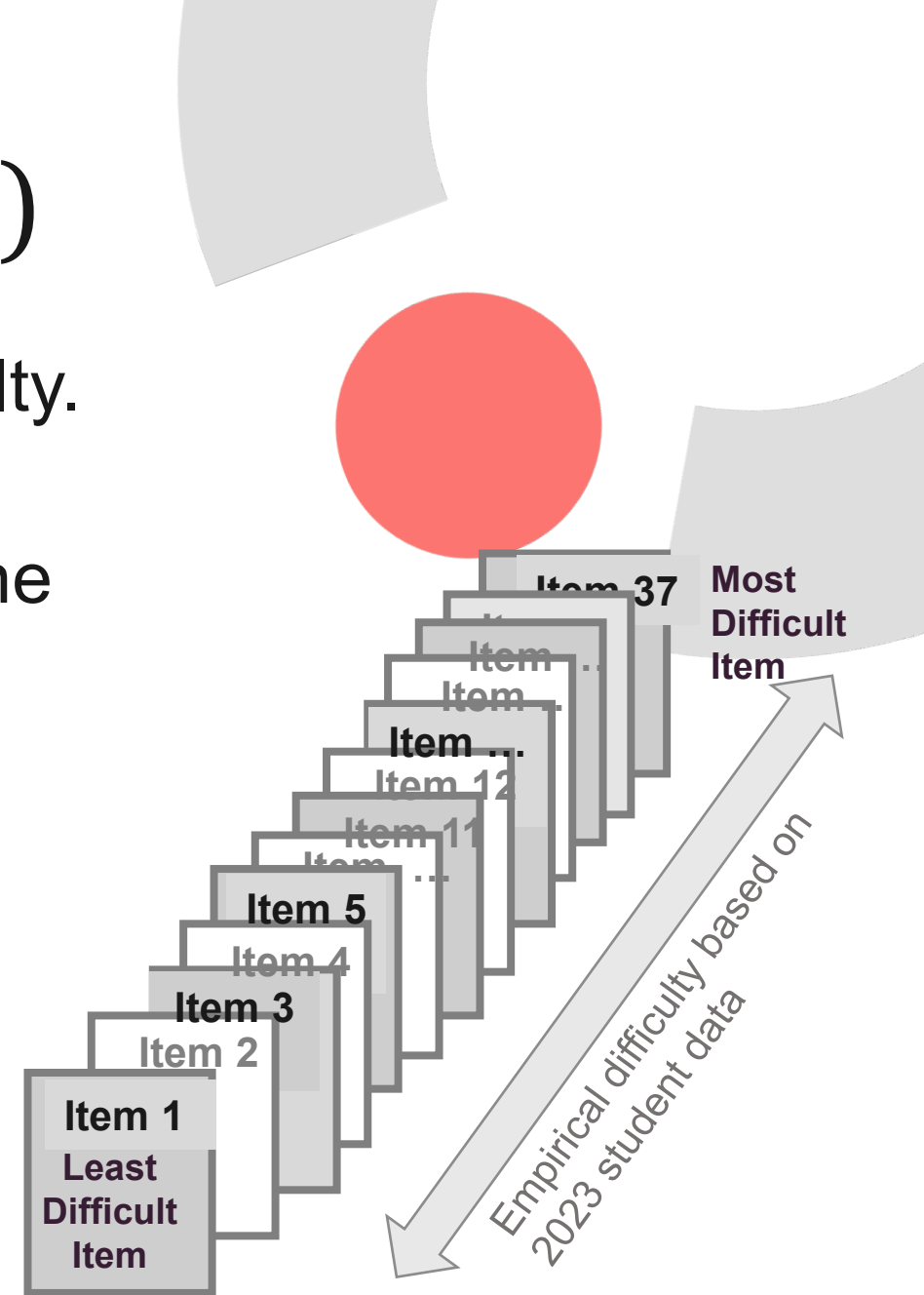


Key Concepts & Process

Ordered Item Booklet (OIB) and Item-Descriptor Matching (ID-Matching)

Ordered Item Booklet (OIB)

- OIB contains test items ordered by difficulty.
- Each OIB page represents an item.
- The difference in difficulty is not exactly the same between each pair of neighboring items.
- Difficulty is based on data from the AY23 OSTP students who took the test
- 2-Point items:
 - Will appear twice in the OIB – once for each point



OIB in the Toolkit

Item Review













OK OSTP Science 8 Step 3 Round 1 Judgements

Hide Documents

Standards

Range PLDs

Borderline PLDs

Position	Asset ID	Description	Point Value	Relevant KSAs	Notes	Item Descriptor Match Level	
1	638903	CL03_PS2-1_TEI1	1			-- ▼	Detail
2	494991	Trilobites and Ammonites	1			-- ▼	Detail
3	788060	Sharks and Dolphins_LS4-3_MC2	1			-- ▼	Detail
4	300154A	Whale Ancestors	1			-- ▼	Detail
5	188317A	Trilobites and Ammonites	1			-- ▼	Detail
6	788002	Coconuts_LS1-4_MC1	1			-- ▼	Detail

ID Matching: Your Judgmental Task

- Review each item in the OIB
- For each item: Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?
 - Use range and borderline PLDs to make a judgment about items in the ordered item booklet (OIB).
 - Consider the Standards described in the PLDs.
 - Consider the knowledge and skills demand of an item.
- As you review items, write down brief content-based reasons for your item-PLD matches
- If an item seems to be aligned in the border between two PLDs, select the PLD that most closely matches the KSAs AND write notes about the item to later inform discussions

From Judgments to Cut Scores

1. You are presented with Items ordered from least to most difficult in the OIB based on student data

2. You will proceed through the items in order of difficulty and make the following judgment:

Match knowledge, skills, and abilities required by an item with a Performance Level Descriptor

3. As you go, you will

- Write content-based reasons for your judgement
- Note when an item seems to align between two PLDs

4. At the completion of the round, we will feed all the item-PLD matches from every panelist into an analysis to calculate three threshold regions.

OIB	PLD Alignment
Item 15	Basic
Item 16	Basic
Item 17	Basic
Item 18	Proficient
Item 19	Basic
Item 20	Proficient
Item 21	Proficient
Item 22	Basic
Item 23	Proficient
Item 24	Proficient
Item 25	Proficient
Item 26	Proficient
Item 27	Proficient

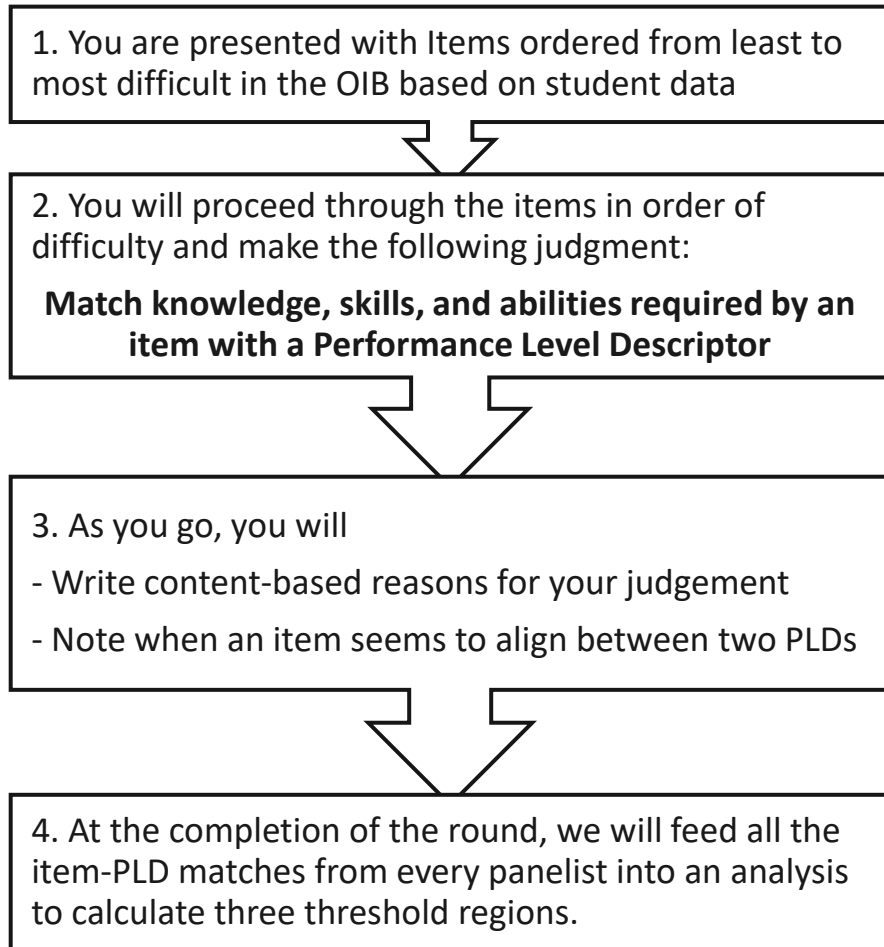
Summarize across all panelists using statistics.

OIB
Item 15
Item 16
Item 17
Item 18
Item 19
Item 20
Item 21
Item 22
Item 23
Item 24
Item 25
Item 26
Item 27

Panelist item-PLD Alignment **Transition Region**

Calculated item-PLD Alignment **Threshold Region** for the entire group

From Judgments to Cut Scores (Cont.)



OIB	PLD Alignment
Item 15	Basic
Item 16	Basic
Item 17	Basic
Item 18	Proficient
Item 19	Basic
Item 20	Proficient
Item 21	Proficient
Item 22	Basic
Item 23	Proficient
Item 24	Proficient
Item 25	Proficient
Item 26	Proficient
Item 27	Proficient

Summarize across all panelists using statistics.

OIB
Item 15
Item 16
Item 17
Item 18
Item 19
Item 20
Item 21
Item 22
Item 23
Item 24
Item 25
Item 26
Item 27

- **Threshold region:** Area where items most likely tip from one PLD level to the next
- We will calculate 3 threshold regions: Basic, Proficient, and Advanced
- After each round the regions will shrink
- At the conclusion of Round 3, we will pinpoint the specific cut scores (points in the OIB) for the Basic, Proficient, and Advanced cuts.



Modeling and Practice

Preparation for Round 1

Modeling & Practice of the ID-Matching Judgmental Task

- We will look at 3 sample items
- For each item: Answer the following question:
 - What does a student need to know or be able to do to respond to this item?
- Match each item to a PLD
 - Explain how the item response demands align with expectations described in PLDs.

A reminder: Content-based Judgments



Good

- Based on Content
- Links items to PLDs
- Refers to specific knowledge, skills, and abilities (KSAs)



Bad

- Based on something other than the content
- Too general
- Based on a specific student or class

Content-based Judgment - Examples

- A good example:
 - The item require XYZ; XYZ are described in the Proficient PLD and not in the Basic PLD.
- A bad example:
 - The items match the Proficient PLD and do not match the Advanced PLD.

Practice Round Process



Complete item-PLD alignment task for 3 sample items.



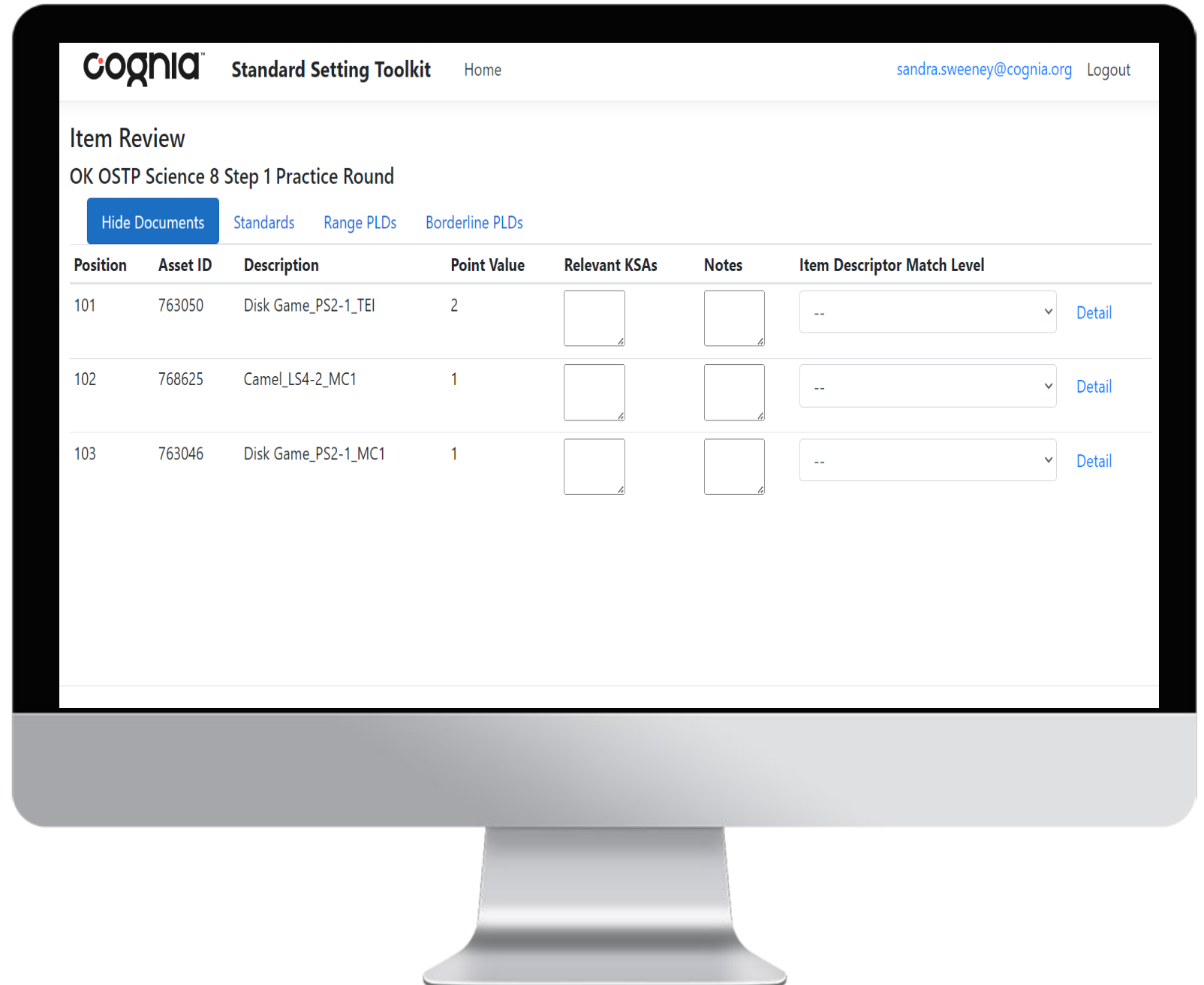
Discuss matches.



Discuss and clarify range and borderline PLDs as needed.

Practice Round

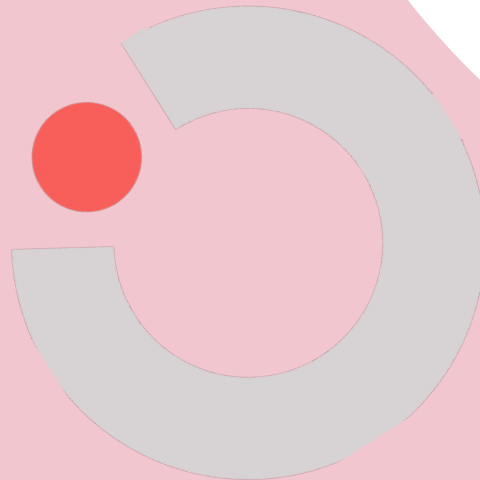
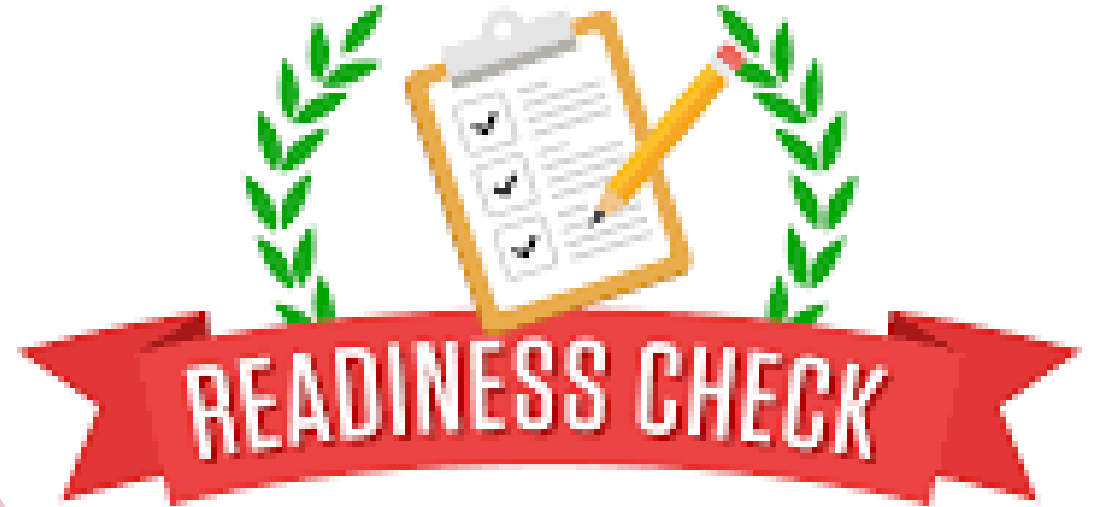
- In the Toolkit you will automatically be redirected to the practice round
- You will see 3 practice items



Practice Round - Review

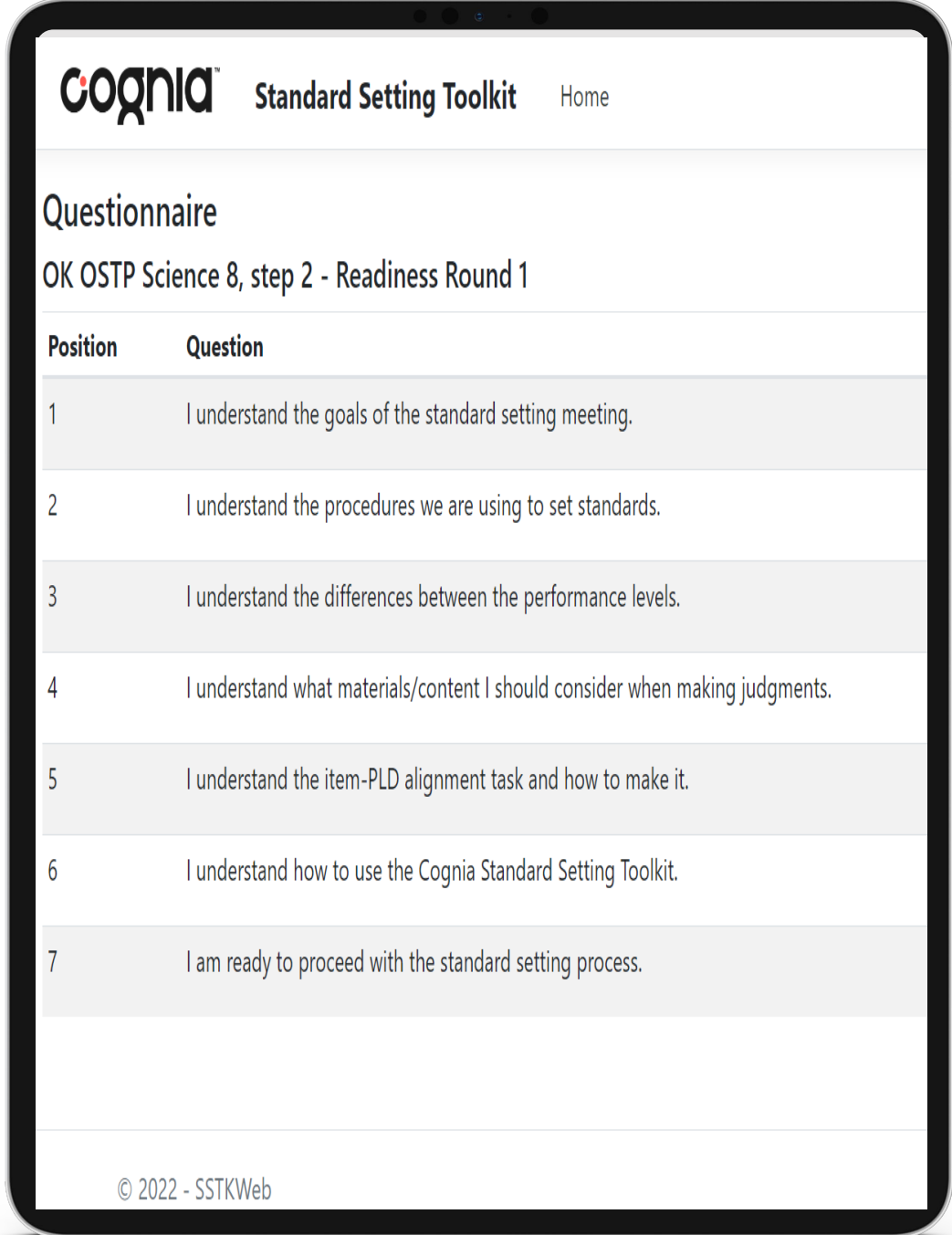
- Reviewed three sample items and for each one:
 - Reviewed the item
 - Considered the Knowledge, Skills, and Abilities required by the item
 - Matched the item to either the Basic, Proficient, or Advanced PLD
- Borderline considerations
 - Some items seem to be in the border between two adjacent PLDs
 - Select the PLD that most closely matches the item
 - Make notes for yourself next to these items to inform discussions later
- Remaining questions or concerns?

Complete Round 1 Readiness Survey



Round 1 – Readiness Survey

- In a moment, you will be redirected in the Toolkit to a short survey
- Goal: Determine if everyone understands the task at hand and is ready to proceed
- Read each question and answer yes/no
- Once everyone has completed the survey, we will review responses and proceed accordingly.

The image shows a tablet displaying the Cognia Standard Setting Toolkit interface. At the top, the Cognia logo is on the left, and 'Standard Setting Toolkit' and 'Home' are on the right. Below the header, the title 'Questionnaire' is displayed, followed by the subtitle 'OK OSTP Science 8, step 2 - Readiness Round 1'. A table with two columns, 'Position' and 'Question', lists seven items. Each item is on a separate row with alternating light gray and white backgrounds. At the bottom of the screen, the copyright notice '© 2022 - SSTKWeb' is visible.

Position	Question
1	I understand the goals of the standard setting meeting.
2	I understand the procedures we are using to set standards.
3	I understand the differences between the performance levels.
4	I understand what materials/content I should consider when making judgments.
5	I understand the item-PLD alignment task and how to make it.
6	I understand how to use the Cognia Standard Setting Toolkit.
7	I am ready to proceed with the standard setting process.

© 2022 - SSTKWeb

Round 1 Judgments

- You will now be redirected to Round 1
 - In the toolkit you will see the full list of items
- Reminder: Your task
 - Review each item
 - Consider the KSAs and match the item to one of the PLDs
 - Write content-based reasons in the “KSAs” box as you go
 - Use the “Notes” box for additional notes (for example: when an item seems to be in-between two PLDs)
- Item-PLD alignment is an individual activity – please do not discuss your work with your colleagues at this time.



OSTP Science Standard setting – Day 2

Standard Setting – Day 2

Facilitator: David Harrison

Content specialist: Mary-Alice Corliss



Day 2 - Agenda

- Debrief Day 1
- Complete Round 1 Judgments
- Feedback and Discussion of Round 1 Results
- Complete Round 2 Judgments
- Feedback and Discussion of Round 2 Results
- Complete Round 3 Judgments
- Final Workshop Evaluation Survey

Round 1 Judgments - Continue



- You will now be redirected to Round 1
 - In the toolkit you will see the full list of items
- Reminder: Your task
 - Review each item
 - Consider the KSAs and match the item to one of the PLDs
 - Write content-based reasons in the “KSAs” box as you go
 - Use the “Notes” box for additional notes (for example: when an item seems to be in-between two PLDs)
- Item-PLD alignment is an individual activity – please do not discuss your work with your colleagues at this time.



Feedback/Discussion of Round 1 Results



Preparation for Round 2

Introduce Benchmarks

- Content-based information from Cognia content specialists
- Benchmarks serve as additional information for your consideration.
- Will be presented as shaded rows in the OIB
 - Yellow → Basic Region
 - Green → Proficient Region
 - Blue → Advanced Region

























Content-Based Benchmarks: Visual Presentation

SSTKWeb Home Admin ▾

sandra.sweeney@cognia.org Logout

Item Review

1. OK OSTP Science 8 Step 5 Round 2 Judgements

Position	Asset ID	Description	Point Value	Relevant KSAs	Notes	Item Descriptor Match Level	
1	638903	CL03_PS2-1_TEI1	1			-- ▾	Detail
2	494991	Trilobites and Ammonites	1			-- ▾	Detail
3	788060	Sharks and Dolphins_LS4-3_MC2	1			-- ▾	Detail
4	300154A	Whale Ancestors	1			-- ▾	Detail
5	188317A	Trilobites and Ammonites	1			-- ▾	Detail
6	788002	Coconuts_LS1-4_MC1	1			-- ▾	Detail
7	300160A	Whale Ancestors	1			-- ▾	Detail
8	638901	CL03_PS2-1_MC2	1			-- ▾	Detail
9	788146	Eclipse_ESS1-1_B_TEI_2	1			-- ▾	Detail
10	494991	Trilobites and Ammonites	2			-- ▾	Detail
11	638899	CL03_PS2-1_MC1	1			-- ▾	Detail
12	300153A	Whale Ancestors	1			-- ▾	Detail

Content-Based Benchmarks

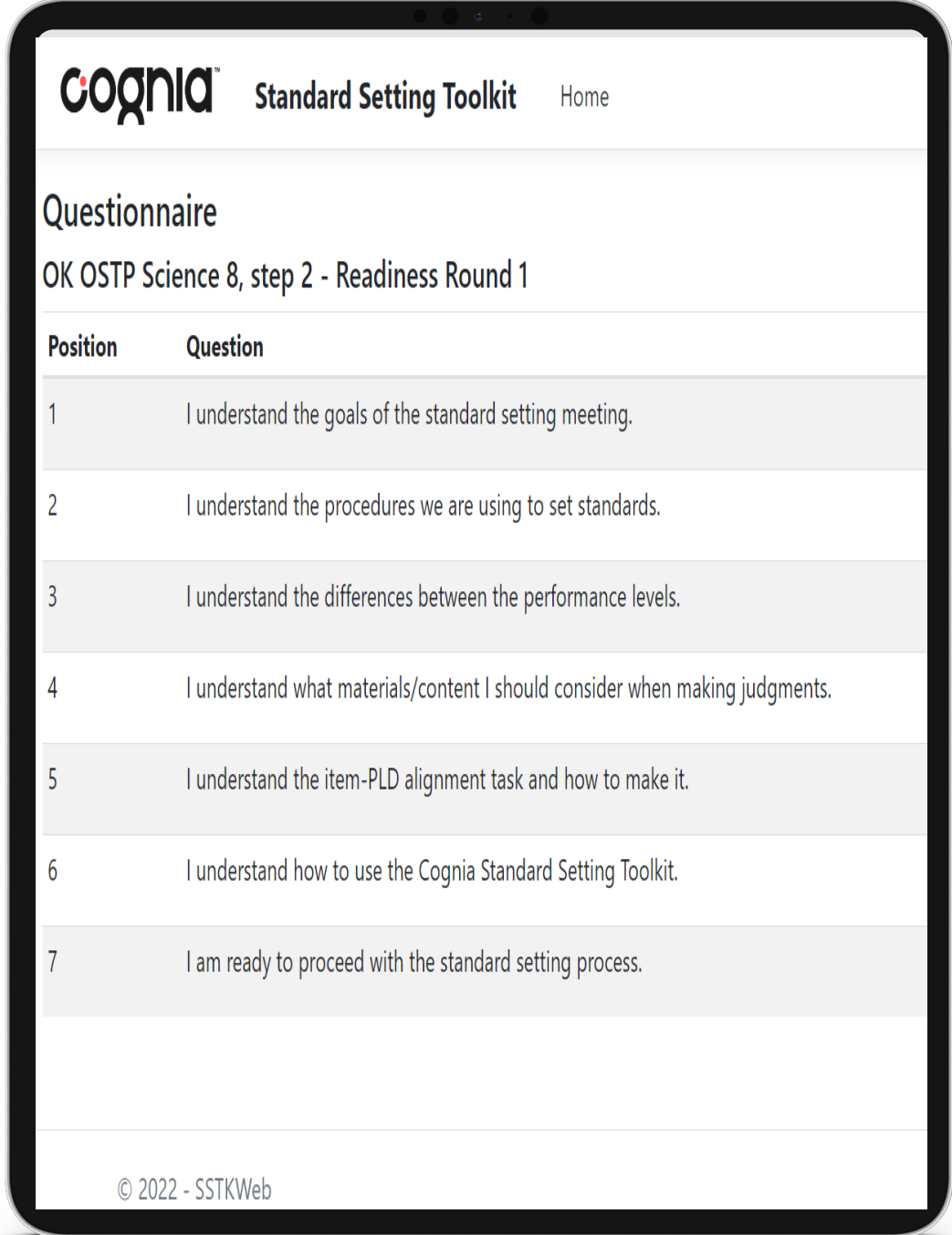
- The shaded regions are calculated based on judgments from other content experts like yourselves.
- This region represents a likely transition between where they were aligning content between two adjacent levels.
- The experts making those judgments are Cognia item writers.
- It is vital that we have the input of educators who teach to these standards and population.
- To that end, your results may very well differ from theirs.
- The content-based benchmarks provide additional information for your consideration but is not meant to constrain or persuade your judgements



Complete Round 2 Readiness Survey

Round 2 – Readiness Survey

- In a moment, you will be redirected in the Toolkit to a short survey
- Goal: Determine if everyone understands the task at hand and is ready to proceed
- Read each question and answer yes/no
- Once everyone has completed the survey, we will review responses and proceed accordingly.



The screenshot shows a web application titled "Cognia Standard Setting Toolkit" with a "Home" link. The main heading is "Questionnaire" followed by "OK OSTP Science 8, step 2 - Readiness Round 1". Below this is a table with two columns: "Position" and "Question". The table contains seven rows of questions. At the bottom, there is a copyright notice: "© 2022 - SSTKWeb".

Position	Question
1	I understand the goals of the standard setting meeting.
2	I understand the procedures we are using to set standards.
3	I understand the differences between the performance levels.
4	I understand what materials/content I should consider when making judgments.
5	I understand the item-PLD alignment task and how to make it.
6	I understand how to use the Cognia Standard Setting Toolkit.
7	I am ready to proceed with the standard setting process.

© 2022 - SSTKWeb

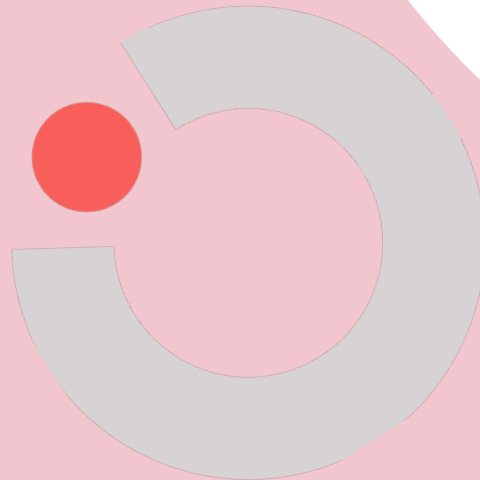
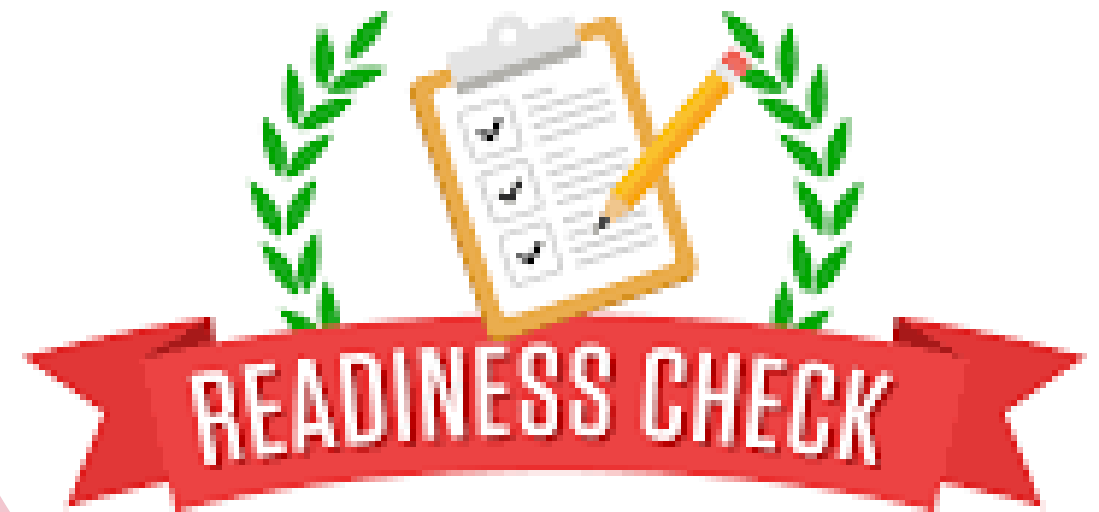
Round 2 Judgments

- You will now be redirected to Round 2
 - In the toolkit you will see the same full list of items with your work from round 1 (notes and judgments)
 - You will also see the shaded regions for the content-based benchmarks
- Reminder: Your task
 - Review items in the benchmark regions and items you were previously unsure about
 - Consider the KSAs and decide to keep or change your initial PLD Match
- Item-PLD alignment is an individual activity – please do not discuss your work with your colleagues at this time.



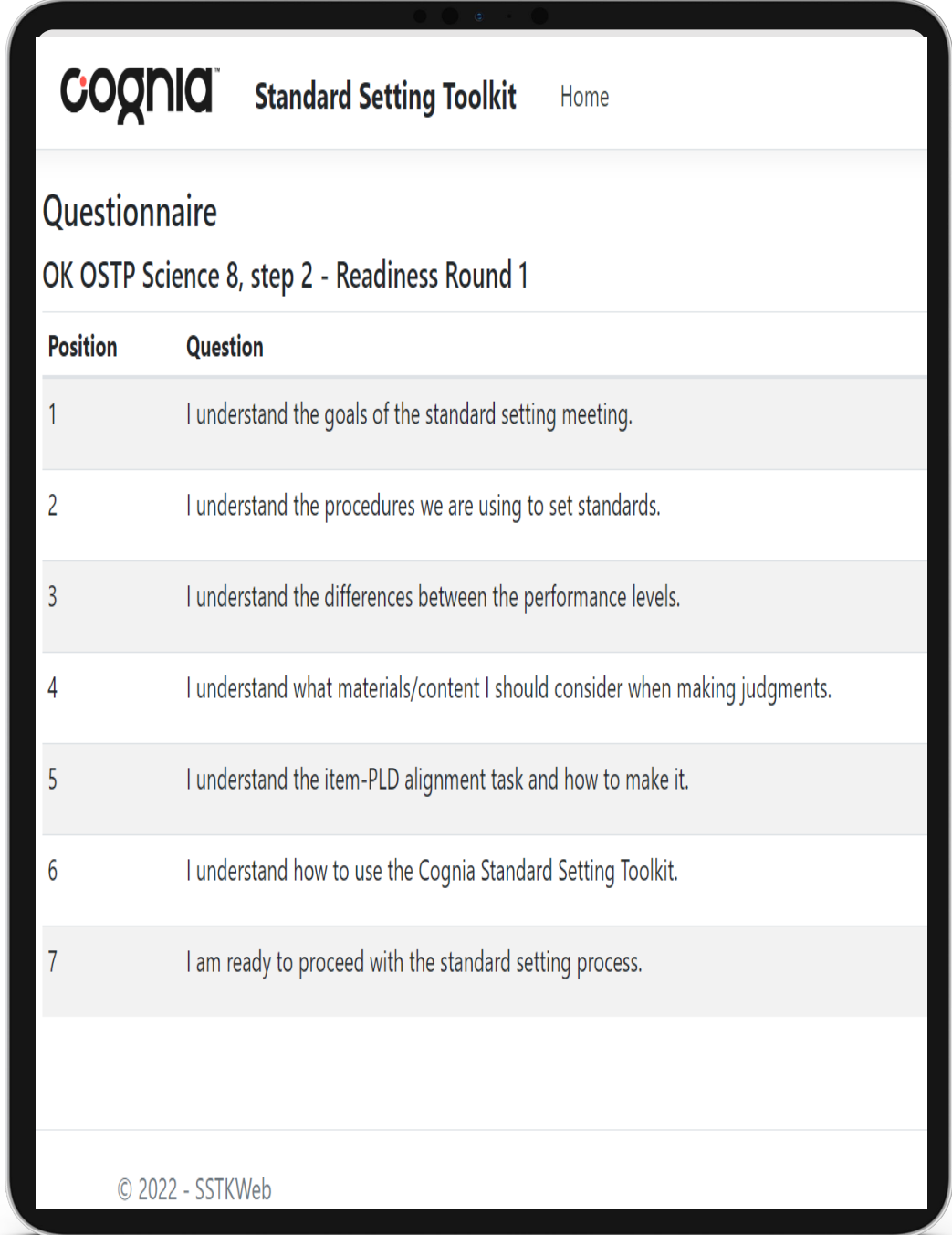
Feedback/Discussion of Round 2 Results

Complete Round 3 Readiness Survey



Round 3 – Readiness Survey

- In a moment, you will be redirected in the Toolkit to a short survey
- Goal: Determine if everyone understands the task at hand and is ready to proceed
- Read each question and answer yes/no
- Once everyone has completed the survey, we will review responses and proceed accordingly.

The image shows a tablet displaying the Cognia Standard Setting Toolkit interface. At the top, the Cognia logo is on the left, and 'Standard Setting Toolkit' and 'Home' are on the right. Below the header, the title 'Questionnaire' is displayed, followed by 'OK OSTP Science 8, step 2 - Readiness Round 1'. A table with two columns, 'Position' and 'Question', lists seven items. The questions are: 1. I understand the goals of the standard setting meeting. 2. I understand the procedures we are using to set standards. 3. I understand the differences between the performance levels. 4. I understand what materials/content I should consider when making judgments. 5. I understand the item-PLD alignment task and how to make it. 6. I understand how to use the Cognia Standard Setting Toolkit. 7. I am ready to proceed with the standard setting process. At the bottom, the copyright notice '© 2022 - SSTKWeb' is visible.

cognia Standard Setting Toolkit Home

Questionnaire

OK OSTP Science 8, step 2 - Readiness Round 1

Position	Question
1	I understand the goals of the standard setting meeting.
2	I understand the procedures we are using to set standards.
3	I understand the differences between the performance levels.
4	I understand what materials/content I should consider when making judgments.
5	I understand the item-PLD alignment task and how to make it.
6	I understand how to use the Cognia Standard Setting Toolkit.
7	I am ready to proceed with the standard setting process.

© 2022 - SSTKWeb

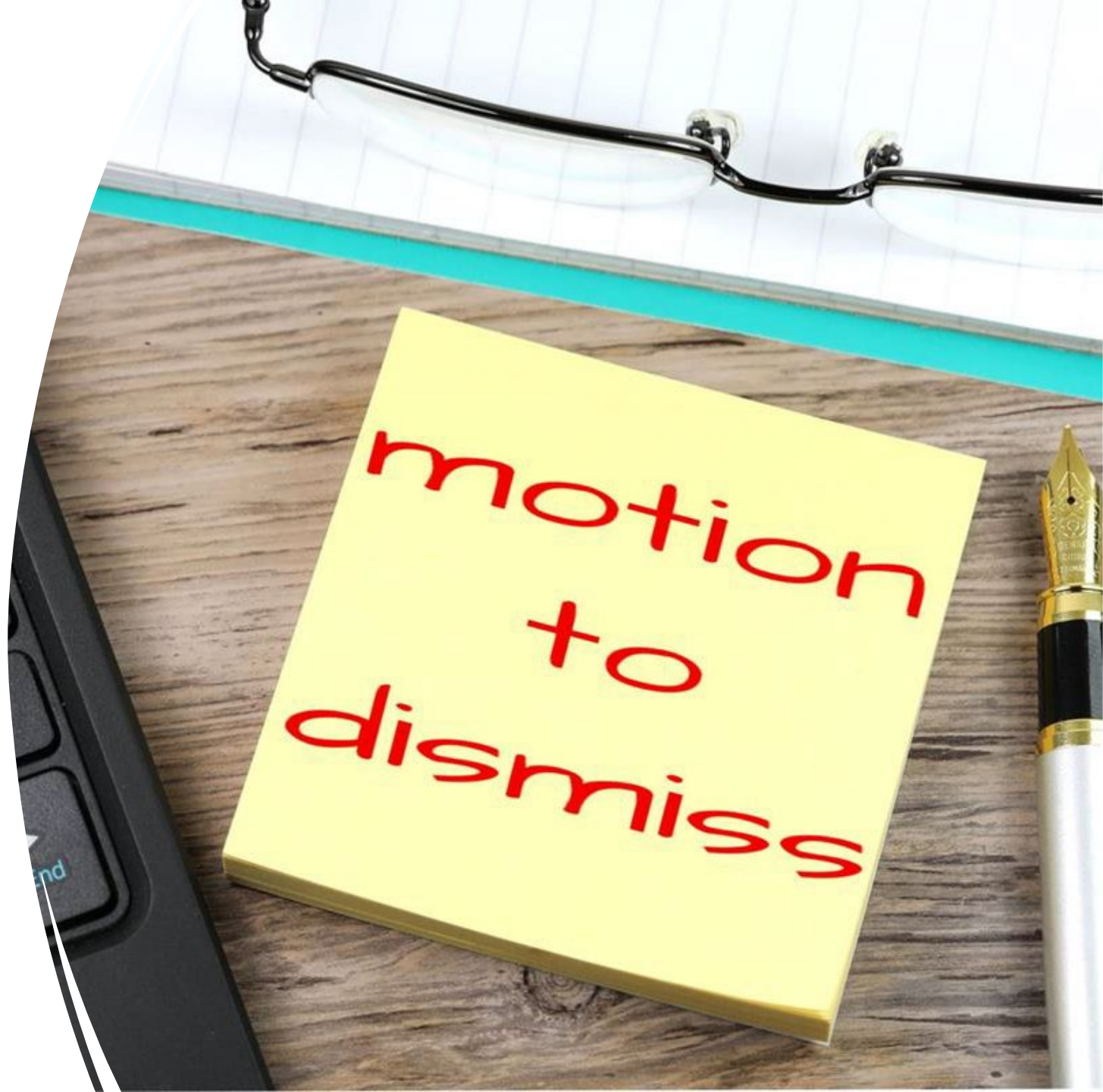
Round 3 Judgments

- You will now be redirected to Round 3
 - In the toolkit you will see the same full list of items with your work from round 2 (notes and judgments)
 - You will also see the shaded regions for the content-based benchmarks
- Reminder: Your task
 - Review items in the benchmark regions and items you were previously unsure about
 - Consider the KSAs and decide to keep or change your initial PLD Match
- Item-PLD alignment is an individual activity – please do not discuss your work with your colleagues at this time.

Complete workshop Evaluation Survey



Dismiss



APPENDIX—D
MEETING AGENDA

Oklahoma School Testing Program Standard Setting Meeting

Science Grade 8

Meeting Agenda

Day 1: June 22, Thursday

Time	Activity/Session
08:00 – 09:00	Registration and Breakfast
09:00 – 09:45	Orientation: Introductions and overview: Welcome, workshop goals, OSTP Science exam; standard setting, the ID Matching method
09:45 – 11:15	Review range and borderline PLDs, content standards (brief)
11:15 – 12:00	Experience the Test
12:00 – 01:00	Lunch
01:00 – 02:30	Training on the ID Matching method; Practice: Facilitator models cognitive-judgment task; Panelist practice and discussion; Prepare for round 1: Complete readiness survey
02:30 – 04:00	Begin Round 1
04:00	Adjourn for the day.

Day 2: June 23, Friday

Time	Activity/Session
08:00 – 08:45	Check-in and Continental breakfast
08:45 – 10:00	Complete Round 1 Judgments.
10:00 – 10:15	Break and Analysis of Round 1 data
10:15 – 11:15	Feedback and discussion of round 1 results.
11:15 – 12:00	Prepare for Round 2: Introduce benchmarks and complete readiness survey
12:00 – 1:00	Lunch
01:00 – 02:00	Complete round 2 Judgments
02:00 – 02:45	Feedback and discussion of round 2 results. Prepare for round 3: complete readiness survey
02:45 – 03:00	Break
03:00 – 03:30	Complete round 3
03:30 – 04:00	Review final results; Impact data; Complete workshop evaluation; Dismissal

APPENDIX—E
NONDISCLOSURE
AGREEMENT



Nondisclosure Agreement

Oklahoma State Testing Program
Science Standard Setting
June 22-23, 2023

The undersigned is an employee, contractor, assessment committee member, or person otherwise authorized to view secure state assessment materials. The undersigned hereby agrees to be bound to the terms of this agreement restricting the disclosure of said materials.

It is essential to the integrity of this item development project and testing program that all test items remain secure. To maintain this security, only authorized persons are permitted to view the test questions. With the exception of materials released by the Oklahoma State Department of Education for informational purposes, all test questions (draft or final) in hardcopy or electronic format and associated materials must be regarded as secure documents. As a result, such materials may not be reproduced, electronically transmitted, discussed, used in classroom instruction, or in any way released or distributed to unauthorized persons. All materials including items and item drafts must be returned at the end of the meeting.

I understand that I am responsible for test materials security. By breaching test materials security as described here, I am breaching professional testing ethics and may be subject to additional penalties under law.

Name: _____

Signature: _____

Date: _____

APPENDIX—F
PERFORMANCE LEVEL DESCRIPTORS



OKLAHOMA STATE DEPARTMENT OF
EDUCATION
— CHAMPION EXCELLENCE —



Oklahoma Grade 8 Science Performance Level Descriptor Tables

Policy PLDs

Policy PLDs define the knowledge and skill level expectations for the Oklahoma Academic Standards for Science.

Advanced

Students demonstrate superior performance on challenging subject matter.

Proficient

Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level.

Basic

Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level.

Below Basic

Students have not performed at least at the Basic level. Students scoring at the Below Basic level should be given comprehensive science instruction.

Borderline PLDs

Borderline PLDs describe the knowledge and skills that students within each proficiency level are just barely expected to be able to demonstrate. In line with the Oklahoma Academic Standards for Science, the statements combine the subject matter for science that students are expected to demonstrate.

Advanced

Students at the borderline of the **Advanced** level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the **Advanced** level can do the following more than 67% of the time:

- **evaluate, revise, or predict a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.**
- **evaluate or modify investigations about: stability and change of forces and motion; the effect of fields on force interactions.**
- **analyze, infer, relate, or identify complex relationships within a system to construct or evaluate explanations for: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.**
- **modify the solution to a problem with new information involving energy transfer, forces, and motions in systems where objects collide.**
- **evaluate, develop, or apply reasoning to support or refute new arguments or counterarguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.**
- **revise questions about data based on new evidence to determine factors that affect the strength of electric and magnetic forces.**
- **analyze mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.**
- **evaluate data to: compare patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.**
- **compare competing claims or scientific explanations to communicate how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.**

Proficient

Students at the borderline of the **Proficient** level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the **Proficient** level can do the following more than 67% of the time:

- develop or use a model to describe: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.
- identify, describe, or explain: a plan to investigate stability and change of forces and motion; how to conduct and evaluate investigations about the effect of fields on force interactions.
- identify, describe, or compare evidence to construct explanations for: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.
- design or revise a solution to a problem involving energy transfer, forces, and motions in systems where objects collide.
- use reasoning to show that evidence supports or refutes arguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.
- use reasoning to develop questions about data to determine factors that affect the strength of electric and magnetic forces.
- use mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.
- analyze and interpret data to: compare patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.
- gather, use, synthesize, or integrate information to communicate and support claims about how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.

Basic

Students at the borderline of the **Basic** level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the **Basic** level can do the following more than 50% of the time:

- identify or describe basic components or concept(s) of a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.
- identify or describe basic steps or processes within investigations about: stability and change of forces and motion; the effect of fields on force interactions.
- identify or describe basic relationships shown in evidence of: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.
- identify or describe basic relationships in a design solution involving energy transfer, forces, and motions in systems where objects collide.
- identify evidence that supports arguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.
- determine factors that affect the strength of electric and magnetic forces.
- identify components of mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.
- use data to: recognize patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.
- describe information to support claims about how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.

Below Basic

Students scoring **Below Basic** have not demonstrated they can perform at the Basic level. Students scoring at the Below Basic level should be given comprehensive science instruction. Students scoring at the Basic level typically:

- identify or describe basic components or concept(s) of a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.
- identify or describe basic steps or processes within investigations about: stability and change of forces and motion; the effect of fields on force interactions.
- identify or describe basic relationships shown in evidence of: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.
- identify or describe basic relationships in a design solution involving energy transfer, forces, and motions in systems where objects collide.
- identify evidence that supports arguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.
- determine factors that affect the strength of electric and magnetic forces.
- identify components of mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.
- use data to: recognize patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.
- describe information to support claims about how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.

LS3.1 LS3.2 ESS1.1 ESS1.2	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Develop and Use Models DCI <ul style="list-style-type: none"> LS3.A Inheritance of Traits LS1.B Growth and Development of Organisms LS3.B Variation of Traits ESS1.A The Universe and Its Stars CCC <ul style="list-style-type: none"> Structure and Function Cause and Effect Patterns Systems and System Models 		identify or describe basic components or concept(s) of a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.	develop or use a model to describe: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.	evaluate, revise, or predict a model involving: the relationship between gene structure and protein structure; the effect of reproduction on genetic variation; cyclic patterns in relation to the position of the Earth, Sun, and Moon; the role of gravity within galaxies and the solar system.

PS2.2 PS2.5	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Planning and Carrying Out Investigations DCI <ul style="list-style-type: none"> PS2.A Forces and Motion PS2.B Types of Interactions CCC <ul style="list-style-type: none"> Cause and Effect Stability and Change 		identify or describe basic steps or processes within investigations about: stability and change of forces and motion; the effect of fields on force interactions.	identify, describe, or explain: a plan to investigate stability and change of forces and motion; how to conduct and evaluate investigations about the effect of fields on force interactions.	evaluate or modify investigations about: stability and change of forces and motion; the effect of fields on force interactions.

LS1.5 LS4.2 LS4.4	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
<p>Constructing Explanations</p> <p>DCI</p> <ul style="list-style-type: none"> • LS1.B Growth and Development of Organisms • LS4.A Evidence of Common Ancestry and Diversity • LS4.B Natural Selection <p>CCC</p> <ul style="list-style-type: none"> • Cause and Effect • Patterns 		<p>identify or describe basic relationships shown in evidence of: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.</p>	<p>identify, describe, or compare evidence to construct explanations for: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.</p>	<p>analyze, infer, relate, or identify complex relationships within a system to construct or evaluate explanations for: the effect of environmental and genetic factors on growth; the common ancestry of organisms based on patterns in anatomy or the chronological order of fossils; the effect of trait variation in populations on natural selection.</p>

PS2.1	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Designing Solutions DCI <ul style="list-style-type: none"> PS2.A Forces and Motion CCC <ul style="list-style-type: none"> Systems and System Models 		identify or describe basic relationships in a design solution involving energy transfer, forces, and motions in systems where objects collide.	design or revise a solution to a problem involving energy transfer, forces, and motions in systems where objects collide.	modify the solution to a problem with new information involving energy transfer, forces, and motions in systems where objects collide.

LS1.4 PS2.4	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
<p>Engaging in Argument from Evidence</p> <p>DCI</p> <ul style="list-style-type: none"> LS1.B Growth and Development of Organisms PS2.B Types of Interactions <p>CCC</p> <ul style="list-style-type: none"> Cause of Effect Systems and System Models 		<p>identify evidence that supports arguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.</p>	<p>use reasoning to show that evidence supports or refutes arguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.</p>	<p>evaluate, develop, or apply reasoning to support or refute new arguments or counterarguments about how: the structures of plants and behaviors of animals affect the likelihood of successful reproduction; gravitational interactions depend on the masses of interacting objects in a system.</p>

PS2.3	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Asking Questions DCI <ul style="list-style-type: none"> PS2.B Types of Interactions CCC <ul style="list-style-type: none"> Cause and Effect 		determine factors that affect the strength of electric and magnetic forces.	use reasoning to develop questions about data to determine factors that affect the strength of electric and magnetic forces.	revise questions about data based on new evidence to determine factors that affect the strength of electric and magnetic forces.

PS4.1 LS4.6	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Using Mathematics and Computational Thinking DCI <ul style="list-style-type: none"> PS4.A Wave Properties LS4.C Adaptation CCC <ul style="list-style-type: none"> Patterns Cause and Effect 		identify components of mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.	use mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.	analyze mathematical representations to: describe patterns in wave models to show the relationship between amplitude and energy; explain how natural selection affects the distribution of traits in populations.

LS4.3 LS4.1 ESS1.3	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
Analyzing and Interpreting Data DCI <ul style="list-style-type: none"> LS4.A Evidence of Common Ancestry and Diversity ESS1.B Earth and the Solar System ETS1: Interdependence of Science, Engineering, and Technology CCC <ul style="list-style-type: none"> Patterns Scale, Proportion, and Quantity 		use data to: recognize patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.	analyze and interpret data to: compare patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.	evaluate data to: compare patterns of embryological similarities between species; identify how patterns in the fossil record indicate the history of life on Earth; determine the scale properties of objects in the solar system.

LS4.5 PS4.3	Below Basic: Students have not demonstrated they can perform at the Basic level.	Basic: Students at the borderline of the Basic level can demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level more than 50% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Below Basic level rather than the Basic level, students scoring at the Basic level can do the following more than 50% of the time:	Proficient: Students at the borderline of the Proficient level can demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of skills at the Basic level rather than the Proficient level, students scoring at the Proficient level can do the following more than 67% of the time:	Advanced: Students at the borderline of the Advanced level can demonstrate superior performance on challenging subject matter more than 67% of the time on the assessment. While these students sometimes may only demonstrate understanding and application of knowledge and skills at the Proficient level rather than the Advanced level, students scoring at the Advanced level can do the following more than 67% of the time:
<p>Obtaining, Evaluating, and Communication of Evidence</p> <p>DCI</p> <ul style="list-style-type: none"> • LS4.B Natural Selection • ETS2.A Interdependence of Science, Engineering, and Technology • PS4.C Information Technologies and Instrumentation <p>CCC</p> <ul style="list-style-type: none"> • Structure and Function • Cause and Effect 		<p>describe information to support claims about how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.</p>	<p>gather, use, synthesize, or integrate information to communicate and support claims about how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.</p>	<p>compare competing claims or scientific explanations to communicate how: humans affect trait inheritance through artificial selection; the structure and function of digital signals contributes to those signals reliably transmitting information.</p>

APPENDIX G
READINESS SURVEYS

Readiness Surveys

Round 1 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I understand the goals of the standard setting meeting.		
I understand the procedures we are using to set standards.		
I understand the differences between the performance levels.		
I understand what materials/content I should consider when making judgments.		
I understand the item-PLD alignment task and how to make it.		
I understand how to use the Cognia Standard Setting Toolkit.		
I am ready to proceed with the standard setting process.		

Round 2 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I understand the round 1 feedback.		
I understand that I should use the round 1 feedback as information, not persuasion, for me to consider as I make my judgements in round 2.		
I understand what the content-based benchmarks represent.		
I understand that I can use the content-based benchmarks as additional information, not persuasion, for me to consider as I make my judgements in round 2.		
I understand that I should make my own judgments about matching items to Performance Level Descriptors (PLDs), with consideration of the insights of my colleagues, but independently of the views and opinions of my colleagues.		
I am ready to proceed with Round 2 of the standard setting process.		

Round 3 Readiness Survey

Survey Questions	Response Options	
	Yes	No
I understand the round 2 feedback.		
I understand that I should use the round 2 feedback as information, not persuasion, for me to consider as I make my judgements in round 3.		
I understand that I should make my own judgments about matching items to Performance Level Descriptors (PLDs), with consideration of the insights of my colleagues, but independently of the views and opinions of my colleagues.		
I am ready to proceed with Round 3 of the standard setting process.		



APPENDIX H

ROUND RESULTS

OK OSTP Science Grade 8 Standard Setting Round Results

Figure 1. Round 1 - Frequency of Panelist Judgments across Basic, Proficient, and Advanced Levels

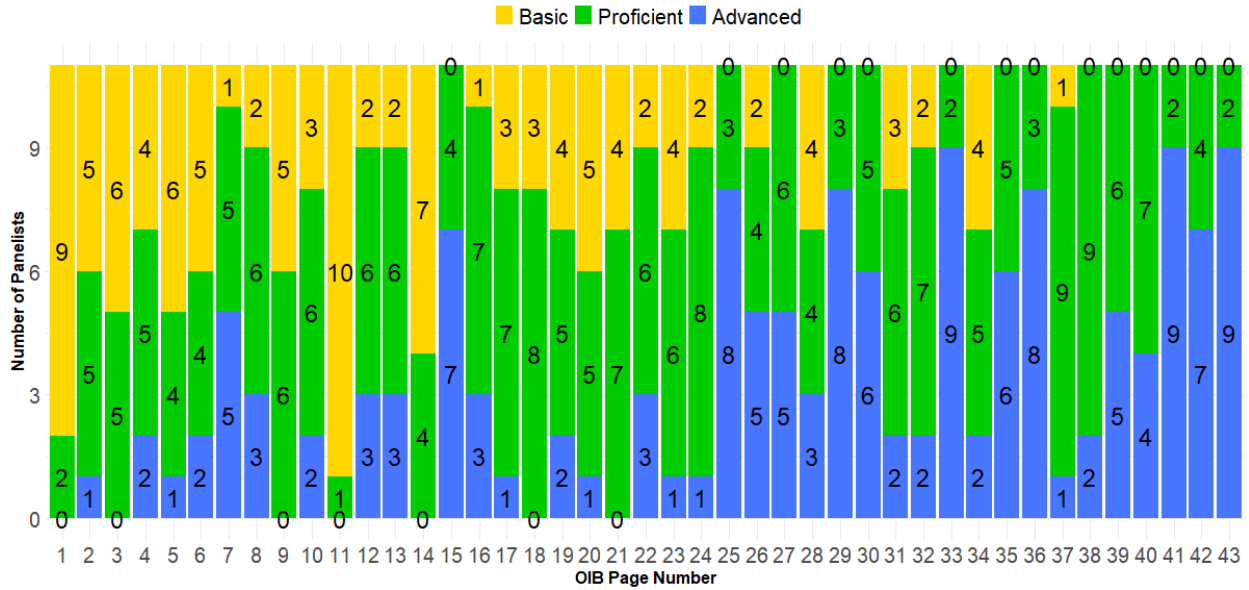


Figure 2. Round 2 - Frequency of Panelist Judgments across Basic, Proficient, and Advanced Levels

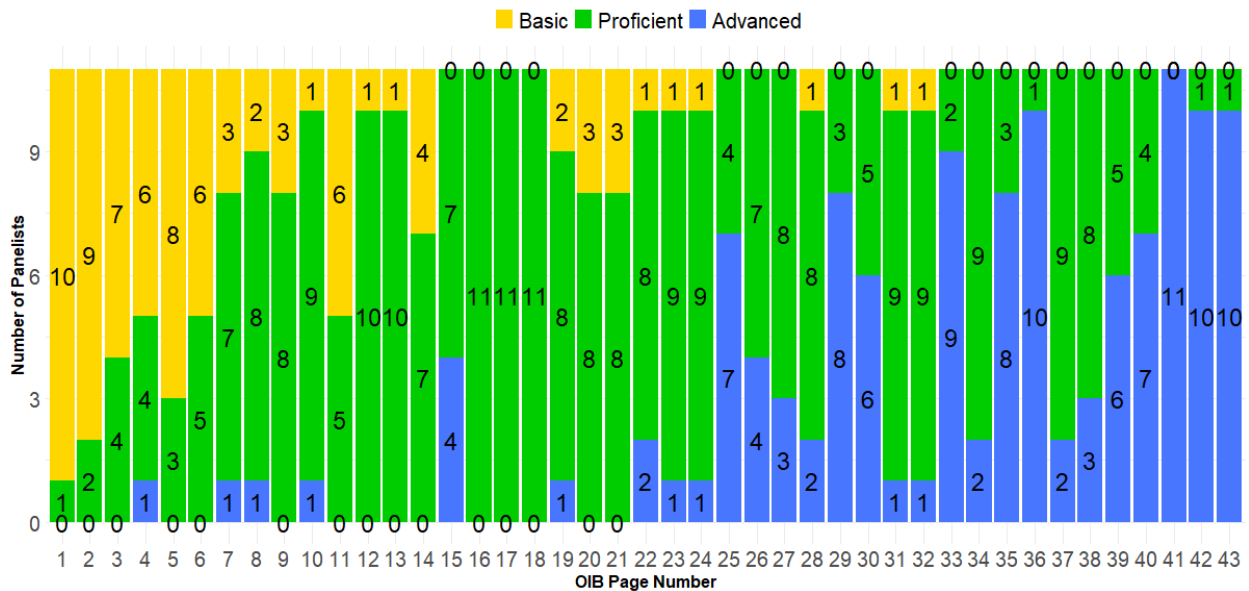
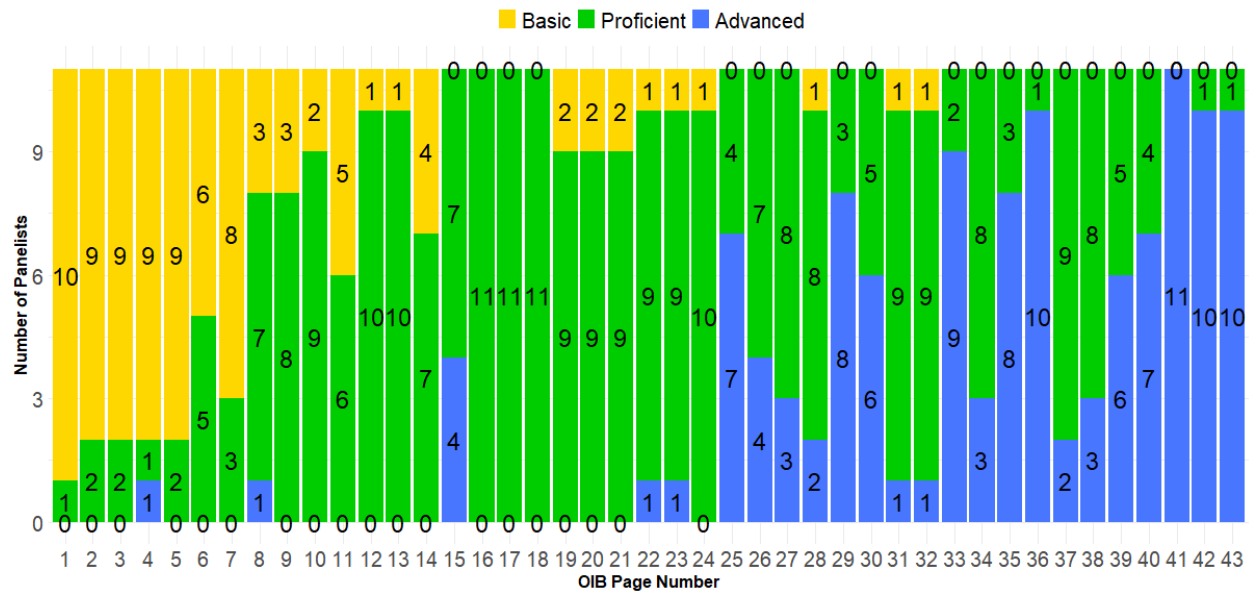


Figure 3. Round 3 - Frequency of Panelist Judgments across Basic, Proficient, and Advanced Levels



APPENDIX—I
WORKSHOP EVALUATION SURVEY

OK OSTP Science Grade 8 Standard Setting

Final Workshop Evaluation

Questions 1 – 20 were selected response items on the following Likert-type scale: Strongly Disagree, disagree, undecided, agree, strongly agree, or not applicable.

1. I understood the goals of the standard setting workshop.
2. I understood the procedures we followed to set standards.
3. I understood that my role was to make content-based judgements about the alignment between the items and the performance level descriptors.
4. The workshop procedures made sense to me, and I learned how to apply them efficiently.
5. I am confident about my understanding of this standard setting process.
6. The workshop facilitator explained things clearly to us.
7. The workshop facilitator encouraged us to raise questions and put our understandings into our own words.
8. The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.
9. The workshop facilitator took steps to help the standard setting process run smoothly.
10. Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.
11. I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.
12. I became sufficiently familiar with the assessment to make item-PLD judgements, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.
13. I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item response demands to PLDs.
14. I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.
15. I understood how to use the standard setting tool to record my item-PLD alignment judgements.
16. I understood how to write content-based rationales for my item-PLD alignment judgements.
17. I understood that the cut scores were calculated based on all item-PLD alignment judgements from all panelists.
18. I understood how to use the feedback after round 1, in preparation for round 2.
19. I understood what the content-based benchmarks, introduced in round 2, represented.
20. I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgements.

Question 21 – 23 were open response questions.

21. Please indicate any parts of the standard setting training and process that we should improve.
22. Please indicate any parts of the standard setting training and process that you felt worked really well.
23. Please note any other feedback you would like us to consider.



APPENDIX—J
WORKSHOP EVALUATION RESULTS

OK OSTP Science Grade 8 Standard Setting Workshop Evaluation Results

Table 1. Frequency of Evaluation Responses (N = 11)

Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
1. I understood the goals of the standard setting workshop.	--	--	--	2	9	--
2. I understood the procedures we followed to set standards.	--	--	--	2	9	--
3. I understood that my role was to make content-based judgements about the alignment between the items and the performance level descriptors.	--	--	--	3	8	--
4. The workshop procedures made sense to me, and I learned how to apply them efficiently.	--	--	1	4	6	--
5. I am confident about my understanding of this standard setting process.	--	--	--	4	7	--
6. The workshop facilitator explained things clearly to us.	--	1	1	2	7	--
7. The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	--	--	--	4	7	--
8. The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	--	--	1	4	6	--
9. The workshop facilitator took steps to help the standard setting process run smoothly.	--	--	2	3	6	--
10. Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	1	4	--	3	3	--
11. I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	--	--	1	3	7	--
12. I became sufficiently familiar with the assessment to make item-PLD judgements, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	--	--	--	5	6	--
13. I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item response demands to PLDs.	--	--	--	5	6	--
14. I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	--	--	--	2	9	--
15. I understood how to use the standard setting tool to record my item-PLD alignment judgements.	--	--	--	1	10	--
16. I understood how to write content-based rationales for my item-PLD alignment judgements.	--	--	--	4	7	--
17. I understood that the cut scores were calculated based on all item-PLD alignment judgements from all panelists.	--	--	--	3	8	--
18. I understood how to use the feedback after round 1, in preparation for round 2.	--	--	--	4	7	--
19. I understood what the content-based benchmarks, introduced in round 2, represented.	--	--	1	4	6	--
20. I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgements.	--	--	1	4	6	--
21. I am satisfied with the final results and cut scores	1	--	3	5	2	--



Table 2. Open-ended responses

Questions	Responses
Please indicate any parts of the standard setting training and process that we should improve.	"We waited too long on data to be returned and subsequently were delayed in ending. I personally spent 10 hours on the process today so this should have been split over 3 days instead of 2."
	"With the reading load that is needed on the science items, I felt like we were not given enough time. This needs to be a 3-day committee workshop, not a 2-day. That way people don't feel rush and we don't have to go over the scheduled time."
	"here should be less time spent on the repetitive explanation of the process on day one. I think there should be more time focused on the action verbs of the PLDs as it relates to KSAs before Round 1. The borderline PLD document was unnecessary and could be added to the range PLD document on one page. I think if possible, the down time waiting on the analysis of all panelist data could be sped up"
	"perhaps go into detail about the different levels and how to relate them to KSA more for those that didn't understand it."
	"Was this in the past a 3-day thing? Why do I feel like it was?"
	"The explanation of the borderline PLD's was a little confusing at first. I didn't really understand it until we got deep into the process"
	"Everything else was great."
Please indicate any parts of the standard setting training and process that you felt worked really well.	"I felt that the chunking or breaking up of the process into rounds worked well."
	"review rounds"
	"think it all went very well. The David, Mary-Alice and Frank worked very well together and made the process very easy."
If you would recommend changing any of the final cut scores, please indicate which cut scores (Basic, Proficient, and/or Advanced) you would recommend changing. For each recommended cut score change, please also note if you would recommend moving it earlier or later in the OIB and by how many pages.	"Basic needs to be expanded. More than 48 % of the students in Oklahoma need to pass this exam."
	"I felt that the chunking or breaking up of the process into rounds worked well."
	"I would change the cut scores specifically Basic and Proficient."
	"I feel like the line between Basic and Proficient was/is a little blurry (but I do realize that is why we had to do round 4). I'm OK with where the cut scores were placed because I feel like that was the consensus of the group, but I would also understand if it was moved by 1-2 questions later in the OIB."
	"I do not think I would move them"

APPENDIX K
STANDARD SETTING MEMO

Oklahoma Standard Setting Memo

OSTP Science Grade 8

June 22-23, 2023

Overview

Cognia and the Oklahoma Department of Education convened a panel of science teachers during June 23-24, 2023, to establish Basic, Proficient, and Advanced cut scores to enable reporting of student performance on the OSTP Science Grade 8 assessment. Eleven educators from around the state participated in two days of training and decision-making with Cognia standard setting specialists. The standard setting panelists reviewed test content and performance level descriptors and followed the Item-Descriptor (ID) Matching standard setting method.

The purpose of this memo is to present the cut scores and associated impact data that resulted from the standard setting meeting.

Methods

Standard Setting Procedure

During the standard setting meeting, the panelists were trained on and followed the ID-Matching method. Each panelist reviewed each item in the ordered item booklet (OIB) and considered the knowledge, skills, and abilities required by the item. Panelists then matched those item response demands to the knowledge and skill expectations in the performance level descriptors for the Basic, Proficient, and Advanced levels. Working independently, the standard setting panelists conducted the ID matching process in three rounds and made item-PLD alignment judgements for each item. Before each round, panelists completed a round readiness survey. After rounds 1 and 2, the Cognia workshop facilitator led panelists through a discussion of agreements and disagreements among the panelists and rationales for their various item-PLD alignment judgements. The ensuing discussion enabled panelists to consider their colleagues' insights about item response demands and rationales for matching items to descriptors, and to consider adjusting their judgements in rounds 2 and 3.

At the beginning of round 2, content-based benchmarks were introduced to panelists which served as additional information for panelists to consider as they made their item-PLD alignment judgements in rounds 2 and 3. At the completion of round 3, the resulting cut scores and associated impact data were presented to panelists. Impact data are the percentages of students who would be sorted into the Below Basic, Basic, Proficient, and Advanced performance levels, using their scores from the 2023 administration of the OSTP Science grade 8 assessment, and based on the cut scores calculated after round 3. Panelists then completed a round 4 evaluation step during which the facilitator guided the panel through the procedure of writing group level content-based rationales for item-PLD alignments where panelists disagreed with the content-based benchmark alignments. Finally, panelists completed the final evaluation survey about their overall experience and satisfaction with the standard setting workshop.

Analyses Procedure

First, Cognia conducted statistical analyses of panelists' item-PLD alignment data by calculating the percent exact, adjacent, and discrepant for each panelist on each performance level. Panelists with the least percentage exact were identified as showing statistically aberrant behavior.

Next, an independent subject matter expert (SME) reviewed the qualitative data for all panelists identified as statistically aberrant. The SME reviewed panelists' notes on the knowledge, skills, abilities required by the items, as well as their content-based rationales to determine if the panelists were on task.

After analyses and qualitative review, the data of panelists that were determined to be statistically and qualitatively aberrant were removed before proceeding with the rest of the analyses.

The next phase of the analyses included conducting logistical regression to calculate cut scores. Since the logistical regression method is sensitive to statistical outliers and the presence of such outliers violate the assumptions of the model, an outlier analysis was performed in the form of visual inspection of the initial logistic regression curves for any statistical outliers. Statistical outliers were identified, and the associated data points were removed and then the final logistic regression analyses were conducted to calculate the proficient and advanced cut scores. After calculating the proficient and advanced cut scores, the TCC method was used to calculate the Basic cut score.

Finally, the resulting cut scores were applied to student data from the spring 2023 administration of the OSTP science grade 8 assessment to calculate the impact data (i.e., the percentage of students that would be classified into each performance level based on the standard setting cut scores).

Results

After the statistical analyses and qualitative review of panelist data, one panelist was determined to be statistically and qualitatively aberrant. Consequently, their data were removed from the final analyses.

Visual inspection of the initial logistic regression curves for the proficient and advanced cuts scores revealed 7 statistical outlier data points. Figure 1 shows the initial logistic regression curve for the proficient level. The row of dots at the top and bottom of the curve represents 1 or more item-PLD judgements. Data points to the far left and right (i.e., in the tails) that are circled in red represent statistical outliers, and the numbers above or below the circles indicate how many item-PLD judgements are located at that specific location. The 7 data points were removed from the final analyses.

Figure 1. Initial Logistic Regression Curve for the Proficient Cut Score

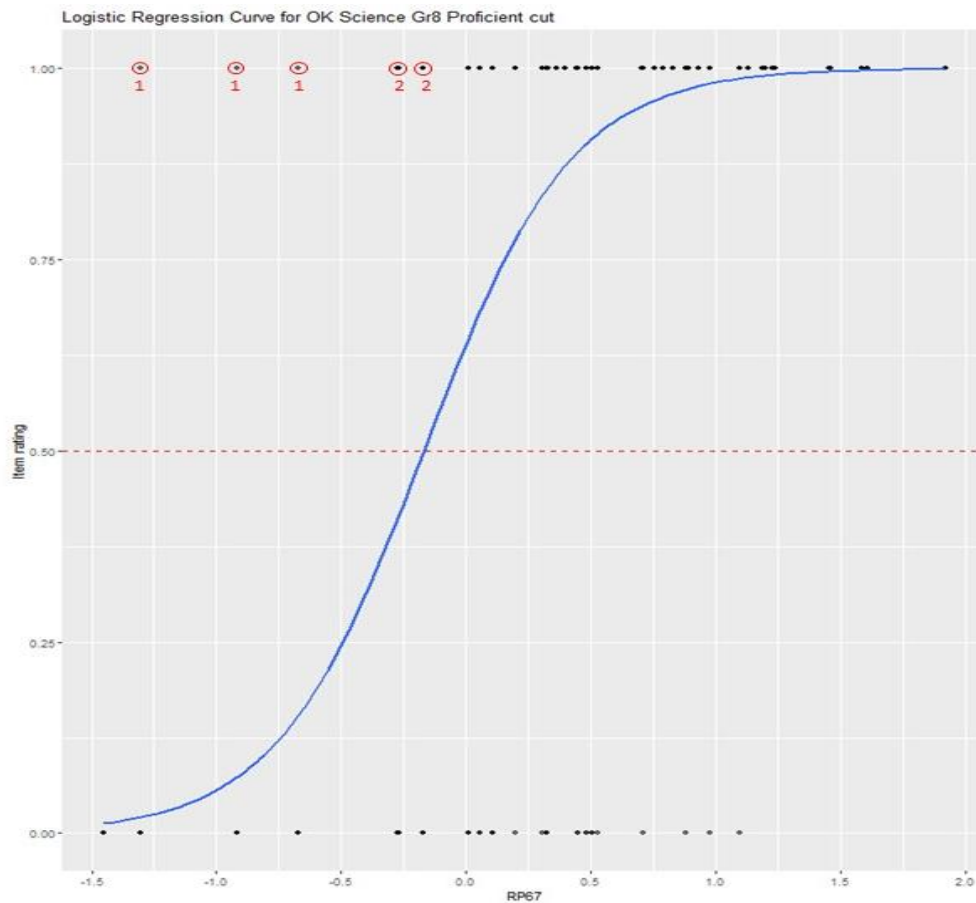


Table 1 shows the cut scores that resulted from the standard setting meeting and analyses, as well as the associated impact data for OSTP science grade 8. The percentage of Oklahoma students in each grade is shown for each performance level (Below Basic, Basic, Proficient, and Advanced), as well as for the combined proficient and advanced performance levels. Figure 2 gives a visual representation of the impact data.

In addition to the cut scores and impact data, Table 1 also lists standard errors associated with each cut score. Note that these standard errors are based on the round 1 judgement data because it is most reflective of the discrepancy between panelists. After round 1, panelists enter group discussions and thus their judgements begin to violate the assumption of dependency.

Table 1. OSTP Science Grade 8 Standard Setting Cut Scores and Impact Data

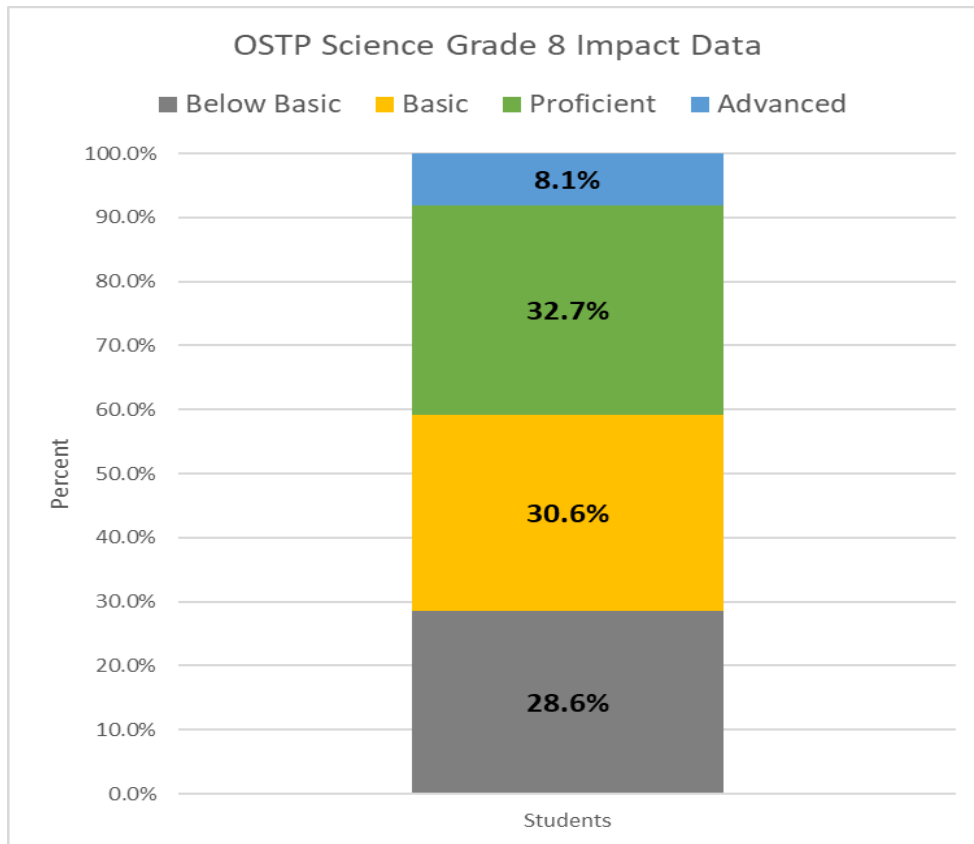
Performance Level	OIB #	Theta	*Standard Error	% Students
Below Basic	--	--	--	28.6
Basic	3 – 4	-0.89	0.115	30.6
Proficient	6 – 7	-0.07	0.155	32.7
Advanced	34 – 35	1.14	0.131	8.1
Proficient + Advanced	--	--	--	40.8

Note. OIB = Ordered Item Booklet

**Standard Error based on round 1 panelist data*



Figure 2. OSTP Science Grade 8 Impact Data based on Standard Setting Cut Scores



In the final workshop evaluation, panelists expressed generally positive support for the workshop overall, workshop facilitation, training, practice, and the workshop process. Table 2 shows the frequency of panelist selected responses to the following final evaluation statement: “I am satisfied with the final results and cut scores.” In addition, quotes from panelists associated with the follow up question to the above statement are also presented in Table 2.

As shown in the table, 1 panelist strongly disagreed with the above statement, 3 panelists were undecided, 5 panelists agreed, and 1 panelist strongly agreed. Three panelists specifically mentioned concerns about the Basic and Proficient cut scores and/or the impact data related to the Basic and Proficient range. In addition, one panelist specifically mentioned concerns about the impact data in the advanced performance level.

Table 2. Frequency of Responses (and related panelist quotes) for Survey Question: “I am satisfied with the final results and cut scores.”

Question Response Options	N (# Panelists)	Related Panelist Quotes
Strongly Disagree	1	“I would change the cut scores specifically Basic and Proficient”
Disagree	--	
Undecided	3	“...My initial thoughts are that it seems there should be more students in the basic and proficient range” “I am undecided” “I recommend taking the data provided by those who are actually in the classroom into much more consideration for validity of reasoning”
Agree	5	“Maybe make the Advance range slightly larger” I feel like the line between Basic and Proficient was/is a little blurry (but I do realize that is why we had to do round 4). I’m OK with where the cut scores were placed because I feel like that was the consensus of the group, but I would also understand if it was moved by 1-2 questions later in the OIB.
Strongly Agree	2	“I do not think I would move them”

Considerations for Articulation and Policy Review

SDE can accept the standard setting cut scores and adopt them as is. Or the department may choose to make “policy adjustments” to the standard setting cut scores, using several criteria.

A common psychometric approach. Adjust the cut scores based on the standard errors of the cut scores. Table 3 shows the cut scores, standard errors, and impact data based on the standard setting results. In addition, the table shows the theta cuts and impact data if the standard setting cuts were to be adjusted by 1 standard error. For additional reference and consideration, the last two columns of Table 3 show the theta cuts and impact data based on cut scores derived from Cognia’s content specialists’ item-PLD alignment work.

Policy and communication approach: Adjust the cut scores to achieve vertical and/or horizontal articulation. For reference, Table 4 shows the cut scores and impact data across grades 5, 8, and 11. See Figure 3 for a visual representation of the impact data across grades. Note that the impact data for grade 8 is based on the standard setting cut scores.

As SDE conducts their policy review, we also encourage SDE to consider the panelists’ thoughts and opinions on the standard setting cut score results as outlined in Table 2 of the results section.

We at Cognia are grateful for the opportunity to contribute to this important conversation and will be pleased to do discuss the contents of this memo and advise SDE on psychometrically defensible ways to make policy adjustments.

Table 3. OSTP Science Grade 8 Cut Scores and Impact Data based on Psychometric Adjustments.

Performance Level	Standard Setting Cut Scores			Cut Score adjusted DOWN by 1 Standard Error		Cut Score adjusted UP by 1 Standard Error		Cut Scores based on Cognia Content Specialists Data	
	Theta	*Standard Error	% Students	Theta	% Students	Theta	% Students	Theta	% Students
Below Basic	--	--	28.6	--	24.9	--	32.7	--	40.7
Basic	-0.89	0.115	30.6	-1.01	28.6	-0.78	31.8	-0.56	28.1
Proficient	-0.07	0.155	32.7	-0.22	36.1	0.09	29.3	0.22	27.7
Advanced	1.14	0.131	8.1	1.01	10.5	1.27	6.2	1.51	3.6
Proficient + Advanced	--	--	40.8	--	46.5	--	35.5	--	31.2

Note. OIB = Ordered Item Booklet

*Standard Error based on round 1 panelist data

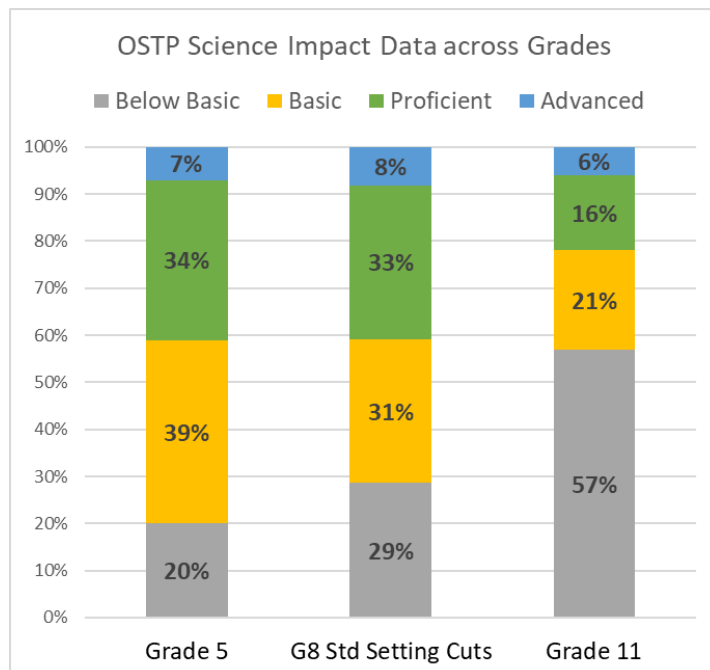
Table 4. OSTP Science Cut Scores and Impact Data across Grades

Performance Level	Grade 8 (Standard Setting Cut Scores)			Grade 5		Grade 11	
	Theta	*Standard Error	% Students	Theta	% Students	Theta	% Students
Below Basic	--	--	28.6	--	20.0	--	57.0
Basic	-0.89	0.115	30.6	-0.91	39.0	0.17	21.0
Proficient	-0.07	0.155	32.7	0.18	34.0	0.80	16.0
Advanced	1.14	0.131	8.1	1.32	7.0	1.53	6.0
Proficient + Advanced	--	--	40.8	--	40.0	--	22.0

Note. OIB = Ordered Item Booklet

*Standard Error based on round 1 panelist data

Figure 3. OSTP Science Impact Data across Grades



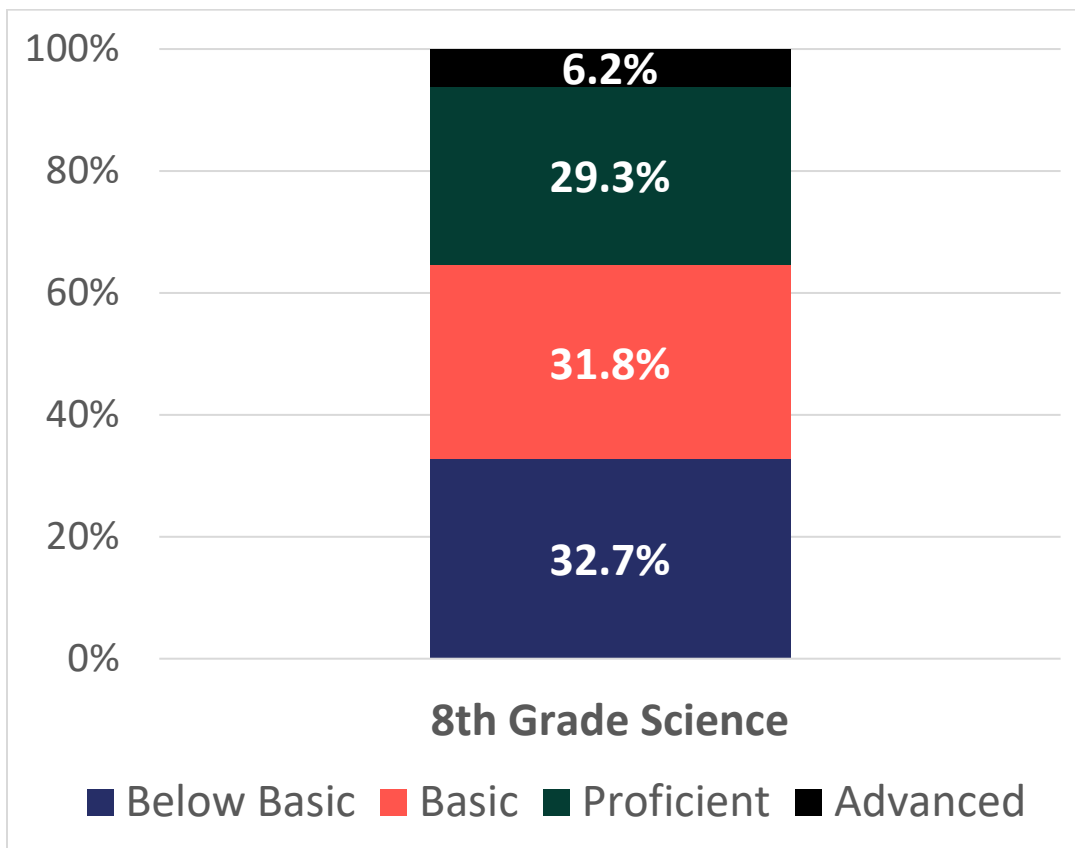
APPENDIX—L
FINAL CUT POINTS

OK OSTP Science Grade 8 Standard Setting Final Cut Points

Table 1. OK OSTP Science Grade 8 Standard Setting Final Cut Points

Performance Level	OIB Page #	Theta	Percent Students
Below Basic	--	--	32.7
Basic	3 – 4	-0.78	31.8
Proficient	10 – 11	0.09	29.3
Advanced	31 – 32	1.27	6.2
Proficient + Advanced	--	--	35.5

Figure 1. OK OSTP Science Grade 8 Impact Data based on Final Cut Points



APPENDIX Q
2024 OSTP ELA AND
MATHEMATICS STANDARD
SETTING REPORT



2024 Oklahoma Standard Setting Report

OSTP ELA and Mathematics—Grades 3–8

June 17–21, 2024—Broken Arrow, Oklahoma

Prepared by Cognia for the Oklahoma Department of Education

TABLE OF CONTENTS

CHAPTER 1. OVERVIEW OF STANDARD SETTING PROCEDURES..... 4

CHAPTER 2. TASKS COMPLETED PRIOR TO STANDARD SETTING..... 5

2.1 CREATION OF PERFORMANCE LEVEL DESCRIPTORS 5

2.2 PREPARATION OF MATERIALS 6

2.2.1 Ordered Item Booklets (OIBs)..... 6

2.2.2 Content-Based Benchmarks..... 7

2.2.3 Cognia Standard Setting Toolkit..... 7

2.2.4 Panelist Materials 8

2.2.5 Presentation Materials..... 9

2.2.6 Data, Information and Analysis Materials..... 9

2.3 SELECTION OF PANELISTS.....10

CHAPTER 3. DURING THE STANDARD SETTING MEETING11

3.1 OVERVIEW OF THE ID MATCHING METHOD11

3.2 MEETING LOGISTICS11

3.2.1 Standard Setting Panelists and Workshop Staff.....11

3.2.2 Standard Setting Meeting Schedule.....14

3.2.3 Standard Setting Meeting Security15

3.2.4 ID Matching Standard Setting Procedure.....15

3.3 DATA REVIEW, CUT SCORE CALCULATION, AND ANALYSES15

3.3.1 Initial Data Review.....15

3.3.2 Cut Score Calculation.....16

3.3.3 Analysis Procedure.....16

3.4 OPENING SESSION AND GENERAL ORIENTATION.....17

3.5 STANDARD SETTING BREAKOUT SESSIONS17

3.5.1 Overview and Introductions.....18

3.5.2 Experience the Test18

3.5.3 Use of the Cognia Standard Setting Toolkit.....18

3.5.4 Review of the Standards and Performance Level Descriptors18

3.5.5 Training on the ID-Matching Judgmental Task.....19

3.5.6 Modeling and Practice.....19

3.5.7 Judgment Rounds and Feedback.....20

3.5.8 Standard Setting Results and Impact Data23

3.5.9 Standard Setting Workshop Evaluation.....24

3.6 ARTICULATION MEETINGS25

3.6.1 Introduction, Overview, and Key Concepts25

3.6.2 Modeling of Standard Setting Panel Decisions25

3.6.3 Familiarization with Standards, Blueprints, and PLDs26

3.6.4 Expectations for Between-Grade Transitions26

3.6.5 Presentation of Impact Data and Discussion26

3.6.6 Closing and Articulation Evaluation Survey27

CHAPTER 4. TASKS COMPLETED AFTER THE STANDARD SETTING MEETING28

4.1 REVIEW AND ARTICULATION ADJUSTMENTS28

4.2 POLICY REVIEW AND APPROVAL OF FINAL CUT SCORES.....28

4.3 PREPARATION OF STANDARD SETTING REPORT29

REFERENCES.....30

APPENDICES.....31

APPENDIX A—PERFORMANCE LEVEL DESCRIPTORS
APPENDIX B—ORDERED ITEM BOOKLET BLUEPRINTS
APPENDIX C—LOGISTIC REGRESSION CALCULATION
APPENDIX D—COGNIA STANDARD SETTING TOOLKIT
APPENDIX E—FACILITATION POWERPOINT PRESENTATION
APPENDIX F—PANELIST INFORMATION
APPENDIX G—MEETING AGENDA
APPENDIX H—NON-DISCLOSURE AGREEMENT
APPENDIX I—ORIENTATION POWERPOINT PRESENTATION
APPENDIX J—READINESS SURVEYS & ROUND BY ROUND RESULTS
APPENDIX K—STANDARD SETTING EVALUATION SURVEY & RESULTS
APPENDIX L—ARTICULATION POWERPOINT PRESENTATION
APPENDIX M—ARTICULATION EVALUATION SURVEY & RESULTS
APPENDIX N—STANDARD SETTING MEMO
APPENDIX O— FINAL CUT POINTS
APPENDIX P— COMMISSION FOR EDUCATIONAL QUALITY AND ACCOUNTABILITY (CEQA) PRESENTATION

Chapter 1. Overview of Standard Setting Procedures

The purpose of this report is to summarize the activities involved in the Standard Setting process for the Oklahoma School Testing Program (OSTP) in English Language Arts (ELA) and mathematics grades 3–8 on behalf of the Oklahoma State Department of Education (OSDE). Changes in the Oklahoma Academic Standards for ELA and mathematics were implemented in Fall 2021 and 2022 respectively, prompting the need to reset standards. The primary goal of the standard setting was to determine the knowledge, skills, and abilities (KSAs) that students must demonstrate to be classified into one of the performance levels (i.e., Advanced, Proficient, Basic, or Below Basic).

The standard setting process used was a modified version of the Item-Descriptor (ID) Matching method (Ferrara & Lewis, 2012; Cizek & Bunch, 2007). The ID Matching method was selected because it reduces cognitive burden on panelists as compared to other standard setting methods that require probability judgments about hypothetical high- and low-performing students, and it most clearly translates content standards into performance categories as compared to other methods of standard setting (Cizek, Bunch, & Koons, 2004).

The standard setting meeting was held from June 17th through June 21st of 2024. In all, 66 panelists participated in the process and were organized into six grade-band panels. Each panel completed the standard setting activities for two grades. Within the breakout sessions, panelists were organized into three tables of 3–4 panelists each plus a facilitator provided by Cognia. At the end of the week, two articulation panels were convened (one each for ELA and mathematics) that consisted of 10–12 panelists from the original standard setting panels.

This report is organized into three major sections, describing tasks completed prior to, during, and after the standard setting meeting.

Chapter 2. Tasks Completed Prior to Standard Setting

2.1 Creation of Performance Level Descriptors

Oklahoma State Statute: Title 70. Schools, Chapter 22 – Testing and Assessment, Section 1210.541 – Student Performance Levels and Cut Scores – Accountability System mandates the adoption of “a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.” The law states that performance levels must be labeled and defined as follows:

1. Advanced, which shall indicate that students demonstrate superior performance on challenging subject matter;
2. Proficient, which shall indicate that students demonstrate mastery over appropriate grade-level subject matter and that students are ready for the next grade, course, or level of education, as applicable;
3. Basic, which shall indicate that students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level or course; and
4. Below Basic, which shall indicate that students have not performed at least at the Basic level.

Cognia collaborated with the Oklahoma State Department of Education (OSDE) to develop Range performance level descriptors (PLDs) for OSTP ELA and mathematics grades 3–8. Prior to this collaboration, Policy PLDs were established by the OSDE to define the knowledge and skill level expectations for the Oklahoma Academic Standards for ELA (OAS-ELA) and mathematics (OAS-M).

In developing the draft Range PLDs, Cognia worked collaboratively with OSDE and took into consideration the content standards and the achievement construct the PLDs represent, and used statements developed for the OSTP ELA and mathematics grades 3–8 assessments to organize Range PLDs for each assessable OSTP standard and objective. Cognia reviewed the content standards to select (a) verbs that define ELA and mathematics skills and thinking processes, (b) nouns to identify knowledge and understanding of ELA and mathematics facts and concepts, and (c) when necessary, modifiers (i.e., adverbs, adjectives) that indicate levels of frequency, consistency, or quality of student performance. Following the framework described in Egan et al. (2012), Cognia collaborated with OSDE and Oklahoma educators to review the draft Range PLDs (i.e., knowledge and skill expectations for all students who have achieved the range of scores in a performance level). Lastly, Cognia and OSDE worked together to approve final Range PLDs ahead of the standard setting meeting. The final Range PLDs were approved by OSDE in April of 2024.

See Appendix A for the final approved Range PLDs for OSTP ELA and mathematics grades 3–8.

2.2 Preparation of Materials

Preparing for the standard setting meeting involved analyzing operational test data and organizing key materials. The materials that were prepared prior to the standard setting meeting included the following:

- Ordered Item Booklets (OIBs)
- Content-based benchmarks
- The Cognia Standard Setting Toolkit
- Panelist materials
- Presentation materials
- Data, information, and analysis materials

Details related to the materials preparation for each of the above categories are provided below.

2.2.1 Ordered Item Booklets (OIBs)

The standard setting was conducted using test items from prior administrations of the OSTP ELA and mathematics grades 3–8 assessments. The initial OIBs comprised operational test items, which were ordered in terms of difficulty. Item difficulty, as defined by its scale location given a response probability (RP) value, was calculated based on data from OSTP ELA and mathematics grades 3–8 students during the prior test administrations. Items ascended in order of difficulty through the OIB. Easier items appeared earlier in the OIB, and more difficult items appeared later.

Response probability (RP) criterion. The RP 67 criterion, defined by the Item Response Theory (IRT) scale value associated with a 67% chance of answering the item correctly, was used to order items in the OIB for the OSTP ELA and mathematics standard setting meeting.

Collection of items for the OIB. To ensure that the items included in the OIB spanned the difficulty continuum—from easy to difficult—and that items were found around the points on the test scale where cut scores were likely to appear, the following procedure was used for building the final OIBs that were used during the standard setting meeting:

- Start with an operational test form: Cognia ordered the items from the Spring 2024 operational test form. Operational items that fell below the statistical thresholds for psychometric adequacy were replaced with items from the same domain that did meet the thresholds.
- Augment the OIB with additional items: As needed, Cognia chose additional items for the OIB from previously field-tested items. For example, if the OIB did not have many items near the point in the test scale where the Proficient benchmark was expected, then items were added to the OIB that had locations around this point based on availability of such items in the pool.
- Review the balance of content against the blueprint: Since additional items were substituted in or added to the OIB, Cognia confirmed that the items had a balance of content consistent with the

test blueprint to ensure that individual content strands were less likely to be over or underrepresented in the OIB through the augmentation process.

Appendix B includes tables showing the blueprints for each subject- and grade-specific OIB.

2.2.2 Content-Based Benchmarks

In standard setting, benchmarks refer to any content- or policy-based information that comes from an external source and is presented to panelists. The exact way that the benchmarks are used in the standard setting depends upon the methodology implemented. However, the general use is the same: standard setting panelists see and consider information from these external measures as they engage in the standard setting meeting activities.

Content-based benchmarks were used for the OSTP ELA and mathematics grades 3–8 standard setting. The procedure for determining the content-based benchmarks was as follows:

- Prior to the standard setting meeting, Cognia and SDE content teams reviewed each item in the OIB and matched the items to one of three PLD levels (Basic, Proficient, or Advanced). Note that the content specialists did not assign any items to the Below Basic PLD. This is because the Below Basic performance level is described simply as the inability to perform at the Basic level.
- Cognia psychometricians then compiled the content specialists' item-PLD alignments and calculated threshold regions through logistic regression. Specifically, the regions were calculated by combining the item-PLD judgments to derive a set of cut scores with a margin of error added around each cut score. See Appendix C for calculation details.
- The above process resulted in content-based benchmark regions for the Proficient and Advanced levels.

Special Considerations for the Basic Benchmark Region. As mentioned previously, the Below Basic performance level is described as the inability to perform at the Basic level; therefore, items were not written to the Below Basic level and, by extension, it was not feasible to align items to the Below Basic level. Since there were no Below Basic item-PLD alignments, the above logistic regression method could not be employed to calculate a cut and corresponding region for the Basic level.

Thus, to facilitate the Basic level cut score identification, Cognia psychometricians empirically derived the cut score by constructing a mini-Test Characteristic Curve (TCC) based on items that were aligned to the Basic PLD. Cognia calculated a theta value that was associated with 50% beyond chance of the expected score of the mini-TCC. The '50% beyond chance' criterion is reflected in the performance level descriptor and takes guessing into account. Three OIB pages were added below and above the empirical cut score to create an empirical threshold region for the Basic level.

2.2.3 Cognia Standard Setting Toolkit

This section provides details about the Cognia Standard Setting Toolkit that panelists used to complete the main standard setting activities during the meeting. The Cognia Standard Setting Toolkit was

developed, tested, and set up by Cognia prior to the meeting and included digital ordered item booklets with integrated item lists, judgment forms, readiness surveys, and the final workshop evaluation survey.

The Cognia Standard Setting Toolkit consisted of a digital interface that first presented the ordered item list view (i.e., a list of items separated by rows with the easiest item at the top and the most difficult at the bottom). From the initial screen, panelists could toggle to the corresponding item detail view and use navigation arrows to move ‘up’ or ‘down’ in the booklet. The item detail view showed a PDF of the full item with the response options, as well as any stimuli or rubrics associated with the item. The ordered item booklets were created as discussed in a previous section of this document. Integrated judgment forms were available within both the item list and detail views. The judgment forms provided space for users to note (1) the relevant knowledge, skills, and abilities (KSAs) needed to answer the item, (2) any additional information that came to mind as panelists undertook the judgment task for each item, and (3) item descriptor matches. Any notes entered by the user in the item list view screen persisted when the user switched to the item detail view screen and vice versa. In addition to the above, the Cognia Standard Setting Toolkit included the round-specific readiness surveys that panelists completed before undertaking each judgment round. Finally, the toolkit included the final workshop evaluation survey that panelists completed at the conclusion of the standard setting meeting.

Additional details and screenshots of the Cognia Standard Setting Toolkit are available in Appendix D.

2.2.4 Panelist Materials

Cognia developed specific and relevant materials that were used by panelists during the meeting. Because panelists utilized the Cognia Standard Setting Toolkit for most of the standard setting activities, some of the materials were presented digitally within the Toolkit. Table 2-1 includes a list of the materials developed for the panelists and their mode of presentation.

Table 2-1. Panelist Materials Prepared Prior to the Standard Setting Meeting

Panelist Material	Paper	Digital Online	Digital Within the Toolkit
Meeting Agenda	✓	✓	
Non-disclosure Agreement	✓		
OSTP ELA or Mathematics Test		✓	
Performance Level Descriptors (PLDs)	✓		✓
ELA and Mathematics Standards	✓		✓
Formula Sheets (Mathematics Grades 6-8)	✓		
Definition Sheets (ELA)	✓		
Practice Items and Judgment Forms			✓
Round Readiness Surveys			✓
Ordered Item Booklets (OIBs)			✓
Integrated Item Map and Judgment Forms			✓
Workshop Evaluation Survey			✓

2.2.5 Presentation Materials

Several PowerPoint presentations were used throughout the duration of the meeting. An orientation PowerPoint presentation was delivered during the opening session of the standard setting meeting, while panel-specific facilitation PowerPoint presentations guided the facilitators through the distribution of information and materials during the main portion of the standard setting meeting. Finally, content-specific PowerPoint presentations were used during the ELA and mathematics articulation meetings that occurred at the conclusion of the standard setting portion of the meeting. Cognia developed the initial presentations and OSDE reviewed and approved the presentations prior to the standard setting meeting.

Notes and scripts that coincided with the PowerPoint slides were added within the presentations to guide facilitators. The notes and scripts for the meeting provided information, including procedural steps, talking points, definitions to explain concepts to panelists, answers to commonly asked questions, and specific materials to distribute to panelists. Copies of the facilitation, orientation, and articulation PowerPoint presentations are available in Appendices E, I, and L, respectively.

2.2.6 Data, Information and Analysis Materials

Prior to the standard setting meeting, all necessary data, information, and other relevant analysis materials were generated for use during the meeting. Table 2-2 shows a list of materials that were generated, as well as the purpose of each.

Table 2-2. Data, Information, and Analysis Materials Generated Before the Standard Setting Meeting

Data, Information, and Analysis Materials	Description/Purpose
Ordered Item Booklets (OIBs)	Each OIB comprised a set of items ordered by item difficulty and was generated according to the procedures outlined in section 2.2.1 of this report. Panelists worked within the OIBs to review items and follow the ID Matching process.
Content-based benchmark regions	Benchmark regions were calculated according to the procedures outlined in section 2.2.2 of this document. Panelists viewed and considered information from these benchmark regions as they engaged in the standard setting meeting activities.
Cognia Standard Setting Toolkit	A digital platform that was set up and tested prior to the meeting and included all necessary item data and information, as well as information related to the standards and PLDs.
Student Test Data	Student test data from the Spring 2024 administration of the OSTP ELA and mathematics grades 3-8 test were prepared to enable the calculation of impact data during and after the meeting.
Programming	Cognia created and tested programming for computing the following: <ul style="list-style-type: none">- Theta cut scores: Cut scores on the theta scale based on panelists' judgments after each judgment round.- Various statistics: Standard errors, percent exact and adjacent (based on differences between judgments from panelists and content specialists).- Panelist judgment frequency distributions: Computed for all panelists after each round. The code also produced presentation artifacts for use during the discussion session after each round.- Impact data: Code that used the theta cut scores and student test data to calculate the percentage of students in each performance level category.

2.3 Selection of Panelists

As emphasized in Cizek and Bunch (2007), regardless of the method used, the selection of panelists is a principal factor in determining standard setting outcomes and maximizing the validity of the standard-setting process. The guidance provided by *Standards for Educational and Psychological Testing* (AERA et al., 2014) states that “a sufficiently large and representative group of judges should be involved to provide reasonable assurance that results would not vary greatly if the process were repeated.”

Consistent with the above guidance and respecting practical considerations regarding the maximum size of a group that can be successfully managed, the goal was to recruit standard setting panels of 10–12 members per grade-band panel representing different stakeholder groups to set standards for OSTP ELA and mathematics. Targets for the size and composition of the panel were also consistent with federal guidelines as described in *Standards and Assessment Peer Review Guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001* (U.S. Department of Education, 2009).

Two goals were proposed for recruiting standard setting panelists: (a) diverse experience and points of view regarding students, student learning, and Oklahoma content standards and (b) diverse representation among panelists in years of teaching, geographic regions in the state, school system sizes, school system urbanicity, and the racial/ethnic make-up of the student populations.

Chapter 3. During the Standard Setting Meeting

3.1 Overview of the ID Matching Method

The Item-Descriptor (ID) Matching method is appropriate for setting standards for standards-aligned assessments like the OSTP ELA and mathematics grades 3–8 assessments. Assessment programs around the world have used ID Matching (e.g., Delaware, Massachusetts, Maryland, Mississippi, New Mexico, New York, South Carolina, and West Virginia; the Chicago and Philadelphia Public Schools; and programs in Brazil, Germany, and Finland).

ID Matching has advantages over Bookmark, Angoff, and other standard setting methods. Specifically, its cognitive-judgmental task requires that standard setting panelists, who are typically classroom educators, undertake a judgmental task that they are well suited for—matching item knowledge and skill response demands with knowledge and skill expectations in performance level descriptors (PLDs). The Bookmark and other methods require panelists to make probability judgments—something that people in general do not do well (e.g., Murphy, 2002). In addition, panelists do not need to hold a hypothetical borderline student in mind when they match items to descriptors and recommend cut scores, so the cognitive load and complexity of ID Matching is more manageable.

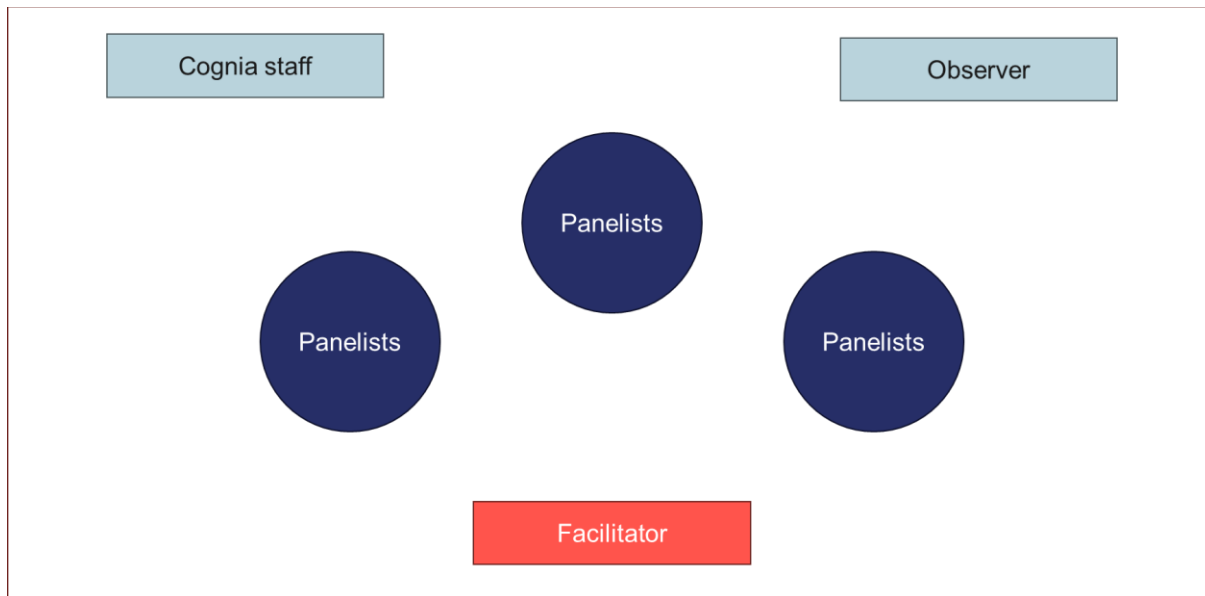
During standard setting using ID Matching, panelists review test items to identify the response demands of each item and then use the PLDs as their guide to match the item response demands to one of the performance level descriptors. The structure of the PLDs provides a general characterization of expected student knowledge and skill at each level and examples of the knowledge and skills that students at each performance level can be expected to demonstrate. By matching test items to specific claims from the Proficient PLD, for example, panelists identify the evidence in test items that supports the claims in that descriptor. Supporting the claims represented in the Proficient PLD contributes to the validity of interpretations of student performance, based on the PLDs, and to the overall validity argument that a student who achieves that level on the assessment has demonstrated adequate understanding of essential concepts with respect to the standards being measured. This logic applies to all cut scores and performance levels.

3.2 Meeting Logistics

3.2.1 Standard Setting Panelists and Workshop Staff

Participants of the OSTP ELA and mathematics standard setting meeting included meeting facilitators, content specialists, panelists, observers, and psychometricians. For the main standard setting activities, each of the six panels convened in a separate breakout room. Figure 3-1 illustrates the general setup for the breakout rooms.

Figure 3-1. Standard Setting Breakout Room Setup



Facilitators

Each standard setting panel was led by a facilitator. The facilitators were members of Cognia's staff who have experience facilitating standard setting meetings and were responsible for leading the panelists through the standard setting process.

The facilitators, with support from Cognia psychometricians and content specialists, ensured that appropriate standard setting processes were followed throughout all phases of the meeting and verified that panelists had a solid understanding of the tasks they were being asked to complete. The facilitators underwent preparatory training to lead the standard setting meeting. Psychometric staff from Cognia conducted the training, which included:

- OSTP ELA or mathematics overview: The facilitators were provided with an overview of the OSTP ELA or mathematics tests, including the different item types, scoring rules, and performance levels.
- Cognia Standard Setting Toolkit: The Cognia Standard Setting Toolkit was used by panelists throughout the standard setting meeting. The facilitators became familiar with the Toolkit to lead the standard setting process.
- Standard setting process: Facilitators participated in a walkthrough of the standard setting meeting, with a focus on specific issues for these meetings, such as time management, the use of the Cognia Standard Setting Toolkit, and communicating feedback information.
- Training slides and presentation script/notes: As part of the walkthrough of the standard setting process, facilitators reviewed the standard setting training slides. Notes in the standard setting training slides and a presentation script provided the facilitators guidance, including when specific language was to be used.

Content Specialists

Two Cognia content specialists, one each for ELA and mathematics, supported the standard setting meeting throughout the week. They presented information during the orientation session related to the development of the tests, procedures for scoring the items, and development/organization of the PLDs. In addition, the content specialists supported the facilitators throughout the standard setting process. Finally, the content specialists were co-facilitators during the articulation meetings.

Panelists

The OSDE selected panelists prior to the standard setting meeting. The goal for panel selection was to include participants who were primarily teachers, but also to include school administrators, higher education personnel, and stakeholders from other interest groups. Moreover, to the extent possible, panelists were selected to reflect a balance of gender, race/ethnicity, and geographic location. Finally, panelists were selected who were familiar with the relevant ELA or mathematics grades 3–8 subject matter. Table 3-1 provides summary information about the panelists that participated in the OSTP ELA and mathematics standard setting.

Appendix F contains detailed panelist information for each panel, including districts represented along with the gender and ethnicity breakdowns.

Table 3-1. Number of Panelists Overall and across Years of Teaching Experience

Panel	Overall	Years of Teaching Experience			
		1 – 5	6 – 10	11 – 20	21+
ELA Grades 3-4	11	8	2	1	--
ELA Grades 5-6	10	5	1	3	1
ELA Grades 7-8	10	5	3	1	1
Mathematics Grades 3-4	11	4	2	3	2
Mathematics Grades 5-6	12	4	--	4	4
Mathematics Grades 7-8	12	2	4	3	3
ELA Articulation	10	7	1	2	--
Mathematics Articulation	12	2	2	4	4

Observers

The purpose of the observers was to allow select individuals the opportunity to observe the standard setting process and, in some cases, provide feedback. Two types of observers, general and independent, were present during the meeting. The general observers consisted primarily of OSDE staff members that were assigned to specific breakout rooms and observed in those rooms for the duration of the meeting. In addition, three independent observers (two for ELA and one for mathematics) were also present during the meeting in an official observer capacity. The goal of the independent observers was to observe and take notes during the standard setting meeting and then write a report based on their observations. Cognia supplied the independent observers with Cognia Chromebooks, as well as specific observer-status log in credentials for the Cognia Standard Setting Toolkit. Within the Toolkit, the observers had the same access and permissions as a panelist; however, any actions they took or data they entered were excluded from the analyses and proceedings. Thus, the observers were able to follow along with the standard setting process. During the meeting, the independent observers floated between breakout sessions, timing their entries and exits to coincide with natural breaks to minimize any disturbances. The

independent observers also had access to the approved standard setting plan, PowerPoint presentations, facilitation scripts, PLDs and any other documents that were used during the meeting.

Psychometricians

Three Cognia psychometricians were on site to manage the Cognia Standard Setting Toolkit, complete real-time analyses during the meeting, and support the facilitators as needed throughout the standard setting process. A fourth Cognia psychometrician worked off-site to replicate the analyses conducted by the on-site psychometrician, thereby ensuring accuracy of the results. The lead psychometrician presented measurement-related information, as well as a broad overview of standard setting concepts, to the panelists during the orientation session. During the breakout sessions, the psychometricians floated between breakout rooms and answered any measurement questions or provided support to the facilitators as needed. In addition, they performed all calculations throughout the standard setting and presented during debrief meetings with OSDE whenever results were available. Finally, psychometricians presented impact data to panelists at the conclusion of the standard setting portion of the meeting and co-facilitated the articulation meetings.

3.2.2 Standard Setting Meeting Schedule

The standard setting portion of the meeting consisted of four days of activities. The meeting started with an opening session on the morning of day one before continuing with training, practice, and round one for the upper grade associated with each grade-band panel. Panelists engaged in the standard setting activities until they completed three rounds for each grade in their respective grade-band panels. After completing the activities for both grades, panelists completed a final standard setting workshop evaluation survey. The standard setting portion of the meeting concluded midday on day four for the mathematics panels, while the ELA panels concluded at the end of day four. At the conclusion of the standard setting portion of the meeting, select panelists from each panel convened for half a day to complete content-specific articulation activities. Table 3-2 presents an overview of the schedule for the standard setting meeting. A detailed meeting agenda can be found in Appendix G.

Table 3-2. Overview of Schedule for OSTP ELA and Mathematics Standard Setting Meeting

Meeting Day	Time	Sessions							
Day 1	AM	General Orientation Session (All Participants)							
	PM	ELA 3 – 4 panel standard setting breakout session	ELA 5 – 6 panel standard setting breakout session	ELA 7 – 8 panel standard setting breakout session	Mathematics 3 – 4	Mathematics 5 – 6	Mathematics 7 – 8		
Day 2	AM				panel standard setting	panel standard setting	panel standard setting		
	PM				breakout session	breakout session	breakout session		
Day 3	AM				Mathematics 3 – 4 panel standard setting breakout session	Mathematics 5 – 6 panel standard setting breakout session	Mathematics 7 – 8 panel standard setting breakout session		
	PM								
Day 4	AM								
	PM								
Day 5	AM	ELA Vertical Articulation Meeting						Mathematics Vertical Articulation Meeting	

3.2.3 Standard Setting Meeting Security

During the meeting, panelists reviewed operational test items, preliminary cut score recommendations, and associated impact data. Due to the nature of this information, security was a critical component of the meeting. Specific procedures were established to ensure the security of all materials was maintained.

As part of the meeting, facilitators reviewed the process for maintaining the security of materials, discussions, and preliminary results from the meeting. Panelists were not permitted to share or discuss secure materials and information outside of meeting rooms. To confirm that the panelists understood and agreed to the security conditions, they signed security and non-disclosure agreements (an example is provided in Appendix H).

To preserve the security of the materials and activities within the Cognia Standard Setting Toolkit, each panelist was provided a Chromebook and unique login credentials. The supporting Cognia psychometrician-controlled panelist access to each section of the Toolkit throughout the meeting. Access to the Toolkit was disabled at the conclusion of the standard setting meeting and the Chromebooks were wiped clean of all data.

3.2.4 ID Matching Standard Setting Procedure

Over the course of four days, panelists engaged in standard setting activities, starting with an opening session on day one. The opening session was followed by the main standard setting session during which panelists received training and engaged in a practice round. Next, panelists engaged in three consecutive judgment rounds for the upper grade associated with their respective grade-band panels, with preparation and discussion between rounds. Panelists then engaged in the same activities for the lower grade associated with their respective grade-band panels. The standard setting portion of the meeting was concluded after the third round for the lower grade, at which point a final workshop evaluation survey was administered.

3.3 Data Review, Cut Score Calculation, and Analyses

3.3.1 Initial Data Review

Given the content-based nature of the standard setting method, it was critically important that panelists remained on task (i.e., made content-based judgments) while engaging in the standard setting process. While the panelist training was targeted and special emphasis was placed on the content-based nature of the work, content specialists were also on hand to review panelists' initial data during the judgment rounds of the standard setting meeting.

Content specialists reviewed panelists' notes on the knowledge, skills, and abilities required by the items, as well as their content-based reasoning to determine whether the panelists were on task. This qualitative evaluation process served as an initial check and allowed for early intervention and adjustment of training procedures as needed.

3.3.2 Cut Score Calculation

To calculate the Proficient and Advanced cut scores during the standard setting meeting, all item-PLD alignment judgments from each panelist were gathered and used as input in a logistic regression calculation (see Appendix C for details).

For example, to calculate the Proficient cut score all items that were aligned to the Basic level were coded as 0, while all items aligned to Proficient and above were coded as 1. The 0/1 coding was required as input for the regression analysis. The result from the above analyses was a theta cut score between the Basic and Proficient performance levels (i.e., the Proficient cut).

To facilitate the Basic level cut score identification, Cognia psychometricians empirically derived the cut score by constructing a miniature Test Characteristic Curve (TCC) based on items that were aligned to the Basic PLD.

Note that during the first round of standard setting, panelists made item-PLD alignments for each item. During rounds 2 and 3, they could revise or retain their item-PLD alignments as they saw fit. Thus, the above process was used to calculate cuts during each round of the standard setting by using the complete set of panelists' judgments for that specific round.

The cut score calculation process was repeated for each grade within each content area.

3.3.3 Analysis Procedure

Cognia psychometricians conducted a series of analyses on the final set of item-PLD alignment data for each grade within each content area. These analyses aimed to identify aberrant and/or outlier data and were performed as follows:

1. Cognia conducted statistical analyses of panelists' item-PLD alignment data by calculating the percent exact, adjacent, and discrepant for each panelist on each performance level, as compared to the results from SDE and Cognia content specialists. Panelists with the least percentage exact were identified as showing statistically aberrant behavior.
2. Content specialist(s) then reviewed the qualitative data for all panelists identified as statistically aberrant. The specialist(s) reviewed panelists' notes on the knowledge, skills, and abilities required by the items, as well as their content-based reasoning to determine if the panelists were on task.
3. After analyses and qualitative review, none of the panelists were identified as both statistically and qualitatively aberrant; therefore, all panelist data were retained.
4. The next phase of the analyses included conducting logistical regression to calculate cut scores. Since the logistical regression method is sensitive to statistical outliers and the presence of such outliers violates the assumptions of the model, an outlier analysis was performed in the form of visual inspection of the initial logistic regression curves to identify any

statistical outliers. Data points identified as statistical outliers were removed before final cut scores were calculated.

5. Final logistic regression analyses were conducted to calculate the Proficient and Advanced cut scores, and the TCC method was used to calculate the Basic cut scores.
6. The resulting cut scores were applied to student data from the spring 2024 administration of the OSTP ELA and mathematics assessments to calculate the impact data (i.e., the percentage of students that were classified into each performance level based on the standard setting cut scores).

3.4 Opening Session and General Orientation

The opening session on day one was the panelists' first opportunity to meet OSDE and Cognia staff. It was important that the panelists felt appreciated and valued for their content expertise. A copy of the orientation session PowerPoint presentation is available in Appendix I.

Cognia representatives set the tone for the workshop in the opening session by

1. welcoming all panelists and expressing appreciation for their commitment to the process.
2. describing the development of the OSTP ELA and mathematics assessments, as well as the associated performance level descriptors.
3. explaining expectations for outcomes they anticipated from the standard setting process.
4. explaining procedures that would be used to review and approve the cut scores.

3.5 Standard Setting Breakout Sessions

After the general orientation session, panelists and relevant staff convened in their assigned grade band and subject-specific breakout sessions. A copy of the general facilitation PowerPoint presentation is available in Appendix E. During the breakout sessions, panelists were organized such that three to four panelists were assigned to each table. Chromebooks, supplied by Cognia and set up for the standard setting, were distributed to all panelists. Facilitators guided panelists through the following activities:

- Overview and introductions
- Experience the test
- Use of the Cognia Standard Setting Toolkit
- Review of the standards and PLDs
- Training on the ID Matching process
- Modeling and practice
- Judgment rounds and feedback
- Final workshop evaluation survey

3.5.1 Overview and Introductions

To begin the breakout sessions, the individuals in each room introduced themselves. After introductions, the facilitator reviewed the security and non-disclosure information. The facilitator then provided a high-level overview of the process. Facilitators also reiterated some of the important points raised during the orientation session as needed. The panelists had an opportunity to ask questions before proceeding.

3.5.2 Experience the Test

After the overview and introductions, panelists experienced the OSTP ELA or mathematics test. Using individual Chromebooks provided by Cognia, panelists were instructed on how to log into their Chromebooks and navigate to the testing platform site. Cognia staff provided panelists with unique login credentials and once they successfully accessed the testing platform, panelists experienced the test the same way students do to become familiar with the test from the students' perspective.

In the interest of time and efficiency, panelists completed the 'Experience the Test' activity only once during the standard setting meeting and a maximum of 45 minutes was allocated for this activity. Except for the ELA 5–6 panel, all panels experienced the test based on the upper grade in their respective breakout session. For example, panelists in the mathematics 3–4 group experienced the grade 4 mathematics test. In the case of the ELA 5–6 panel, panelists experienced the ELA grade 5 test so that panelists in this grade-band panel were exposed to the writing prompt that was part of the grade 5 test (but not part of the grade 6 test; OSTP only administers writing prompts in ELA in grades 5 and 8).

The purpose behind this activity was for panelists to have a sense of the test and testing platform from the student perspective. Panelists were encouraged to experience the test but were directed not to linger over items or spend time evaluating any items.

3.5.3 Use of the Cognia Standard Setting Toolkit

The facilitator guided panelists through the steps needed to log in and access the Cognia Standard Setting Toolkit. Each panelist used their email and an initial assigned password to access the site. After their initial log in, panelists were directed to change their passwords, and then prompted to log back into the system with their new passwords. Their emails and individual passwords were used to access the Toolkit for the duration of the standard setting meeting. Once everyone completed the log in procedure, they viewed an initial screen with tabs that linked to the standards and PLDs.

3.5.4 Review of the Standards and Performance Level Descriptors

Before engaging in their item judgment tasks, panelists studied the standards and the performance level descriptors (PLDs). This important step was designed to ensure that panelists thoroughly understood the knowledge, skills, and abilities (KSAs) needed for students to be classified into the four performance levels (Below Basic, Basic, Proficient, and Advanced).

Throughout the standard setting process, panelists studied the standards and PLDs associated with the OSTP ELA or mathematics assessments relevant to the content area and grades for their respective breakout sessions. Panelists were asked to consider the KSAs detailed in the standards, and how they were reflected in the PLDs.

Facilitators used their PowerPoint training slides and associated scripts to guide panelists through an in-depth review of the PLDs after viewing the standards. Facilitators encouraged panelists to pay attention to the verbiage in the descriptors with the goal of reaching a common understanding of the meaning behind the verbiage, and the elements that distinguished the different performance levels from each other.

Within each content area and grade band, panelists reviewed the standards and PLDs before starting the judgment rounds for each of the two grades. To begin, panelists focused on the standards and PLDs for the upper grade relative to their breakout session. For example, panelists in the ELA 3–4 group first focused on the standards and PLDs for ELA grade 4. Once they completed all training and the standard setting activities (including three rounds of judgment) for grade 4, the panelists in the ELA 3–4 group then moved on to ELA grade 3. Facilitators guided panelists through an in-depth review of the ELA grade 3 standards and PLDs before panelists completed the three judgment rounds for the grade. This same sequential process was followed in each of the six breakout sessions.

The PLDs across all grades and content areas are provided in Appendix A.

3.5.5 Training on the ID-Matching Judgmental Task

Once panelists reviewed and discussed the standards and PLDs associated with the upper grade level within their breakout session (e.g., grade 8 for the mathematics 7–8 group), the facilitator led them through more detailed training on the ID-Matching judgmental task. The facilitator used a customized PowerPoint slide deck and script to explain the following concepts: the ordered item booklet (OIB), how to review items and what information to consider while doing so, and how to make item-descriptor matches. The facilitator emphasized the importance of considering the knowledge, skills, and abilities (KSAs) required by an item, as well as the information in the PLDs, to make their item-descriptor matches.

After explaining the main concepts and the process for making item-descriptor matches, the facilitator provided a high-level description of the round-by-round judgment procedures and what to expect before (i.e., readiness surveys), during (i.e., judgmental tasks and, when relevant, consideration of benchmarks), and after (i.e., presentation of results and discussion) each round.

During the training, facilitators provided clear explanations and directions while ensuring that the panelists had all the information and support needed to undertake the standard setting process. The facilitators encouraged panelists to ask questions during the training but also reminded panelists that they would have the opportunity to practice before beginning the first round. In addition, the facilitators reminded panelists that they would review concepts as needed throughout the standard setting process.

A generalized version of the breakout session PowerPoint presentation is available in Appendix E. Note that the generalized version of the PowerPoint presentation was used as the foundation but was customized for each panel within each content area to account for grade or content specific needs. The PowerPoint presentations were also accompanied by facilitation scripts.

3.5.6 Modeling and Practice

After training on the ID-Matching process, the facilitator provided a brief demonstration of the Cognia Standard Setting Toolkit. A Cognia psychometrician, with dedicated access to a management screen

within the Cognia Standard Setting Toolkit, was responsible for managing aspects related to the system. Once all panelists successfully accessed the system, the Cognia psychometrician advanced all participants to the practice round.

Before proceeding with modeling and practice, the facilitators took some time to make sure panelists knew how to navigate within the Cognia Standard Setting Toolkit. Specifically, the facilitators pointed out that the first screen presented the item list view (a list of items ordered by difficulty) and then demonstrated how to: use the text boxes and item-descriptor dropdown menu, navigate to the item detail view, and use available tabs to access any additional item information when relevant (i.e., stimuli or rubrics).

After the demonstration of the Cognia Standard Setting Toolkit, facilitators proceeded with the practice round which consisted of three sample items. Facilitators used the three sample items to model the judgmental task and guided panelists through making their own item-descriptor matches. During this practice round, the facilitators reinforced the training concepts.

The three sample items were chosen such that (1) none of the items were part of the OIB, (2) the first two items were relatively easy to identify in terms of item-PLD alignment, and (3) the last item was more challenging to identify in terms of item-PLD alignment (i.e., the item was expected to fall in a borderline region). Using sample items that were not part of the OIB allowed the facilitator to avoid undue influence over panelists' judgmental tasks in the standard setting rounds. In addition, the mix of items allowed panelists the opportunity to experience different levels of cognitive load while making their judgments, as would be the case once they considered the full set of items contained in the OIB.

Additionally, in the case of the ELA grades, the sample items were chosen such that a 2-point constructed response item was part of the sample set for grades where these items appeared on the operational test. This allowed panelists the opportunity to be exposed to this item type and practice how to engage with a multi-point item type during judgment rounds. During the modeling and practice session, panelists also had the opportunity for discussion with each other, to ask questions, and become more familiar with the Toolkit.

3.5.7 Judgment Rounds and Feedback

During the main portion of the standard setting meeting, panelists completed three consecutive rounds of judgments for each of the two grades relevant to the content area and grade band of their respective breakout sessions. Each panel began with the upper grade and concluded with the lower grade.

Each judgment round consisted of three distinct sessions: preparation, judgment, and feedback/discussion. This was an iterative process during which the outcomes of each judgment round were considered during the next judgment round. Table 3-3 provides a crosswalk of the activities, analyses, and outcomes for each session within each judgment round.

Table 3-3. Crosswalk of Activities, Analyses, and Outcomes by Judgment Round

Round	Session	Panelist Activities	Analyses	Outcomes
1	Preparation	Training; modeling and practice. Complete Round 1 readiness survey.	Determine if all panelists are ready to proceed.	--
	Judgment	Review all items. Determine the KSAs required to respond to the item and align each item to a PLD.	1. Calculate cut scores and standard errors 2. Calculate % exact agreement on OIB items 3. Create presentation artifacts	1. Initial cut scores 2. Presentation artifacts
	Feedback & Discussion	Discuss Round 1 results: focus on items with the most disagreement.	--	--
2	Preparation	Introduce content-based benchmarks. Complete Round 2 readiness survey.	Determine if all panelists are ready to proceed.	--
	Judgment	Review items (with special attention to items discussed in Round 1 feedback) and make changes to item-PLD alignments as desired.	1. Calculate cut scores and standard errors 2. Calculate % exact agreement on OIB items 3. Create presentation artifacts	1. Revised cut scores 2. Presentation artifacts
	Feedback & Discussion	Discuss Round 2 results: focus on items with the most disagreement, and benchmark regions.	--	--
3	Preparation	Briefly reiterate judgement process. Complete Round 3 readiness survey.	Determine if all panelists are ready to proceed.	--
	Judgment	Review items (with special attention to items discussed in Round 2 feedback) and make changes to item-PLD alignments as desired.	1. Complete series of analyses as described 2. Calculate cut scores and standard errors 3. Calculate associated impact data 4. Create presentation artifacts	1. Cut scores and impact data 2. Presentation artifacts
	Feedback & Discussion	Present final cut scores and impact data to panelists	--	

Readiness Surveys: Before each judgment round, panelists completed a readiness survey that consisted of questions about whether they felt prepared to undertake the upcoming round of judgments. Responses to the survey questions were reviewed before proceeding with the judgment round. If one or more panelists' answers indicated that they were not ready or did not understand one or more of the concepts, such information was relayed to the facilitator who then reviewed the necessary concepts with the panel. Panelists were then asked to complete the readiness survey again. Panelists moved on to the judgment round only when everyone indicated that they were ready to do so. The readiness survey for each round is available in Appendix J.

Feedback and Discussion: After each judgment round, Cognia psychometricians calculated a variety of statistics as described previously. In addition, the psychometricians created presentation artifacts in the form of frequency charts. During the feedback and discussion portion that followed each judgment round, the facilitator presented the frequency chart to the panelists and used it to facilitate table and room discussions. The discussions focused on items that showed the most disagreement between panelists, and panelists were encouraged to share their thoughts and viewpoints. Panelists were also encouraged to refer to training materials (e.g., OIB, item information, PLDs, and standards) as well as their own notes (taken within the Toolkit) throughout this discussion. Panelists were reminded that the goal of the discussion was not to persuade or influence others. Instead, the discussion centered around sharing their own reasoning for their PLD matches and listening to other panelists' reasons as additional information to consider.

Round 1 Judgments

During the first round, panelists worked individually with the PLDs, the standards, and the ordered item booklet (OIB). For each item in the OIB, panelists considered the knowledge, skills, and abilities (KSAs) needed to respond to the item (i.e., asking themselves ‘what does a student need to know and be able to do to respond to this item?’). After identifying the KSAs required by the item, panelists then assigned an item descriptor match (i.e., basic, proficient, or advanced) to the item. They continued in this manner until they reviewed all items in the OIB. All panelists made their round 1 judgments individually and without discussion.

As panelists completed their round 1 work, content specialist(s) were on hand to review their data. Specifically, specialists reviewed panelists’ notes on the knowledge, skills, and abilities required by the items, as well as their content-based reasoning to determine if the panelists were on task. This qualitative evaluation process served as an initial reasonableness check and allowed for early intervention and adjustment of training procedures as needed.

At the conclusion of round 1 judgments, Cognia psychometricians compiled all judgments from all panelists to calculate cut scores, associated standard errors, and various other statistics as described in Section 3.3.3 of this document. In addition, the psychometricians produced the presentation artifact (i.e., a graphical representation of results) that was handed off to the facilitator for use during the round 1 discussion.

Round 2 Judgments and Introduction of Content-based Benchmarks

Before starting the second round of judgments, the panelists were introduced to the content-based benchmarks. Facilitators, with support from a psychometrician, described how the benchmarks were calculated, demonstrated how they would be presented within the Cognia Standard Setting Toolkit, and explained how panelists should consider the information represented by the benchmarks as they engaged in round 2 of the standard setting activities. Panelists were reminded that benchmarks were provided for their consideration, and not to influence their judgments. Next, panelists completed the round 2 readiness survey and once all panelists indicated that they were ready to proceed, they continued to round 2 of the judgment task.

During the second round, panelists once again worked individually with the PLDs, the standards, and the ordered item booklet (OIB). Taking into consideration the feedback and discussion after round 1, as well as the additional information represented by the content-based benchmarks, panelists reviewed their work from round 1. Panelists could keep their judgments from round 1 or revise them. All panelists made their round 2 judgments individually and without discussion. At the conclusion of round 2 judgments, Cognia psychometricians again compiled all judgments from all panelists to calculate cut scores and associated standard errors. In addition, the psychometricians produced the presentation artifact (i.e., a graphical representation of results) that was handed off to the facilitators for use during the round 2 discussion.

Round 3 Judgments

After the round 2 feedback and discussion portion, but before round 3 judgments, panelists once again completed a readiness survey. Once all panelists indicated that they were ready to proceed, they continued to round 3 of the judgment task. During the third round, panelists once again worked individually with the PLDs, the standards, and the ordered item booklet (OIB). Taking into consideration the feedback and discussion after round 2, panelists reviewed their work from round 2. Panelists could keep their judgments from round 2 or revise them. All panelists made their round 3 judgments individually and without discussion.

At the conclusion of the round 3 judgments, Cognia psychometricians again compiled all judgments from all panelists and, using the same procedures already detailed in previous sections, used the panelists' item-PLD judgments to calculate the final cut scores, as well as associated impact data.

3.5.8 Standard Setting Results and Impact Data

The frequency charts of panelists item-PLD judgments across the basic, proficient, and advanced levels for each of the three rounds across all grades and content areas are available in Appendix J. Note that these frequency charts are the same graphical displays that were presented to panelists after each round.

Once panelists completed the standard setting activities for both grades in their respective grade band panels, the final cut scores and associated impact data were calculated. A Cognia psychometrician presented the impact data for the relevant grades to each panel. Table 3-4 shows the standard setting results for ELA and mathematics grades 3–8 and include the OIB page range, theta values, and standard errors associated with the cut scores. In addition, the table includes the impact percentage for each performance level based on the standard setting cut scores.

Table 3-4. Standard Setting Results for OSTP ELA and Mathematics Grades 3-8

Grade	Performance Level	ELA Results				Mathematics Results			
		OIB #	Theta	Standard Error	Impact %	OIB #	Theta	Standard Error	Impact %
3	Below Basic	--	--	--	30.0	--	--	--	27.3
	Basic	3 - 4	-0.890	--	19.2	11 - 12	-1.000	--	36.3
	Proficient	11 - 12	-0.288	0.035	40.0	21 - 22	0.106	0.041	21.0
	Advanced	41 - 42	0.949	0.042	10.8	42 - 43	0.739	0.058	15.4
	Proficient + Advanced	--	--	--	50.8	--	--	--	36.4
4	Below Basic	--	--	--	36.1	--	--	--	31.9
	Basic	4 - 5	-0.700	--	16.7	5 - 6	-0.770	--	28.3
	Proficient	17 - 18	-0.225	0.042	38.1	12 - 13	0.092	0.023	30.7
	Advanced	35 - 36	0.941	0.043	9.1	47 - 48	1.180	0.076	9.1
	Proficient + Advanced	--	--	--	47.2	--	--	--	39.8
5	Below Basic	--	--	--	22.8	--	--	--	35.5
	Basic	5 - 6	-1.120	--	18.0	7 - 8	-0.660	--	27.2
	Proficient	11 - 12	-0.531	0.042	32.7	18 - 19	0.141	0.025	27.0
	Advanced	42 - 43	0.315	0.038	26.5	45 - 46	1.109	0.017	10.3
	Proficient + Advanced	--	--	--	59.2	--	--	--	37.3
6	Below Basic	--	--	--	41.6	--	--	--	42.8
	Basic	2 - 3	-0.670	--	15.6	9 - 10	-0.480	--	20.3
	Proficient	9 - 10	-0.232	0.044	38.6	19 - 20	0.078	0.027	32.6
	Advanced	45 - 46	1.222	0.059	4.2	48 - 49	1.503	0.120	4.2
	Proficient + Advanced	--	--	--	42.8	--	--	--	36.9
7	Below Basic	--	--	--	51.3	--	--	--	54.7
	Basic	8 - 9	-0.380	--	14.3	6 - 7	-0.180	--	16.5
	Proficient	15 - 16	0.015	0.070	32.2	14 - 15	0.314	0.026	15.3
	Advanced	47 - 48	1.551	0.124	2.2	32 - 33	0.881	0.024	13.5
	Proficient + Advanced	--	--	--	34.5	--	--	--	28.8
8	Below Basic	--	--	--	40.3	--	--	--	58.8
	Basic	8 - 9	-0.740	--	20.1	6 - 7	-0.090	--	16.9
	Proficient	10 - 11	-0.207	0.068	37.3	10 - 11	0.416	0.021	13.8
	Advanced	50 - 51	1.351	0.172	2.3	32 - 33	0.971	0.028	10.6
	Proficient + Advanced	--	--	--	39.6	--	--	--	24.4

3.5.9 Standard Setting Workshop Evaluation

At the conclusion of the standard setting meeting, panelists completed a final workshop evaluation survey and gave their feedback on various aspects of the standard setting meeting. Overall, panelists indicated that they felt positive about how Cognia conducted the workshop and about their final recommendations. Specifically, panelists expressed support for the workshop overall; workshop facilitation; training, practice, and the workshop process; the Cognia Standard Setting Toolkit; and other details in the workshop process. The standard setting evaluation questions and results are available in Appendix K.

3.6 Articulation Meetings

At the conclusion of the standard setting meeting, a vertical articulation of the standard setting cut scores was completed. The purpose of the articulation was to allow a subset of panelists from the initial six standard setting panels to review the results from the standard setting meeting and determine if they represented reasonable expectations. This review was completed across grades within each of the two content areas. The two (ELA and mathematics) articulation panels were made up of 3–5 panelists from each of the initial grade-band panels, for a total of 10–12 educators in each content-specific articulation panel. The articulation meetings were co-facilitated by a Cognia psychometrician and either the ELA or mathematics content specialist.

Given the content-based nature of the standard setting, the vertical articulation process consisted of a qualitative review and discussion regarding performance expectations across grades based on the performance level descriptors (see Appendix L for a PowerPoint presentation). Articulation facilitators guided panelists through the following activities:

- Introduction, overview, and key concepts
- Modeling of standard setting panel decisions
- Familiarization with standards, blueprints, and PLDs
- Expectations for between-grade transitions
- Presentation of impact data and discussion
- Articulation workshop evaluation survey

3.6.1 Introduction, Overview, and Key Concepts

Panelists and articulation facilitators briefly introduced themselves. Next, the articulation facilitators provided an overview of the goals and expected outcomes of the articulation meeting. Finally, the facilitators reviewed key concepts related to the articulation process. Specifically, the facilitators addressed the “why” and “how” of the articulation process, as well as the shift to a consensus-based process for articulation compared to the independent judgment process for standard setting. Panelists had the opportunity to ask questions and were encouraged to describe concepts in their own words to facilitate their understanding.

3.6.2 Modeling of Standard Setting Panel Decisions

The content specialist facilitated the modeling and discussion of standard setting panel decisions so that articulation panelists became more familiar with the work done in the panels and grades unfamiliar to them. The standard setting judgment task was modeled for three items (one item for each of the original three grade band panels). As the facilitator presented and modeled each item, articulation panelists followed along in the Cognia Standard Setting Toolkit. Panelists who were participants in specific standard setting panels (e.g., mathematics grade 3–4 panel) were encouraged to share their thoughts and experiences when an item relevant to their specific panel was modeled. Panelists from the other panels were encouraged to ask questions and engage in a discussion with each other throughout the

process. The same process was used until an item relevant to each of the original standard setting panel's work was modeled and discussed.

3.6.3 Familiarization with Standards, Blueprints, and PLDs

Next, panelists engaged in a review and discussion of the standards, test blueprints, and PLDs across the six grades. In the interest of time, the content specialist asked each table to focus on a specific strand or objective. The panelists then engaged in table discussions about their strand/objective across the grades and performance levels. After table discussions, there was a brief discussion among all panelists about the activity and any patterns they noticed across grades. The purpose of this review was to have the panelists familiarize themselves with the standards, blueprints, and PLDs of the grades unfamiliar to them, as well as across the grades at the different performance levels.

3.6.4 Expectations for Between-Grade Transitions

Next, panelists discussed their expectations for student performance relative to between-grade transitions. The discussion was facilitated with guided questions to consider for each of the five grade transitions (i.e., from grade 3 to 4, 4 to 5, 5 to 6, 6 to 7, and 7 to 8). For each grade transition the guided question that panelists were asked to consider followed the same pattern. For example, when considering the first transition (from grade 3 to 4), the question posed to panelists was: "How much more/less challenging is it for 4th graders to demonstrate proficiency in a 4th-grade test (blueprint), assessing 4th-grade standards, as described by 4th grade PLDs THAN IT IS for 3rd graders to demonstrate proficiency on the blueprint, standards, and PLDs of their grade?"

Panelists engaged in a group discussion about the question. Response options for the transition questions were on a Likert-type scale: (1) Much less challenging, (2) less challenging, (3) about the same, (4) more challenging, or (5) much more challenging. Panelists were asked for a consensus response with associated rationale for their response. When consensus could not be reached, the majority response was recorded. Two Cognia staff members took notes of the discussion and recorded responses in the Cognia Standard Setting Toolkit for reference.

The results and qualitative data relevant to the between-grade transition questions and discussions are included in the Cognia Standard Setting Memo (see Appendix N).

3.6.5 Presentation of Impact Data and Discussion

Following the between-grade discussion of performance expectations, panelists were shown impact data across tests from the spring 2024 administration. These impact data were based on the standard setting cut scores. The facilitator led a discussion about the reasonableness of the cut score recommendations, when comparing student performance and performance level classification across tests, in relation to their expectations they had identified in the previous discussion.

With one clear exception, panelists generally agreed that the impact data aligned with the grade transition expectations they had discussed. ELA grade 5 was the only grade for which panelists recommended a significant adjustment. During the grade-transition discussion, the ELA articulation panel determined that it was more challenging for 5th graders to attain proficiency on the 5th grade test than it was for fourth

graders to attain proficiency on the 4th grade test. The following text is an excerpt from the notes in the discussion:

“Especially in standard 3, this seemed to be a big leap; there are harder concepts in the standards. For example, 4.R.1 describing the purpose, vs. 5th grade more evaluation of achieving the purpose. ii. Writing is essentially the same, but reading is more challenging. iii. More inference required in grade 5. iv. Votes for more challenging: consensus”

Given the expectation outlined above, panelists expected impact data to show that fewer students were categorized as proficient and above in 5th grade compared to 4th grade; however, the standard setting impact data showed the opposite with many more students categorized as proficient and above in 5th grade compared to 4th grade. After considerable discussion and review of PLDs and content relative to ELA grades 4, 5, and 6, the articulation panel agreed that an adjustment was needed to bring the result in line with performance expectations they identified.

3.6.6 Closing and Articulation Evaluation Survey

At the end of the articulation meeting, panelists were reminded of the review and approval process their recommendations would go through and the nondisclosure agreement they previously signed. Panelists also completed an evaluation of the process used during the articulation meeting. The articulation evaluation survey questions and results for both articulation panels are available in Appendix M.

Chapter 4. Tasks Completed After the Standard Setting Meeting

Upon conclusion of the standard setting meeting, several important tasks were completed. These tasks centered on the following: reviewing the standard setting process and addressing issues presented by the outcomes, making adjustments based on the articulation panel's recommendations, adjusting cut scores based on policy considerations, and final approval of the operational cut scores. Shortly after the standard setting meeting, Cognia provided SDE with a standard setting memo that included an overview of the standard setting process, as well as the provisional cut scores as recommended by the standard setting panels. A copy of the memo is available in Appendix N.

4.1 Review and Articulation Adjustments

The standard setting literature considers evaluation of the workshop and its results to be another product of the standard setting process (e.g., Reckase and Chen, 2012), as it provides important validity evidence supporting the cut scores that are obtained. To that end, a review and analysis of the standard setting results was conducted. In addition, to provide evidence of the participants' views of the standard setting process, a review and analysis of panelists' feedback on the workshop evaluation surveys was also conducted. This review did not reveal any anomalies in the standard setting process. Panelist responses on the evaluation surveys indicated that panelists: understood the content-based judgement task, tools and feedback at each step throughout the process; had adequate time for training and practice as well as opportunities to ask questions; and felt like the facilitators responded to questions and requests for clarification in a clear and timely manner. In general, participants felt that the standard setting method was appropriate and that their judgments were based on appropriate information and decision making.

Based on the data and recommendations from the ELA and mathematics articulation panels, Cognia psychometricians made adjustments to the cut scores to achieve cross-grade articulation. With the exception of ELA grade 5, minor adjustments were made within the margin of error so as to stay consistent with the standard setting and articulation panel results while still ensuring that expectations were articulated across grades. In the case of ELA grade 5, a significant adjustment was made to align with the articulation panel's recommendation. Please refer to Tables 4 and 8 in the Memo (Appendix N) for the details regarding the ELA and mathematics articulation adjustments. The articulated cut scores were presented to OSDE for their consideration.

4.2 Policy Review and Approval of Final Cut Scores

The final part of the standard setting process consisted of a policy review during which policy makers established the final operational cut scores used to classify students into various performance levels. OSDE engaged in a review and discussion of possible policy adjustments. Based on the recommendations of the Oklahoma Technical Advisory Committee, Cognia psychometricians calculated and then presented adjustment options to OSDE for their consideration. After discussion and review, the OSDE made no policy adjustments to the articulated standard setting results. The full set of cuts, shown in Appendix O, were presented to the Commission for Educational Quality and Accountability (CEQA) at a meeting on July 10, 2024, and were approved for use assigning students to performance levels in the

2023–2024 Oklahoma ELA and mathematics grades 3–8 assessments. See Appendix P for the CEQA PowerPoint Presentation.

4.3 Preparation of Standard Setting Report

Following the final compilation of standard setting results, Cognia prepared this report, which documents the procedures and results of the 2024 Oklahoma Standard Setting Meeting that was held to establish performance standards for the OSTP ELA and mathematics grades 3–8 assessments.

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological Testing*. Washington, DC: American Educational Research Association.
- Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: Establishing and evaluating performance standards on tests*. Thousand Oaks, CA: Sage Publications.
- Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: The MIT Press
- Reckase, M.D. (2001). Innovative methods for helping Standard Setting participants to perform their task: The role of feedback regarding consistency, accuracy, and impact. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives* (pp. 159–173). Mahwah, NJ: Lawrence Erlbaum Associates.
- U.S. Department of Education. (2009). *Standards and assessments peer review guidance: Information and examples for meeting requirements of the No Child Left Behind Act of 2001*. Washington, DC: U.S. Department of Education Office of Elementary and Secondary Education. Retrieved June 10, 2010, from the World Wide Web: www2.ed.gov/policy/elsec/guid/saaprguidance.pdf.

Appendices

APPENDIX—A
PERFORMANCE LEVEL DESCRIPTORS

OSTP ELA Grade 3 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: <i>Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:</i>	OK Policy PLD Proficient: <i>Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:</i>	OK Policy PLD Advanced: <i>Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:</i>
Reading & Writing Process			
3.2.R.1	Identify the main idea of a portion of text and identify key details related to the main idea of a portion of the text.	Determine the main idea and supporting details of a text.	Explain how key details support the main idea of a text.
3.2.R.2	Identify elements of fiction and nonfiction texts.	Identify elements of various genres in fiction, poetry, and nonfiction texts.	Compare elements of various genres in fiction, poetry, and nonfiction texts and provide supporting details.
3.2.R.3	Summarize portions of a text or sequence the main events of a story (first, next, last).	Summarize and sequence the important events of a story.	Analyze a story to summarize and correctly sequence the events in a story; evaluate the best summary; explain why details are included in a summary.
3.2.R.4	Summarize facts and details in a portion of an informational text.	Summarize facts and details from an informational text.	Explain why certain facts and details are included in an informational text.
3.2.W.1	Identify the parts of and use the writing process to prewrite, organize, and develop narrative, informative, and opinion drafts of a single paragraph.	Prewrite, organize, and develop narrative, informative, and opinion drafts that display evidence of paragraphing.	
3.2.W.2	Use a process to revise content in a paragraph for correct organization (e.g., logical order and transitions) and clarity.	Revise content for clarity, coherence, and organization (e.g., logical order and transitions).	
3.2.W.3	Edit drafts of a sentence for punctuation (end marks), capitalization (beginnings of sentences), and correctly spelled high-frequency words, using resources as needed.	Edit drafts for punctuation, capitalization, and correctly-spelled grade-level words, using resources as needed.	
Critical Reading & Writing			
3.3.R.1	Identify if the author's purpose is to entertain, inform, or persuade.	Determine if the author's purpose is to entertain, inform, or persuade.	Analyze a text to determine whether the author's purpose is to entertain, inform, or persuade.

Objective	Basic	Proficient	Advanced
3.3.R.2	Identify features of first- or third-person point-of-view texts.	Determine whether a grade-level literary text is narrated in first- or third-person point of view.	Determine whether a grade-level literary text is narrated in first- or third-person point of view and provide evidence to support their determination.
3.3.R.3	Identify literary elements: <ul style="list-style-type: none"> • setting • plot • characters • characterization 	Find textual examples of literary elements: <ul style="list-style-type: none"> • setting • plot • characters • characterization 	Identify the effect of literary elements: <ul style="list-style-type: none"> • setting • plot • characters • characterization
3.3.R.4	Identify examples of literary devices: <ul style="list-style-type: none"> • personification • simile • alliteration • onomatopoeia 	Find examples of literary devices: <ul style="list-style-type: none"> • personification • simile • alliteration • onomatopoeia 	Identify the effect of literary devices: <ul style="list-style-type: none"> • personification • simile • alliteration • onomatopoeia
3.3.R.5	Answer simple inferential questions from a portion of a text and use evidence from a portion of a text to support inferences.	Answer inferential questions, using a text to support answers.	Answer complex inferential questions, using a text to support answers.
3.3.R.6	Identify fact and opinion in an informational text.	Distinguish fact from opinion in an informational text.	Identify how the fact or opinion supports the main idea of a text.
3.3.R.7	Identify the structure of a portion of an informational text: <ul style="list-style-type: none"> • problem/solution • description • sequential 	Describe the structure of a portion of an informational text: <ul style="list-style-type: none"> • problem/solution • description • sequential 	Analyze the structure of a portion of an informational text: <ul style="list-style-type: none"> • problem/solution • description • sequential
3.3.W.1	Write simple narratives with support (e.g., graphic organizers) that incorporate: <ul style="list-style-type: none"> • setting • plot • characters • characterization 	Write narratives that incorporate: <ul style="list-style-type: none"> • setting • plot • characters • characterization 	Write complex narratives reflecting real or imagined experiences that include: <ul style="list-style-type: none"> • setting • plot • characters • characterization
3.3.W.2	With support (e.g., graphic organizers), write facts about a subject, including a main idea with supporting details, in a paragraph.	Write facts about a subject, including a main idea with supporting details, in multiple paragraphs with transitional words and phrases.	Write complex facts about a subject, including a main idea with supporting details, in multiple paragraphs with transitional words and phrases.
3.3.W.3	With support (e.g., graphic organizers), write an opinion about a topic and provide relevant evidence as support in a paragraph.	Write an opinion about a topic and provide relevant evidence as support in multiple paragraphs with transitional words and phrases.	Write a complex opinion about a topic and provide relevant evidence as support in multiple paragraphs with transitional words and phrases.
Vocabulary			
3.4.R.1	Identify relationships among words, including synonyms, and antonyms.	Identify relationships among words, including synonyms, antonyms, homophones, and homographs.	Identify relationships among words, including synonyms, antonyms, homophones, and homographs.

Objective	Basic	Proficient	Advanced
3.4.R.2	Use context clues to clarify the meaning of words.	Use context clues to clarify the meaning of words.	Use context clues to clarify the meaning of words.
3.4.R.3	Use word parts (e.g., simple affixes, simple Anglo-Saxon roots, stems) to identify the meaning of words.	Use word parts (affixes, Anglo-Saxon roots, and stems) to define and determine the meaning of new words.	Define and determine the meaning of new words by using familiar word parts including affixes, Anglo-Saxon roots, and stems.
3.4.R.4	Consult reference materials (e.g., dictionaries, glossaries) to identify the meaning of words in a text.	Consult reference materials (e.g., dictionaries, glossaries, thesauruses) to demonstrate comprehension of the words in a text.	
3.4.R.5		Acquire new grade-level vocabulary, relate new words to prior knowledge, and apply vocabulary in various contexts.	Use new grade-level vocabulary, relate new words to prior knowledge, and apply vocabulary in various contexts.
3.4.W.1	Use high-frequency vocabulary in writing to clearly communicate ideas.	Use grade-level vocabulary in writing to clearly communicate ideas.	Use above-grade-level vocabulary in writing to clearly communicate ideas.
3.4.W.2	Use precise and vivid basic vocabulary in writing.	Use precise and vivid grade-level vocabulary in writing for the intended mode and effect on the audience.	
Language			
3.5.R.1	Recognize simple sentences.	Recognize simple and compound sentences.	Define the features of simple and compound sentences in grade-appropriate texts.
3.5.R.2	Recognize parts of speech in basic sentences: <ul style="list-style-type: none"> • concrete and abstract nouns • different types of verbs (i.e., action, linking, helping) <ul style="list-style-type: none"> • the subject and predicate of a sentence <ul style="list-style-type: none"> • adjectives • prepositions • possessive pronouns • adverbs • coordinating conjunctions (i.e., and, but, or) 	Recognize parts of speech in sentences: <ul style="list-style-type: none"> • concrete, abstract, and possessive nouns • different types of verbs (i.e., action, linking, helping) and their roles in a sentence • the complete subject and complete predicate of a sentence <ul style="list-style-type: none"> • possessive adjectives • prepositional phrases • possessive pronouns and the nouns they replace • coordinating conjunctions (i.e., for, and, nor, but, or, yet, so) • adverbs of frequency (e.g., always, often, never) 	Analyze parts of speech in complex sentences: <ul style="list-style-type: none"> • concrete, abstract, and possessive nouns • different types of verbs (i.e., action, linking, helping) and their roles in a sentence • the complete subject and complete predicate of a sentence <ul style="list-style-type: none"> • possessive adjectives • prepositional phrases • possessive pronouns and the nouns they replace • coordinating conjunctions (i.e., for, and, nor, but, or, yet, so) • adverbs of frequency (e.g., always, often, never)
3.5.W.1	Compose simple and compound declarative, interrogative, imperative, and exclamatory sentences.	Compose simple and compound declarative, interrogative, imperative, and exclamatory sentences, avoiding and correcting fragments.	
3.5.W.2	Use nouns, verbs, adjectives, and adverbs.	Use nouns, verbs, adjectives, prepositions, and adverbs to add clarity and variety to their writing.	Explain the effect of nouns, verbs, adjectives, prepositions, and adverbs in their writing.
3.5.W.3	Capitalize titles of respect, words in titles, and geographical names.	Capitalize and punctuate titles of respect, words in titles, and geographical names.	Recognize and correct errors in capitalization and punctuation in titles of respect, words in titles, and geographical names.

Objective	Basic	Proficient	Advanced
3.5.W.4	Use periods with declarative and imperative sentences and question marks with interrogative sentences.	Use periods with declarative and imperative sentences, question marks with interrogative sentences, and exclamation points with exclamatory sentences.	Recognize and correct errors in punctuation: periods with declarative and imperative sentences, question marks with interrogative sentences, and exclamation points with exclamatory sentences.
3.5.W.5	Use apostrophes to form simple contractions (e.g., can't, doesn't, isn't) and to show possession.	Use apostrophes to form complex contractions (e.g., should've, won't, y'all) and to show possession.	Recognize and correct errors in apostrophes when forming complex contractions (e.g., should've, won't, y'all) and to show possession.
3.5.W.6	Identify the placement of commas when using a coordinating conjunction and when separating individual words in a series.	Use commas before a coordinating conjunction and to separate individual words in a series.	Recognize and correct errors in comma usage before a coordinating conjunction and to separate individual words in a series.
3.5.W.7		Use a colon to indicate time.	
3.5.W.8	Explain why quotation marks are used.	Use quotation marks to indicate dialogue.	Recognize and revise errors in quotation mark usage when indicating dialogue.
Research			
3.6.R.1	Conduct research to answer assigned questions and to build knowledge.	Conduct research to answer questions, including self-generated questions, and to build knowledge.	Conduct research and evaluate if research questions are fully answered.
3.6.R.2	Identify some text features (e.g., captions, subheadings, charts) to comprehend informational texts.	Identify and use text features (e.g., graphics, captions, subheadings, italics, charts, tables, legends) to comprehend informational texts.	Analyze text features (e.g., graphics, captions, subheadings, italics, charts, tables, legends) to comprehend complex informational texts.
3.6.R.3	Identify relevant sources.	Begin to determine the relevance of the information gathered.	Determine the relevance of more complex information gathered.
3.6.W.1	Identify questions related to a topic.	Choose a topic of interest and generate several questions about it for research.	Choose a topic of interest and generate several valid questions about it for research.
3.6.W.2	With support (e.g., a partially completed graphic organizer), organize information found during research and follow a modified citation style (i.e., author, title, publication year).	Begin to organize information found during research, following a modified citation style (i.e., author, title, publication year).	

OSTP ELA Grade 4 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: <i>Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:</i>	OK Policy PLD Proficient: <i>Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:</i>	OK Policy PLD Advanced: <i>Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:</i>
Reading & Writing Process			
4.2.R.1	Identify the main idea and key supporting details of a text.	Determine the key details that support the main idea of a text.	Evaluate the key details that support the main idea of a text.
4.2.R.2	Identify features of fiction, poetry, and nonfiction to distinguish various genres.	Compare fiction, poetry, and nonfiction to distinguish various genres.	Compare and explain the differences in fiction, poetry, and nonfiction to distinguish various genres.
4.2.R.3	Summarize or sequence the important events in a portion of a story (e.g., first, next, last).	Summarize and sequence the important events of a story.	Analyze a story, summarize and sequence the important events of a story, evaluate for the best summary, and explain why certain details should be included in a summary.
4.2.R.4	Summarize facts and details from a portion of an informational text.	Summarize facts and details from an informational text.	Explain why certain facts and details from an informational text are included in a summary.
4.2.W.1	Identify the parts of and use the writing process to: prewrite, organize, and develop narrative, informative, and opinion drafts of a paragraph.	Use the writing process to prewrite, organize, and develop narrative, informative, and opinion drafts that display evidence of paragraphing.	Use the writing process to prewrite by selecting a strategy, organize by selecting a particular structure, and develop narrative, informative, and opinion drafts that display evidence of paragraphing.
4.2.W.2	Revise content in a paragraph for clarity and organization (e.g., logical order).	Revise content for clarity, coherence, and organization (e.g., logical order and transitions).	Revise content for clarity (using precise language geared toward the audience), coherence, and organization (e.g., logical order and transitions) using effective language.
4.2.W.3	Edit drafts of a sentence for punctuation (end marks), capitalization (beginnings of sentences), and correctly spelled grade-level words, using resources as needed.	Edit drafts for punctuation, capitalization, and correctly spelled grade-level words, using resources as needed.	
Critical Reading & Writing			
4.3.R.1	Identify the author's purpose (i.e., entertain, inform, persuade).	Determine the author's purpose (i.e., entertain, inform, persuade) by identifying key details.	Determine the author's purpose (i.e., entertain, inform, persuade) and determine how key details reveal the author's purpose was achieved.
4.3.R.2	Identify features of first- or third-person point of view.	Determine whether a grade-level literary text is narrated in first- or third-person point of view.	Determine whether a grade-level literary text is narrated in first- or third-person point of view and provide evidence to support their determination.

Objective	Basic	Proficient	Advanced
4.3.R.3	<p>Identify textual evidence of literary elements:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict 	<p>Find textual evidence of literary elements:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict 	<p>Determine the effect of literary elements:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict
4.3.R.4	<p>Identify textual evidence of literary devices:</p> <ul style="list-style-type: none"> ● metaphor ● idiom ● personification ● hyperbole ● simile ● alliteration ● onomatopoeia 	<p>Find textual evidence of literary devices:</p> <ul style="list-style-type: none"> ● metaphor ● idiom ● personification ● hyperbole ● simile ● alliteration ● onomatopoeia 	<p>Determine the effect of literary devices:</p> <ul style="list-style-type: none"> ● metaphor ● idiom ● personification ● hyperbole ● simile ● alliteration ● onomatopoeia
4.3.R.5	Answer simple inferential questions and use evidence from a text to support answers.	Answer inferential questions using evidence from one or more texts to support answers.	Answer complex inferential questions using evidence from one or more texts to support answers.
4.3.R.6	Distinguish fact from opinion in an informational text and identify how reasons and facts support specific points.	Distinguish fact from opinion in an informational text and explain how reasons and facts support specific points.	Distinguish fact from opinion in an informational text and draw a conclusion about their effectiveness.
4.3.R.7	<p>Identify the structures of an informational text:</p> <ul style="list-style-type: none"> ● cause/effect ● problem/solution ● description ● sequential 	<p>Distinguish the structures of an informational text:</p> <ul style="list-style-type: none"> ● cause/effect ● problem/solution ● description ● sequential 	<p>Determine the structure of an informational text:</p> <ul style="list-style-type: none"> ● cause/effect ● problem/solution ● description ● sequential
4.3.W.1	<p>Compose simple narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include a plot with a climax and resolution ● include characters who overcome conflicts and use dialogue ● unfold in chronological sequence ● use some sentence variety and sensory details to create interest ● replicate literary elements and/or literary devices from mentor texts 	<p>Compose narratives that reflect real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include plots with a climax and resolution ● include developed characters who overcome conflicts and use dialogue <ul style="list-style-type: none"> ● use a consistent point of view ● unfold in chronological sequence ● use sentence variety, sensory details, and vivid language to create interest ● model literary elements and/or literary devices from mentor texts 	<p>Compose complex narratives that reflect real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include plots with a climax and resolution ● include developed characters who overcome conflicts and use dialogue <ul style="list-style-type: none"> ● use a consistent point of view ● unfold in chronological sequence ● use sentence variety, sensory details, and vivid language to create interest ● model literary elements and/or literary devices from mentor texts

Objective	Basic	Proficient	Advanced
4.3.W.2	<p>Compose simple informative essays that:</p> <ul style="list-style-type: none"> • introduce and develop a topic • incorporate evidence (e.g., specific facts, examples) • maintain an organized structure • replicate literary devices from mentor texts 	<p>Compose informative essays that:</p> <ul style="list-style-type: none"> • introduce and develop a topic • incorporate evidence (e.g., specific facts, examples) • maintain an organized structure with transitional words and phrases • use sentence variety and word choice to create interest • model literary devices from mentor texts 	<p>Compose complex informative essays that:</p> <ul style="list-style-type: none"> • introduce and develop a topic • incorporate and explain evidence (e.g., specific facts, examples) • maintain an organized structure with complex transitional words and phrases • use sentence variety and precise word choice to create interest • model literary devices from mentor texts
4.3.W.3	<p>Write simple opinion essays that:</p> <ul style="list-style-type: none"> • introduce a topic and state an opinion • incorporate text-based evidence to support the opinion • maintain an organized structure with simple transitional words and phrases 	<p>Write opinion essays that:</p> <ul style="list-style-type: none"> • introduce a topic and state an opinion • incorporate relevant, text-based evidence to support the opinion • use sentence variety and word choice to create interest • maintain an organized structure with transitional words and phrases 	<p>Write complex opinion essays that:</p> <ul style="list-style-type: none"> • introduce a topic and state an opinion • incorporate and explain relevant, text-based evidence to support the opinion • use sentence variety and precise word choice to create interest • maintain an organized structure with complex transitional words and phrases
Vocabulary			
4.4.R.1	Identify relationships among words, including synonyms, antonyms, homophones, and homographs.	Identify relationships among words, including synonyms, antonyms, analogies, homophones, and homographs.	Identify relationships among words, including synonyms, antonyms, analogies, homophones, and homographs.
4.4.R.2	Use context clues to clarify the meaning of words.	Use context clues to clarify the meaning of words.	Use context clues to clarify the meaning of words.
4.4.R.3	Use word parts (e.g., simple affixes, simple Latin roots, stems) to define and determine the meaning of new words.	Use word parts (e.g., affixes, Latin roots, stems) to define and determine the meaning of new words.	Use word parts (e.g., complex affixes, complex Latin roots, stems) to define and determine the meaning of new words.
4.4.R.4	Consult reference materials (e.g., dictionaries, glossaries) to identify the meaning of words in a text.	Consult reference materials (e.g., dictionaries, glossaries, thesauruses) to comprehend the words in a text.	
4.4.R.5		Acquire new grade-level vocabulary, relate new words to prior knowledge, and apply vocabulary in various contexts.	
4.4.W.1		Use grade-level vocabulary in writing to clearly communicate ideas.	
4.4.W.2	Use precise and vivid vocabulary in writing.	Use precise and vivid vocabulary in writing for the intended mode and effect on the audience.	
Language			
4.5.R.1	Recognize simple and compound sentences.	Recognize simple and compound sentences.	Define the features of simple and compound sentences.

Objective	Basic	Proficient	Advanced
4.5.R.2	<p>Recognize parts of speech in sentences:</p> <ul style="list-style-type: none"> • possessive nouns • irregular verbs • subject of a verb • comparative adjectives • prepositional phrases • possessive pronouns • coordinating conjunctions • comparative adverbs 	<p>Recognize parts of speech in sentences:</p> <ul style="list-style-type: none"> • irregular possessive nouns (e.g., children's) • irregular and past participle verbs and verb tense to identify settings, times, and sequences <ul style="list-style-type: none"> • subject and verb agreement • comparative and superlative adjectives <ul style="list-style-type: none"> • prepositional phrases • possessive pronouns and the nouns they replace (i.e., antecedents) <ul style="list-style-type: none"> • coordinating conjunctions • comparative and superlative adverbs • interjections 	
4.5.W.1	Compose simple declarative, interrogative, imperative, and exclamatory sentences, and recognize fragments.	Compose simple and compound declarative, interrogative, imperative, and exclamatory sentences, avoiding and correcting fragments.	
4.5.W.2	Use nouns, verbs, adjectives, prepositions, and adverbs to add variety to their writing.	Use nouns, verbs, adjectives, prepositions, and adverbs to add clarity and variety to their writing.	Explain why nouns, verbs, adjectives, prepositions, and adverbs are included in their writing.
4.5.W.3	Recognize or correct errors in subject and verb agreement.	Recognize and correct errors in subject and verb agreement.	Compose sentences with correct subject and verb agreement.
4.5.W.4	Capitalize familial relations and proper adjectives.	Capitalize familial relations, proper adjectives, conventions of letter writing, and the first letter of a quotation.	
4.5.W.5		Use periods with declarative and imperative sentences, question marks with interrogative sentences, and exclamation points with exclamatory sentences.	Recognize and revise errors in end punctuation including: periods with declarative and imperative sentences, question marks with interrogative sentences, and exclamation points with exclamatory sentences.
4.5.W.6	Use apostrophes to show possession of singular nouns.	Use apostrophes to show possession of singular and plural nouns and recognize and remove apostrophes used to form plurals.	Recognize and revise errors in apostrophe use to show possession of singular and plural nouns and recognize and remove apostrophes used to form plurals.
4.5.W.7	Use commas to separate individual words in a series.	Use commas in greetings and closings in letters and emails, to separate individual words in a series, and to indicate dialogue.	Recognize and revise errors in comma usage in greetings and closings in letters and emails, to separate individual words in a series, and to indicate dialogue.
4.5.W.8	Recognize where a colon should be placed when introducing a list (e.g., Deb only needed three things from the grocery store: milk, eggs, and bread.).	Use a colon to introduce a list (e.g., Deb only needed three things from the grocery store: milk, eggs, and bread.).	Recognize and revise errors when using a colon to introduce a list (e.g., Deb only needed three things from the grocery store: milk, eggs, and bread.).
4.5.W.9	Recognize where quotation marks belong when being used to indicate dialogue and titles of works; explain why quotation marks are used.	Use quotation marks to indicate dialogue, quoted material, and titles of works.	Recognize and revise errors when using quotation marks to indicate dialogue, quoted material, and titles of works.

Objective	Basic	Proficient	Advanced
4.5.W.10	Recognize the correct way to use underlining to indicate titles of works.	Use underlining or italics to indicate titles of works.	Recognize and revise errors when using underlining or italics to indicate titles of works.
Research			
4.6.R.1	Conduct research to answer questions, including self-generated questions, and to build knowledge, using one source (e.g., visual and text reference sources, electronic resources, and/or interviews).	Conduct research to answer questions, including self-generated questions, and to build knowledge, using multiple sources (e.g., visual and text reference sources, electronic resources, and/or interviews).	Conduct research to answer questions, including self-generated questions, and to evaluate knowledge, using multiple sources (e.g., visual and text reference sources, electronic resources, and/or interviews).
4.6.R.2	Identify and/or use some text features (e.g., graphics, captions, subheadings, italicized words, charts, tables, legends) to comprehend informational texts.	Identify and use text features (e.g., graphics, captions, headings/subheadings, bold/italicized words, charts, tables, legends) to comprehend informational texts.	Analyze text features (e.g., graphics, captions, headings/subheadings, bold/italicized words, charts, tables, legends) to comprehend informational texts.
4.6.R.3	Determine the relevance of sources.	Determine the relevance of the information gathered.	Explain the relevance of the information gathered.
4.6.W.1	Identify a viable research question about a topic.	Generate a viable research question about a topic.	Generate more than one viable research question about a topic.
4.6.W.2	With support (e.g., a graphic organizer) organize information found during research.	Organize information found during research, following a modified citation style (i.e., author, title, publication year).	

OSTP ELA Grade 5 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:
Reading & Writing Process			
5.2.R.1	Identify key supporting details that support the main idea of a text.	Explain how key supporting details support the main idea of a text.	Analyze how key supporting details support the main idea of a text.
5.2.R.2	Use features of fiction, poetry, and nonfiction texts to distinguish various genres.	Identify details in fiction, poetry, and nonfiction texts to distinguish various genres.	Compare details within or across fiction, poetry, and nonfiction texts to distinguish various genres.
5.2.R.3	Summarize or sequence the important events from a portion of a story.	Summarize and sequence the important events of a story.	Analyze a complex story, summarize and sequence the important events of a story, evaluate for the best summary, and explain why certain details should be included in a summary.
5.2.R.4	Summarize facts and details from portions of an informational text.	Summarize facts and details from an informational text.	Summarize facts and details from a complex informational text; evaluate for the best summary.
5.2.W.1	Identify the parts of and use the writing process to: prewrite, organize, and develop narrative, informative, and opinion drafts of a paragraph.	Use a recursive process to prewrite, organize, and draft multi-paragraph narrative, informative, and opinion drafts.	Show knowledge of a recursive process to prewrite and organize for intended purpose, and draft multi-paragraph narrative, informative, and opinion drafts.
5.2.W.2	Revise content in a paragraph for clarity and organization (e.g., logical order and transitions).	Revise content for clarity, coherence, and organization (e.g., logical order and transitions).	Revise content for clarity (using precise language geared toward the audience), coherence, and organization (e.g., logical order and effective use of transitions).
5.2.W.3	Edit drafts of a sentence for punctuation (end marks), capitalization (beginnings of sentences), and correctly spelled grade-level words, using resources as needed.	Edit drafts for punctuation, capitalization, and correctly spelled grade-level words, using resources as needed.	
Critical Reading & Writing			
5.3.R.1	Identify the author's purpose (i.e., entertain, inform, persuade).	Determine the author's purpose (i.e., entertain, inform, persuade), and draw conclusions to determine if the author's purpose was achieved.	Analyze key details to determine if the author's purpose was achieved.

Objective	Basic	Proficient	Advanced
5.3.R.2	Determine whether a grade-level literary text is narrated in first- or third-person point of view.	Determine whether a grade-level literary text is narrated in first- or third-person point of view (limited and omniscient) and describe its effect.	Analyze key details to determine if the text is narrated first- or third-person point of view (limited and omniscient) and describe its effect.
5.3.R.3	Identify textual evidence of literary elements: <ul style="list-style-type: none"> • setting • plot • characters (i.e., protagonist, antagonist) • characterization • conflict 	Determine how literary elements contribute to the meaning of a literary text: <ul style="list-style-type: none"> • setting • plot • characters (i.e., protagonist, antagonist) • characterization • conflict • theme 	Using textual evidence, explain how literary elements contribute to the meaning of a literary text: <ul style="list-style-type: none"> • setting • plot • characters (i.e., protagonist, antagonist) • characterization • conflict • theme
5.3.R.4	Identify textual evidence of literary devices: <ul style="list-style-type: none"> • imagery • metaphor • idiom • personification • hyperbole • simile • alliteration • onomatopoeia 	Determine how literary devices contribute to the meaning of a text: <ul style="list-style-type: none"> • imagery • metaphor • idiom • personification • hyperbole • simile • alliteration • onomatopoeia 	Using textual evidence, explain how literary devices contribute to the meaning of a text: <ul style="list-style-type: none"> • imagery • metaphor • idiom • personification • hyperbole • simile • alliteration • onomatopoeia
5.3.R.5	Analyze ideas in a portion of a text, providing textual evidence to support their inferences.	Analyze ideas in one or more texts, providing textual evidence to support their inferences.	Draw evaluative conclusions from one or more texts, providing textual evidence to support their inferences.
5.3.R.6	Identify fact or opinion in an informational text and locate reasons and facts that support specific points.	Distinguish fact from opinion in an informational text and explain how reasons and facts support specific points.	Distinguish relevant fact from opinion in an informational text and explain how reasons and facts support specific points using supporting evidence from the informational text.
5.3.R.7	Identify the structures of informational texts: <ul style="list-style-type: none"> • compare/contrast • cause/effect • problem/solution • description • sequential 	Distinguish the structures of informational texts: <ul style="list-style-type: none"> • compare/contrast • cause/effect • problem/solution • description • sequential 	Analyze the structures of informational texts and provide supporting evidence for that analysis: <ul style="list-style-type: none"> • compare/contrast • cause/effect • problem/solution • description • sequential

Objective	Basic	Proficient	Advanced
5.3.W.1	<p>Compose simple narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> include plots with a climax and resolution include developed characters who overcome conflicts and use dialogue <ul style="list-style-type: none"> unfold in chronological sequence use some sentence variety, sensory details, and vivid language to create interest replicate literary elements and/or literary devices from mentor texts 	<p>Compose narratives experiences reflecting real or imagined that:</p> <ul style="list-style-type: none"> include plots with a climax and resolution include developed characters who overcome conflicts and use dialogue <ul style="list-style-type: none"> use a consistent point of view unfold in chronological sequence use sentence variety, sensory details, and vivid language to create interest model literary elements and/or literary devices from mentor texts 	<p>Compose complex narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> include plots with a climax and resolution including developed characters who overcome conflicts and use dialogue <ul style="list-style-type: none"> use a consistent point of view unfold in chronological sequence use sentence variety, sensory details, and vivid language to create interest model literary elements and/or literary devices from mentor texts
5.3.W.2	<p>Compose simple informative essays that:</p> <ul style="list-style-type: none"> introduce and develop a topic include evidence (e.g., specific facts, examples, charts, and graphs) maintain an organized structure with simple transitional words and phrases use some sentence variety and word choice to create interest replicate literary devices from mentor texts 	<p>Compose informative essays that:</p> <ul style="list-style-type: none"> introduce and develop a topic incorporate evidence (e.g., specific facts, examples, charts, and graphs) maintain an organized structure with transitional words and phrases use sentence variety and word choice to create interest model literary devices from mentor texts 	<p>Compose complex informative essays that:</p> <ul style="list-style-type: none"> introduce and develop a topic incorporate and explain evidence (e.g., specific facts, examples, charts, and graphs) maintain an organized structure with complex transitional words and phrases use sentence variety and precise word choice to create interest model literary devices from mentor texts
5.3.W.2	<p>Write simple opinion essays that:</p> <ul style="list-style-type: none"> introduce a topic and state an opinion include text-based evidence use some sentence variety and word choice to create interest organize writing in a logical sequence with simple transitional words and phrases 	<p>Write opinion essays that:</p> <ul style="list-style-type: none"> introduce a topic and state a clear opinion incorporate relevant, text-based evidence to support the opinion use sentence variety and word choice to create interest organize writing in a logical sequence with transitional words and phrases 	<p>Write complex opinion essays that:</p> <ul style="list-style-type: none"> introduce a topic and state a clear opinion incorporate relevant, text-based evidence to support the opinion use sentence variety and word choice to create interest organize writing in a logical sequence with transitional words and phrases
Vocabulary			
5.4.R.1	Identify relationships among words, including synonyms, antonyms, simple analogies, homophones, and homographs.	Identify relationships among words, including synonyms, antonyms, analogies, homophones, and homographs.	Identify relationships among words, including synonyms, antonyms, complex analogies, homophones, and homographs.
5.4.R.2	Use context clues to clarify the meaning of basic words.	Use context clues to clarify the meaning of words.	Use context clues to clarify the meaning of words and identify supporting evidence.
5.4.R.3	Use word parts (e.g., simple affixes, simple Latin roots, stems) to define and determine the meaning of new words.	Use word parts (e.g., affixes, Latin roots, stems) to define and determine the meaning of new words.	Use word parts (e.g., complex affixes, complex Latin roots, stems) to define and determine the meaning of new words.

Objective	Basic	Proficient	Advanced
5.4.R.4	Choose reference materials (e.g., dictionaries, glossaries, thesauruses) to identify the meanings of words in a text.	Consult reference materials (e.g., dictionaries, glossaries, thesauruses) to comprehend the words in a text.	
5.4.R.5		Acquire new grade-level vocabulary, relate new words to prior knowledge, and apply vocabulary in various contexts.	
5.4.W.1		Use grade-level vocabulary in writing to clearly communicate ideas.	
5.4.W.2	Use precise and vivid vocabulary in writing.	Use precise and vivid vocabulary in writing for the intended mode and effect on the audience.	
Language			
5.5.R.1	Recognize simple and compound sentences.	Recognize simple, compound, and complex (i.e., independent and dependent clauses) sentences.	Determine and explain whether sentences are simple, compound, or complex (i.e., independent and dependent clauses) and identify independent and dependent clauses.
5.5.R.2	Recognize parts of speech in simple sentences: <ul style="list-style-type: none"> • nouns • verb tense to identify settings, times, sequences, and conditions • subject and verb agreement • adjectives • prepositional phrases • intensive pronouns and their antecedents • coordinating conjunctions • adverbs • interjections 	Recognize and explain the impact on meaning of parts of speech in sentences: <ul style="list-style-type: none"> • nouns • verb tense to identify settings, times, sequences, and conditions • subject and verb agreement • adjectives • prepositional phrases • intensive pronouns and their antecedents • coordinating conjunctions • adverbs • interjections 	
5.5.W.1		Compose simple, compound, and complex (i.e., independent and dependent clauses) sentences.	
5.5.W.2		Use nouns, verbs, adjectives, prepositions, adverbs, and pronouns to add clarity and variety to their writing.	
5.5.W.3	Recognize the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and inappropriate shifts in pronoun number and person.	Recognize and correct the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and inappropriate shifts in pronoun number and person.	

Objective	Basic	Proficient	Advanced
5.5.W.6		Use the correct forms of it's/its, you're/your, and they're/there/their.	Recognize and revise errors in the incorrect use of it's/its, you're/your, and they're/there/their.
5.5.W.7	Use commas to separate individual words in a series and to indicate dialogue.	Use commas to separate individual words in a series, to indicate dialogue, and to separate the independent and dependent clauses in a complex sentence.	Recognize and revise errors in comma usage to separate individual words in a series, to indicate dialogue, and to separate the independent and dependent clauses in a complex sentence.
5.5.W.8	Identify sentences that correctly use a colon to introduce a list.	Use a colon to introduce a list.	Recognize and revise errors in colon use to introduce a list.
5.5.W.9	Identify sentences that correctly use quotation marks to indicate dialogue, quoted material, and titles of works.	Use quotation marks to indicate dialogue, quoted material, and titles of works.	Recognize and revise errors when using quotation marks to indicate dialogue, quoted material, and titles of works.
5.5.W.10	Identify sentences that correctly use underlining or italics to indicate titles of works.	Use underlining or italics to indicate titles of works.	Recognize and revise errors when using underlining or italics to indicate titles of works.
5.5.W.11	Identify sentences that correctly use a semicolon to punctuate compound sentences.	Use a semicolon to punctuate compound sentences.	Recognize and revise errors when using a semicolon to punctuate compound sentences.
Research			
5.6.R.1	Conduct research to answer questions, including self-generated questions, and to build knowledge, using one source (e.g., visual and text reference sources, electronic resources, and/or interviews).	Conduct research to answer questions, including self-generated questions, and to build knowledge, using multiple sources (e.g., visual and text reference sources, electronic resources, and/or interviews).	Conduct research to answer questions, including self-generated questions, and to evaluate knowledge, using multiple sources (e.g., visual and text reference sources, electronic resources, and/or interviews).
5.6.R.2	Identify and/or use some text features (e.g., graphics, captions, subheadings, italicized words, charts, tables, legends) to comprehend the structure of informational texts.	Identify and use text features (e.g., graphics, captions, headings/subheadings, bold/italicized words, charts, tables, legends) to analyze the structure of informational texts.	Use text features (e.g., graphics, captions, headings/subheadings, bold/italicized words, charts, tables, legends) and explain how they support the structure of informational texts.
5.6.R.3	Determine the relevance of the information gathered.	Determine the relevance and reliability of the information gathered.	Determine and explain the relevance and reliability of the information gathered.
5.6.W.1	Identify a viable research question about a provided topic.	Formulate a viable research question.	Formulate multiple viable research questions.
5.6.W.2	Organize information found during research.	Organize information found during research, following a modified citation style (i.e., author, title, publication date).	

OSTP ELA Grade 6 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:
Reading & Writing Process			
6.2.R.1	Summarize the important events or information in a text.	Summarize alphabetic and/or multimodal texts, including main idea, to demonstrate comprehension.	Summarize complex alphabetic and/or multimodal texts, including main idea, to demonstrate comprehension; evaluate summaries.
6.2.R.2	Identify details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to distinguish genres and provide supporting evidence for analysis.
6.2.R.3	Paraphrase a sentence in their own words to demonstrate comprehension.	Paraphrase a paragraph in their own words to demonstrate comprehension.	
6.2.W.1	Identify a prewriting strategy.	Prewrite (e.g., develop ideas and plan).	Create a prewriting strategy.
6.2.W.2	Develop ideas to compose a first draft.	Organize and develop ideas to compose a first draft.	Organize and develop ideas related to a thesis to compose a first draft.
6.2.W.3	Revise drafts of paragraphs for logical order and effective transitions.	Revise drafts for intended purpose, audience, and organization (e.g., logical order and transitions).	Evaluate and revise drafts for intended purpose, audience, and organization (e.g., logical order and transitions).
6.2.W.4	Edit for correct grammar, usage, and mechanics, using various resources.	Edit for correct grammar, usage, and mechanics, using various resources.	Use various resources to correct grammar, usage, and mechanics for intended purposes.
Critical Reading & Writing			
6.3.R.1	Compare and contrast stated purposes of authors writing on the same topic from a variety of historical, cultural, ethnic, and global perspectives.	Compare and contrast stated or implied purposes of authors writing on the same topic from a variety of historical, cultural, ethnic, and global perspectives.	Compare and contrast stated or implied purposes of authors writing on the same topic from a variety of historical, cultural, ethnic, and global perspectives in complex texts.
6.3.R.2	Identify how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts.	Evaluate how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts.	Evaluate how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts and provide supporting evidence.

Objective	Basic	Proficient	Advanced
6.3.R.3	<p>Identify how literary elements contribute to the meaning of a literary text:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient) 	<p>Analyze how literary elements contribute to the meaning of a literary text:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient) 	<p>Evaluate how literary elements contribute to the meaning of a literary text:</p> <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient)
6.3.R.4	<p>Identify how literary devices contribute to the meaning of a text:</p> <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) 	<p>Analyze how literary devices contribute to the meaning of a text:</p> <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) 	<p>Evaluate how literary devices contribute to the meaning of a text:</p> <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration)
6.3.R.5	Identify literary elements that impact a text's theme.	Identify literary elements and devices that impact a text's theme.	Evaluate literary elements and devices that impact a text's theme.
6.3.R.6	Identify facts included in an argument as for or against an issue.	Categorize facts included in an argument as for or against an issue.	Determine whether facts strengthen or weaken an argument.
6.3.R.7	<p>Determine how informational text structures support the author's purpose:</p> <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	<p>Analyze how informational text structures support the author's purpose:</p> <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	<p>Analyze and explain how informational text structures support the author's purpose:</p> <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential
6.3.R.8	Identify evidence from a text that supports an inference.	Analyze one or more ideas from a text, providing textual evidence to support their inferences.	
6.3.W.1	<p>Compose simple narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include plots involving characters resolving conflicts ● unfold in chronological sequence ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	<p>Compose narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts ● unfold in chronological or surprising sequence (e.g., foreshadowing) ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	<p>Compose complex narratives reflecting real or imagined experiences that:</p> <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts ● unfold in chronological or surprising sequence (e.g., foreshadowing) ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts

Objective	Basic	Proficient	Advanced
6.3.W.2	<p>Compose simple informative essays or reports that:</p> <ul style="list-style-type: none"> ● introduce and develop a topic ● incorporate evidence (e.g., specific facts and details) ● attempt to maintain an organized structure 	<p>Compose informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) <ul style="list-style-type: none"> ● maintain an organized structure ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style ● emulate literary devices from mentor texts 	<p>Compose complex informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) <ul style="list-style-type: none"> ● maintain an organized structure ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style ● emulate literary devices from mentor texts
6.3.W.3	<p>Compose simple argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce a claim ● attempt to organize claims and evidence in a logical sequence ● attempt to provide relevant evidence to develop arguments, using credible sources 	<p>Compose argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce precise claims ● organize claims and evidence in a logical sequence ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style 	<p>Compose complex argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce precise claims ● organize claims and evidence in a logical sequence ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style
Vocabulary			
6.4.R.1	Identify synonyms, antonyms, and analogies.	Analyze the relationships among synonyms, antonyms, and analogies.	Evaluate the relationships among synonyms, antonyms, and analogies for intended effect.
6.4.R.2	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among simple multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among complex multiple-meaning words.
6.4.R.3	Use word parts (e.g., affixes, Latin roots, stems) to define and determine the meaning of simple words.	Use word parts (e.g., affixes, Latin roots, stems) to define and determine the meaning of increasingly complex words.	
6.4.R.4		Use a dictionary, glossary, or thesaurus to determine or clarify the meanings, syllabication, pronunciation, synonyms, antonyms, and parts of speech of words.	
6.4.W.1	Use simple vocabulary in writing to clearly communicate ideas.	Use precise, grade-level vocabulary in writing to clearly communicate ideas.	Use precise, complex vocabulary in writing to clearly communicate ideas.
6.4.W.2	Select language in writing to create a specific given effect according to purpose.	Select language in writing to create a specific effect according to purpose.	Select complex language in writing to create a specific effect according to purpose.
Language			
6.5.R.1		Recognize simple, compound, and complex sentences.	

Objective	Basic	Proficient	Advanced
6.5.R.2	<p>Recognize and explain the impact on meaning of parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns verb tense to signify various times, sequences, and conditions subject and verb agreement adjectives prepositional phrases adverbs interjections 	<p>Recognize and explain the impact on meaning of parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns verb tense to signify various times, sequences, conditions, and states subject and verb agreement adjectives prepositional phrases reflexive pronouns and their antecedents singular they / them / their subordinating conjunctions adverbs interjections 	
6.5.W.1	Compose simple, compound, and complex sentences.	Compose simple, compound, and complex sentences to add clarity and variety in their writing.	
6.5.W.2	Use nouns, verbs, adjectives, prepositions, adverbs, and pronouns.	Add clarity and variety to their writing with nouns, verbs, adjectives, prepositions, adverbs, and pronouns.	
6.5.W.3	Recognize the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and inappropriate shifts in pronoun number and person.	Recognize and correct the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and inappropriate shifts in pronoun number and person.	Evaluate for and correct the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and inappropriate shifts in pronoun number and person.
6.5.W.7	Identify sentences that correctly use commas to separate an introductory element from the rest of the sentence and to indicate direct address (e.g., Where are you, Sam?).	Use commas to separate an introductory element from the rest of the sentence and to indicate direct address (e.g., Where are you, Sam?).	Edit for commas to separate an introductory element from the rest of the sentence and to indicate direct address (e.g., Where are you, Sam?).
6.5.W.8	Identify sentences that correctly use a colon to introduce a quotation from a source (e.g., According to National Geographic, meerkat homes are quite comfortable: "Each burrow is an extensive tunnel-and-room system that remains cool even under the broiling African sun.").	Use a colon to introduce a quotation from a source (e.g., According to National Geographic, meerkat homes are quite comfortable: "Each burrow is an extensive tunnel-and-room system that remains cool even under the broiling African sun.").	Edit for a colon to introduce a quotation from a source (e.g., According to National Geographic, meerkat homes are quite comfortable: "Each burrow is an extensive tunnel-and-room system that remains cool even under the broiling African sun.").
6.5.W.9	Identify sentences that use quotation marks to indicate dialogue, quoted material, and titles of works.	Use quotation marks to indicate dialogue, quoted material, and titles of works.	Edit for quotation marks to indicate dialogue, quoted material, and titles of works.
6.5.W.10	Identify sentences that use underlining or italics to indicate titles of works.	Use underlining or italics to indicate titles of works.	Edit for underlining or italics to indicate titles of works.
6.5.W.11	Identify sentences that use a semicolon to punctuate compound sentences.	Use a semicolon to punctuate compound sentences.	Edit for a semicolon to punctuate compound sentences.
Research			
6.6.R.1	Identify viable research questions to gather information about a topic.	Use their own viable research questions to gather information about a topic.	

Objective	Basic	Proficient	Advanced
6.6.R.2	Record information from various primary and secondary sources.	Record and organize information from various primary and secondary sources.	Record, organize, and analyze information from various primary and secondary sources.
6.6.R.3	Identify the relevance and reliability of the information gathered.	Determine the relevance and reliability of the information gathered.	Evaluate the relevance and reliability of the information gathered.
6.6.W.1	Identify a viable research question.	Formulate and refine a viable research question.	
6.6.W.2	Identify a clear, concise thesis statement.	Develop a clear, concise thesis statement.	Revise a thesis statement to be clear and concise.
6.6.W.3	Quote findings.	Quote findings following a consistent citation style (e.g., MLA, APA) to avoid plagiarism.	

OSTP ELA Grade 7 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:
Reading & Writing Process			
7.2.R.1	Summarize alphabetic and/or multimodal text, including main idea and key details, to demonstrate comprehension of a text.	Summarize alphabetic and/or multimodal texts, including main idea and key details, to demonstrate comprehension within and between texts.	Summarize alphabetic and/or multimodal texts, including main idea and key details, to demonstrate comprehension between texts; evaluate summaries.
7.2.R.2	Identify details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to distinguish genres and provide supporting evidence.
7.2.R.3	Paraphrase a paragraph in their own words to demonstrate comprehension.	Paraphrase a short passage in their own words to demonstrate comprehension.	
7.2.W.1	Identify a prewriting strategy (e.g., develop ideas and plan).	Prewrite (e.g., develop ideas and plan).	Create prewriting strategy.
7.2.W.2	Minimally plan/organize ideas.	Organize and develop ideas to compose a first draft.	Organize and develop ideas related to a thesis to compose a first draft.
7.2.W.3	Revise provided drafts of paragraphs for intended purpose, audience, organization, and coherence (e.g., consistent point of view).	Revise drafts for intended purpose, audience, organization, and coherence (e.g., consistent point of view).	Revise self-created drafts for intended purpose, audience, organization, and coherence (e.g., consistent point of view) and style.
7.2.W.4	Edit for correct grammar, usage, and mechanics, using various resources.	Edit for correct grammar, usage, and mechanics, using various resources.	Use various resources to correct grammar, usage, and mechanics for intended purposes.
Critical Reading & Writing			
7.3.R.1	Read works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and identify the methods the authors use to achieve their purposes.	Read works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and compare the methods the authors use to achieve their purposes.	Read works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and analyze the methods the authors use to achieve their purposes.
7.3.R.2	Identify how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts.	Evaluate how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts.	Evaluate how perspective (e.g., historical, cultural, ethnic, and global) affects a variety of literary and informational texts and provide supporting evidence.

Objective	Basic	Proficient	Advanced
7.3.R.3	Identify literary elements to support an interpretation of a text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient and second person) 	Analyze literary elements to support an interpretation of a text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient and second person) 	Evaluate literary elements to support an interpretation of a text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient and second person)
7.3.R.4	Identify literary devices to support an interpretation of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal irony 	Analyze literary devices to support an interpretation of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal irony 	Evaluate literary devices to support an interpretation of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal irony
7.3.R.5	Identify literary elements and devices that impact a text's theme.	Identify literary elements and devices that impact a text's theme and mood.	Explain how literary elements and devices impact a text's theme and mood.
7.3.R.6	Identify factual claims in a text.	Distinguish factual claims from opinions.	Evaluate factual claims.
7.3.R.7	Determine how informational text structures support the author's purpose: <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	Analyze how informational text structures support the author's purpose: <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	Analyze and explain how informational text structures support the author's purpose: <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential
7.3.R.8	Identify multiple ideas from a text that support an inference.	Analyze multiple ideas from a text, providing textual evidence to support their inferences.	
7.3.W.1	Compose simple narratives reflecting real or imagined experiences that may: <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts <ul style="list-style-type: none"> ● unfold in chronological or surprising sequence (e.g., foreshadowing) ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	Compose narratives reflecting real or imagined experiences that <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts <ul style="list-style-type: none"> ● unfold in chronological or surprising sequence (e.g., foreshadowing) ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	Compose complex narratives reflecting real or imagined experiences that: <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts <ul style="list-style-type: none"> ● unfold in chronological or surprising sequence (e.g., foreshadowing) ● include a narrator, precise language, sensory details, dialogue, and thoughts to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts

Objective	Basic	Proficient	Advanced
7.3.W.2	<p>Compose simple informative essays or reports that:</p> <ul style="list-style-type: none"> ● introduce and develop topics ● incorporate evidence (e.g., specific facts and details) ● attempt to maintain an organized structure 	<p>Compose informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) <ul style="list-style-type: none"> ● maintain an organized structure ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style ● emulate literary devices from mentor texts 	<p>Compose complex informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) <ul style="list-style-type: none"> ● maintain an organized structure ● use sentence variety and word choice to create clarity <ul style="list-style-type: none"> ● establish and maintain a formal style ● emulate literary devices from mentor texts
7.3.W.3	<p>Compose simple argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce a claim ● attempt to organize the claim and evidence in a logical sequence ● provide evidence to develop arguments, using credible sources 	<p>Compose argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce precise claims ● organize claims and evidence in a logical sequence ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and word choice to create clarity ● establish and maintain a formal style 	<p>Compose complex argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce precise claims ● organize claims and evidence in a logical sequence ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and word choice to create clarity ● establish and maintain a formal style
Vocabulary			
7.4.R.1	Identify synonyms, antonyms, and analogies.	Analyze the relationships among synonyms, antonyms, and analogies.	Evaluate the relationships among synonyms, antonyms, and analogies for intended effect.
7.4.R.2	Use context clues and denotation to determine or clarify the meaning of words or distinguish among simple multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among complex multiple-meaning words.
7.4.R.3	Use word parts (e.g., affixes, Greek roots, stems) to define and determine the meaning of new words.	Use word parts (e.g., affixes, Greek roots, stems) to define and determine the meaning of increasingly complex words.	
7.4.R.4		Use a dictionary, glossary, or thesaurus to determine or clarify the meanings, syllabication, pronunciation, synonyms, antonyms, and parts of speech of words.	
7.4.W.1	Use simple vocabulary in writing to clearly communicate ideas.	Use precise, grade-level vocabulary in writing to clearly communicate ideas.	Use precise, complex vocabulary in writing to clearly communicate ideas.
7.4.W.2	Select simple language in writing to create a specific, given effect according to purpose.	Select language in writing to create a specific effect according to purpose.	Select complex language in writing to create a specific effect according to purpose.
Language			
7.5.R.1	Recognize simple, compound, and complex sentences.	Recognize simple, compound, complex, and compound-complex sentences and explain their effects.	Analyze simple, compound, complex, and compound-complex sentences and explain their effects.

Objective	Basic	Proficient	Advanced
7.5.R.2	<p>Recognize and explain the impact on meaning of parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns subject and verb agreement singular they / them / their <ul style="list-style-type: none"> adverbs interjections 	<p>Recognize and explain the impact on meaning of parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns gerunds subject and verb agreement cumulative and coordinate adjectives demonstrative pronouns vague pronouns (i.e., ones with unclear or ambiguous antecedents) <ul style="list-style-type: none"> singular they / them / their correlative conjunctions <ul style="list-style-type: none"> adverbs interjections 	
7.5.W.1	Compose simple, compound, complex sentences.	Compose simple, compound, complex, and compound-complex sentences to add clarity and variety in their writing.	Compose compound-complex sentences to add clarity, variety, and intended effect in their writing.
7.5.W.2	Use nouns, verbs, adjectives, prepositions, adverbs, and pronouns.	Add clarity and variety to their writing with nouns, verbs, adjectives, prepositions, adverbs, and pronouns.	
7.5.W.3	Identify the following: run-ons, errors in subject and verb agreement, and inappropriate shifts in verb tense.	Recognize and correct the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and vague pronouns (i.e., ones with unclear or ambiguous antecedents).	Evaluate for and correct the following: run-ons, errors in subject and verb agreement, inappropriate shifts in verb tense, and vague pronouns (i.e., ones with unclear or ambiguous antecedents).
7.5.W.7	Identify sentences that correctly use commas to separate words or phrases in a series.	Use commas to separate words or phrases in a series.	Edit for commas used to separate words or phrases in a series.
7.5.W.8	Identify sentences that correctly use a colon to introduce a quotation from a source.	Use a colon to introduce a quotation from a source.	Edit for colons used to introduce a quotation from a source.
7.5.W.9	Identify sentences that correctly use quotation marks to indicate dialogue, quoted material, and titles of works.	Use quotation marks to indicate dialogue, quoted material, and titles of works.	Edit for quotation marks used to indicate dialogue, quoted material, and titles of works.
7.5.W.10	Identify sentences that correctly use underlining or italics to indicate titles of works, thoughts in narratives, and words in a foreign language.	Use underlining or italics to indicate titles of works, thoughts in narratives, and words in a foreign language.	Edit for use of underlining or italics to indicate titles of works, thoughts in narratives, and words in a foreign language.
7.5.W.11	Identify sentences that correctly use a semicolon to punctuate compound and compound-complex sentences.	Use a semicolon to punctuate compound and compound-complex sentences.	Edit for a semicolon to punctuate compound and compound-complex sentences.
Research			
7.6.R.1	Find and comprehend information (e.g., claims, evidence) about a topic and identify a viable research question.	Find and comprehend information (e.g., claims, evidence) about a topic, using their own viable research questions.	

Objective	Basic	Proficient	Advanced
7.6.R.2	Record and organize information from a variety of primary and secondary sources.	Find, record, and organize information from a variety of primary and secondary sources, following ethical and legal guidelines.	Find, analyze, and record and organize information from a variety of primary and secondary sources, following ethical and legal guidelines.
7.6.R.3	Determine the relevance and reliability of the information gathered.	Determine the relevance, reliability, and validity of the information gathered.	Evaluate the relevance, reliability, and validity of the information gathered.
7.6.W.1	Identify a clear and concise research question.	Formulate and refine a viable research question.	
7.6.W.2	Identify a clear, concise thesis statement.	Develop a clear, concise thesis statement.	Revise a thesis statement to be clear and concise.
7.6.W.3	Quote and summarize findings.	Quote and summarize findings following a consistent citation style (e.g., MLA, APA) to avoid plagiarism.	

OSTP ELA Grade 8 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:
Reading & Writing Process			
8.2.R.1	Summarize an alphabetic or multimodal text to demonstrate comprehension of a text.	Summarize alphabetic and/or multimodal texts about similar topics to demonstrate comprehension within and between texts.	Summarize alphabetic and/or multimodal texts about similar topics to demonstrate comprehension within and between texts; evaluate summaries.
8.2.R.2	Identify details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to identify characteristics of genres.	Analyze details in fiction, poetry, and nonfiction texts to identify characteristics of genres and provide supporting evidence.
8.2.R.3	Paraphrase a paragraph in their own words to demonstrate comprehension.	Paraphrase a portion of passage in their own words to demonstrate comprehension.	
8.2.W.1	Identify a prewriting strategy (e.g., develop ideas and plan).	Prewrite (e.g., develop ideas and plan).	Create and use a prewriting strategy.
8.2.W.2	Minimally plan/organize ideas.	Organize and develop ideas to compose a first draft.	Organize and develop ideas related to a thesis to compose a first draft.
8.2.W.3	Revise provided drafts of paragraphs for intended purpose, audience, and organization.	Revise drafts for intended purpose, audience, organization, coherence, and style (e.g., word choice and sentence variety).	Revise self-created drafts for intended purpose, audience, organization, coherence, and style (e.g., word choice and sentence variety).
8.2.W.4	Edit a paragraph for correct grammar and mechanics, using various resources.	Edit for correct grammar, usage, and mechanics, using various resources.	Edit for correct grammar, usage, and mechanics, using various resources; edit mechanics for intended effect and purpose.
Critical Reading & Writing			
8.3.R.1	Analyze works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and compare the methods the authors use to achieve their purposes.	Analyze works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and analyze the methods the authors use to achieve their purposes.	Analyze works written on the same topic from a variety of historical, cultural, ethnic, and global perspectives and evaluate the methods the authors use to achieve their purposes.
8.3.R.2	Determine perspectives (e.g., historical, cultural, ethnic, and global) and describe how they affect various literary and informational texts.	Evaluate perspectives (e.g., historical, cultural, ethnic, and global) and describe how they affect various literary and informational texts.	

Objective	Basic	Proficient	Advanced
8.3.R.3	Identify literary elements to support interpretations of a literary text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient, second person, and unreliable narrator) 	Analyze literary elements to support interpretations of a literary text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient, second person, and unreliable narrator) 	Evaluate literary elements to support interpretations of a literary text: <ul style="list-style-type: none"> ● setting ● plot ● characters (i.e., protagonist, antagonist) ● characterization ● conflict (i.e., internal, external) ● point of view (i.e., third person limited and omniscient, second person, and unreliable narrator)
8.3.R.4	Determine literary devices to support interpretations of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal and situational irony 	Analyze literary devices to support interpretations of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal and situational irony 	Evaluate literary devices to support interpretations of a text: <ul style="list-style-type: none"> ● figurative language (i.e., simile, metaphor, personification, hyperbole, imagery, symbolism, idiom) ● sound devices (i.e., onomatopoeia, alliteration) ● verbal and situational irony
8.3.R.5	Identify literary elements and devices that impact a text's theme and mood.	Identify literary elements and devices that impact a text's theme, mood, and tone.	Identify literary elements and devices that impact a text's theme, mood, and tone.
8.3.R.6	Identify a claim and describe how evidence supports a claim.	Evaluate textual evidence to determine whether a claim is substantiated or unsubstantiated.	
8.3.R.7	Determine how informational text structures support the author's purpose: <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	Analyze how informational text structures support the author's purpose: <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential 	Analyze and evaluate how informational text structures support the author's purpose and explain why one structure was selected over another. <ul style="list-style-type: none"> ● compare/contrast ● cause/effect ● problem/solution ● description ● sequential
8.3.R.8	Compare or contrast ideas within a text, providing textual evidence to support their inferences.	Compare or contrast two or more texts, providing textual evidence to support their inferences.	Analyze two or more texts, providing textual evidence to support their inferences.
8.3.W.1	Compose simple narratives reflecting real or imagined experiences that may: <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts ● unfold in chronological or surprising sequence (e.g., flashback and foreshadowing) ● include a narrator, precise language, sensory details, and dialogue to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	Compose narratives reflecting real or imagined experiences that <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts ● unfold in chronological or surprising sequence (e.g., flashback and foreshadowing) ● include a narrator, precise language, sensory details, and dialogue to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts 	Compose complex narratives reflecting real or imagined experiences that: <ul style="list-style-type: none"> ● include plots involving complex characters resolving conflicts ● unfold in chronological or surprising sequence (e.g., flashback and foreshadowing) ● include a narrator, precise language, sensory details, and dialogue to enhance the narrative <ul style="list-style-type: none"> ● use sentence variety to create clarity ● emulate literary elements and/or literary devices from mentor texts

Objective	Basic	Proficient	Advanced
8.3.W.2	<p>Compose simple informative essays or reports that:</p> <ul style="list-style-type: none"> ● introduce and develop topics ● incorporate evidence (e.g., specific facts, details) ● attempt to maintain an organized structure ● attempt to use sentence variety and word choice to create clarity 	<p>Compose informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) ● maintain an organized structure ● use sentence variety and word choice to create clarity ● establish and maintain a formal style ● emulate literary devices from mentor texts 	<p>Compose complex informative essays or reports that:</p> <ul style="list-style-type: none"> ● objectively introduce and develop topics ● incorporate evidence (e.g., specific facts, details, charts and graphs, data) ● maintain a clear and organized structure using smooth transitions ● use sentence variety and word choice to create clarity ● establish and maintain a formal style ● emulate literary devices from mentor texts
8.3.W.3	<p>Compose simple argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce claims ● attempt to organize claims and evidence in a logical sequence ● provide evidence to develop arguments, using credible sources ● attempt to use sentence variety and word choice to create clarity 	<p>Compose argumentative essays that:</p> <ul style="list-style-type: none"> ● introduce precise claims ● acknowledge counterclaims ● organize claims, counterclaims, and evidence in a logical sequence ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and word choice to create clarity ● establish and maintain a formal style 	<p>Compose complex argumentative essays that:</p> <ul style="list-style-type: none"> ● clearly introduce precise claims ● acknowledge counterclaims ● effectively organize claims, counterclaims, and evidence in a logical sequence using smooth transitions ● provide relevant evidence to develop arguments, using credible sources ● use sentence variety and precise word choice to create clarity ● establish and maintain a formal style
Vocabulary			
8.4.R.1	Identify synonyms, antonyms, and analogies.	Analyze the relationships among synonyms, antonyms, and analogies.	Evaluate the relationships among synonyms, antonyms, and analogies.
8.4.R.2	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among simple multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among multiple-meaning words.	Use context clues, connotation, and denotation to determine or clarify the meaning of words or distinguish among complex multiple-meaning words.
8.4.R.3	Use word parts (e.g., affixes, Greek roots, stems) to define and determine the meaning of simple words.	Use word parts (e.g., affixes, Greek roots, stems) to define and determine the meaning of increasingly complex words.	
8.4.R.4		Use a dictionary, glossary, or thesaurus to determine or clarify the meanings, syllabication, pronunciation, synonyms, antonyms, and parts of speech of words.	
8.4.W.1	Use precise, simple vocabulary in writing to clearly communicate ideas.	Use precise, grade-level vocabulary in writing to clearly communicate ideas.	Use precise, complex vocabulary in writing to clearly communicate ideas.
8.4.W.2	Select language in writing to create a given effect according to purpose.	Select language in writing to create a specific effect according to purpose.	Select complex language in writing to create a specific effect according to purpose.
Language			
8.5.R.1		Recognize active and passive voice and misplaced and dangling modifiers in sentences.	

Objective	Basic	Proficient	Advanced
8.5.R.2	<p>Recognize parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns verbals (i.e., gerunds, participles, infinitives) <ul style="list-style-type: none"> cumulative and coordinate adjectives vague pronouns singular they/them/their coordinating, subordinating, and correlative conjunctions <ul style="list-style-type: none"> adverbs interjections 	<p>Recognize and explain the impact on meaning of parts of speech in sentences:</p> <ul style="list-style-type: none"> nouns verbals (i.e., gerunds, participles, infinitives) <ul style="list-style-type: none"> cumulative and coordinate adjectives vague pronouns singular they/them/their coordinating, subordinating, and correlative conjunctions <ul style="list-style-type: none"> adverbs interjections 	
8.5.W.1	Compose simple, compound, complex, and compound-complex sentences.	Compose simple, compound, complex, and compound-complex sentences to add clarity and variety to their writing.	Compose simple, compound, complex, and compound-complex sentences to add clarity, variety, and contribute to the intended purpose of their writing.
8.5.W.2	Use nouns, verbs, verbals, adjectives, prepositions, adverbs, pronouns, and conjunctions in their writing.	Create clarity and/or add variety to their writing with nouns, verbs, verbals, adjectives, prepositions, adverbs, pronouns, and conjunctions.	Create clarity and add variety to their writing with nouns, verbs, verbals, adjectives, prepositions, adverbs, pronouns, and conjunctions.
8.5.W.3	Recognize and correct vague pronouns.	Recognize and correct the following: misplaced and dangling modifiers, vague pronouns, and second person point of view in formal writing.	Evaluate for and correct the following: misplaced and dangling modifiers, vague pronouns, and second person point of view in formal writing.
8.5.W.7		Use commas to separate coordinate adjectives (e.g., a fascinating, enjoyable movie).	Evaluate for and use commas to separate coordinate adjectives (e.g., a fascinating, enjoyable movie).
8.5.W.8		Use a colon to introduce a quotation from a source.	Edit for colons used to introduce a quotation from a source.
8.5.W.10	Use underlining or italics to indicate titles of works.	Use underlining or italics to indicate titles of works, thoughts in narratives, and words in a foreign language.	Edit for underlining or italics to indicate titles of works, thoughts in narratives, and words in a foreign language.
8.5.W.11	Use a semicolon to punctuate compound sentences.	Use a semicolon to punctuate compound and compound-complex sentences.	Edit for a semicolon to punctuate compound and compound-complex sentences.
Research			
8.6.R.1	Find and comprehend information (e.g., claims, evidence) about a topic and identify viable research questions.	Find and comprehend information (e.g., claims, evidence) about a topic, using their own viable research questions.	
8.6.R.2	Find and organize information from a variety of primary and secondary sources.	Find, record, and organize information from a variety of primary and secondary sources, following ethical and legal guidelines.	Find, analyze, record, and organize information from a variety of primary and secondary sources, following ethical and legal guidelines.

Objective	Basic	Proficient	Advanced
8.6.R.3	Identify the relevance, reliability, and validity of the information gathered.	Determine the relevance, reliability, and validity of the information gathered.	Evaluate information for relevance, reliability, and validity.
8.6.W.1	Identify a viable research question.	Formulate and refine a viable research question.	
8.6.W.2	Identify a clear, concise thesis statement.	Develop a clear, concise, defensible thesis statement.	Revise a defensible thesis statement based on findings for clarity and concision.
8.6.W.3	Quote and summarize findings.	Quote, paraphrase, and summarize findings following a consistent citation style (e.g., MLA, APA) to avoid plagiarism.	

OSTP Math Grade 3 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically.	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically.	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically.	
Numbers & Operations	Represent and describe whole numbers up to 100,000.	Compare and order whole numbers.	Compare and order whole numbers when numbers are given in different forms.	3.N.1.1, 3.N.1.2, 3.N.1.4
	Solve addition and subtraction problems.	Solve multiplication problems. Recognize the relationship between multiplication and division.	Assess the reasonableness of results in addition and subtraction problems.	3.N.1.3, 3.N.2.3, 3.N.2.5, 3.N.2.7, 3.N.2.8
		Round numbers to the nearest thousand, ten thousand, and hundred thousand.	Use rounding to estimate sums and differences.	3.N.1.5, 3.N.2.4
		Represent multiplication and division facts by modeling a variety of approaches.		3.N.2.1, 3.N.2.6
		Demonstrate fluency with multiplication facts.		3.N.2.2
	Read and write fractions. Apply understanding of unit fractions. Represent fractions with models.	Compose and decompose fractions.	Compare and order fractions using models.	3.N.3.1, 3.N.3.2, 3.N.3.3, 3.N.3.4
	Determine the value of a set of coins or a set of bills.			3.N.4.1, 3.N.4.2
Algebraic Reasoning & Algebra	Describe patterns.	Describe the rule for a pattern.	Create and extend patterns.	3.A.1.1, 3.A.1.2, 3.A.1.3

Strand	Basic	Proficient	Advanced	Objective(s)
Algebraic Reasoning & Algebra		Determine unknowns (represented by symbols) in one-step addition, subtraction, and multiplication equations.	Generate real-world situations to represent number sentences.	3.A.2.1
	Identify commutative, identity, and associative properties.	Apply commutative, identity, and associative properties.		3.A.2.2
Geometry & Measurement	Sort three-dimensional figures based on attributes.	Build a three-dimensional figure using unit cubes.	Count cubes to find the number of cubes needed to pack the whole or half of a structure.	3.GM.1.1, 3.GM.1.2, 3.GM.2.3
	Identify right angles.	Classify angles.		3.GM.1.3
		Determine the perimeter of polygons.		3.GM.2.1
		Determine the area of two-dimensional figures.	Analyze why length and width are multiplied to find the area of a rectangle.	3.GM.2.2, 3.GM.2.4
	Choose an appropriate instrument to measure the length of an object.	Measure length.		3.GM.2.5, 3.GM.2.6
		Use an analog thermometer to determine temperature.		3.GM.2.7
	Read and write time from a digital clock.	Read and write time from an analog clock.	Determine elapsed time.	3.GM.3.1, 3.GM.3.2
Data & Probability	Collect data.	Organize a data set using a frequency table, line plot, pictograph, or bar graph with intervals of one.	Organize a data set using a frequency table, line plot, pictograph, or bar graph with intervals other than one.	3.D.1.1
		Solve one-step problems represented with a frequency table, pictograph, or bar graph with scaled intervals.	Solve two-step problems represented with a frequency table, pictograph, or bar graph with scaled intervals.	3.D.1.2

OSTP Math Grade 4 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations	Represent and describe whole numbers up to 1,000,000.	Use place value to compare and order whole numbers.		4.N.1.1, 4.N.1.2, 4.N.1.4
		Apply knowledge of place value to multiply a number by 10, 100, and 1,000.		4.N.1.3
	Demonstrate fluency with multiplication and division facts.	Multiply and estimate 3-digit by 1-digit and 2-digit by 2-digit whole numbers.	Assess the reasonableness of the estimation of 3-digit by 1-digit and 2-digit by 2-digit whole-number products.	4.N.2.1, 4.N.2.2, 4.N.2.3,
		Solve multi-step problems.	Apply and analyze models to solve multi-step problems and assess the reasonableness of results.	4.N.2.4
		Divide a 3-digit dividend by a 1-digit divisor with and without remainder.		4.N.2.5
	Use models to determine equivalent fractions.			4.N.3.1
		Use benchmark fractions to locate additional fractions on a number line.		4.N.3.2
	Use models to compare and order fractions with like denominators.	Use models to compare and order fractions with unlike denominators.		4.N.3.3
	Use models to add and subtract fractions.	Decompose fractions.		4.N.3.4, 4.N.3.5

Strand	Basic	Proficient	Advanced	Objective(s)
Numbers & Operations	Represent tenths and hundredths with models.	Make connections between fractions (tenths and hundredths) and decimals with models.		4.N.3.1, 4.N.3.6
	Read and write decimals up to the hundredths place, including money.	Compare and order benchmark fractions. Compare and order decimals.	Compare and order benchmark fractions to decimals.	4.N.3.7, 4.N.3.8, 4.N.3.9
		Select the fewest number of coins for a given amount of money.		4.N.4.1
	Determine change using whole dollars.	Determine change using coins and dollars.		4.N.4.2
Algebraic Reasoning & Algebra	Create an input/output table.	Determine rules and extend patterns shown in input/output tables.		4.A.1.1, 4.A.1.2
		Define the single operation rule of a pattern involving geometric shapes.	Construct models to show growth patterns involving geometric shapes.	4.A.1.3
	Use the relationships between multiplication and division with the properties of multiplication to solve problems.	Solve for a variable in an equation with addition, subtraction, multiplication, and division of whole numbers.	Analyze models to represent number sentences.	4.A.2.1, 4.A.2.2
		Determine unknown values in equivalent expressions.	Determine unknown values in non-equivalent expressions.	4.A.2.3
Geometry & Measurement	Identify points, endpoints, and angles.	Identify lines, line segments, rays, and parallel and perpendicular lines.		4.GM.1.1
	Describe and recognize quadrilaterals.	Classify quadrilaterals.	Construct quadrilaterals.	4.GM.1.2
	Identify three-dimensional figures.	Compare and contrast the similarities and differences of three-dimensional figures based on their attributes.		4.GM.1.3

Strand	Basic	Proficient	Advanced	Objective(s)
Geometry & Measurement		Measure angles.		4.GM.2.1
		Decompose and determine the area of polygons.		4.GM.2.2
		Develop the concept of volume.	Create models to determine volume.	4.GM.2.3
	Identify appropriate units and tools to measure length. Measure the lengths of objects.	Compare the lengths of objects.	Determine and justify the best use of customary and metric measurements in a variety of situations.	4.GM.2.4, 4.GM.2.5, 4.GM.2.6, 4.GM.2.7
		Convert measurements of time.	Determine elapsed time.	4.GM.3.1, 4.GM.3.2
Data & Probability		Create a frequency table or line plot with whole numbers. Organize data sets to create tables, bar graphs, timelines, and Venn diagrams with whole numbers.	Create a frequency table or line plot with fractions. Organize data sets to create tables, bar graphs, timelines, and Venn diagrams with fractions.	4.D.1.1, 4.D.1.2
		Solve one-step problems by analyzing data in whole-number, decimal, or fraction form in a frequency table and line plot.	Solve two-step problems by analyzing data in whole-number, decimal, or fraction form in a frequency table and line plot.	4.D.1.3

OSTP Math Grade 5 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations	Represent decimal fractions with a model.			5.N.1.1
	Recognize and generate equivalent decimals, fractions, and mixed numbers and represent whole numbers.	Compare and order fractions. Compare and order decimals.	Order a mix of decimals, fractions, mixed numbers, and whole numbers.	5.N.1.2, 5.N.1.3, 5.N.1.4
	Solve division, multiplication, addition, and subtraction problems.	Estimate and solve division problems with the remainder represented as a fraction, decimal, or whole number.	Interpret the remainder of division problems within the context of the problem.	5.N.2.1, 5.N.2.2, 5.N.2.3, 5.N.2.4
	Add and subtract decimals and fractions with like denominators.	Estimate, illustrate, add, and subtract fractions and mixed numbers.	Order a mix of decimals, fractions, mixed numbers, and whole numbers.	5.N.3.1, 5.N.3.2, 5.N.3.3, 5.N.3.4
Algebraic Reasoning & Algebra	Describe patterns of change. Identify the origin and axes in relation to the coordinates.	Graph patterns of change as ordered pairs on a coordinate plane. Use a rule or table to represent ordered pairs.	Make predictions and generalizations about patterns of change.	5.A.1.1, 5.A.1.2
	Generate equivalent numerical expressions.	Evaluate numerical expressions.	Apply the order of operations, commutative property, associative property, and distributive property.	5.A.2.1, 5.A.2.3
	Determine whether an equation involving a variable is true or false for a given value of the variable.	Determine whether an inequality involving a variable is true or false for a given value of the variable.		5.A.2.2

Strand	Basic	Proficient	Advanced	Objective(s)
Geometry & Measurement	Describe and identify triangles.	Classify triangles by their attributes.	Construct triangles.	5.GM.1.1
	Describe, identify, and classify three-dimensional figures when given an image.	Using attributes, describe, identify, and classify three-dimensional figures without a given image.		5.GM.1.2
	Recognize nets for three-dimensional figures.	Construct nets for three-dimensional figures.		5.GM.1.3
		Determine volume of rectangular prisms.	Compare volumes of rectangular prisms.	5.GM.2.1
		Estimate perimeter of polygons and shapes that may include curves.	Justify perimeter of shapes that may include curves.	5.GM.2.2
	Measure angles.	Compare angles.		5.GM.3.1
	Choose an appropriate instrument to measure lengths. Measure the lengths of objects.	Apply the relationship between units to convert and compare objects to solve problems.		5.GM.3.2, 5.GM.3.3, 5.GM.3.4
		Estimate lengths and geometric measurements.		5.GM.3.5
Data & Probability		Calculate the mean, median, mode, and range of a data set.		5.D.1.1
		Create and analyze line and double bar graphs with whole numbers.	Create and analyze line and double bar graphs with fractions or decimals.	5.D.1.2

OSTP Math Grade 6 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations	Represent reflective relationships between integers and their opposites. Explain the meaning of zero.			6.N.1.1
	Read and represent integers or other positive rational numbers.	Order and compare integers or other positive rational numbers.	Explain integers or other positive rational numbers.	6.N.1.2, 6.N.1.3
	Explain that a percent represents parts “out of 100” and ratios “to 100.”	Find equivalent fractions, mixed numbers, decimals, and percents.		6.N.1.3, 6.N.1.4
	Illustrate and compute the addition and subtraction of integers.	Estimate addition and subtraction of integers.	Assess the reasonableness of an answer to addition and subtraction of integers.	6.N.2.1, 6.N.2.2, 6.N.2.3
	Evaluate powers with whole-number bases and exponents.	Identify and represent patterns with whole-number exponents and perfect squares.		6.N.2.4
	Factor whole numbers.	Write positive integers as products of prime factors. Determine greatest common factor and least common multiple.	Use greatest common factor and least common multiple to calculate with fractions, find equivalent fractions, and express the sum of two-digit numbers with a common factor using the distributive property.	6.N.2.5, 6.N.2.6

Strand	Basic	Proficient	Advanced	Objective(s)
Numbers & Operations	Identify ratios.	Use ratios to compare and relate quantities. Determine unit rates. Recognize that multiplicative comparison and additive comparison are different.	Apply the relationship between ratios, equivalent fractions, unit rates, and percents to solve problems in various contexts.	6.N.3.1, 6.N.3.2, 6.N.3.3
	Solve problems involving multiplication and division of fractions and decimals.	Illustrate multiplication and division of fractions and decimals. Estimate solutions involving multiplication and division of fractions and decimals.	Use estimates to assess the reasonableness of solutions involving multiplication and division of fractions and decimals in the context of the problem.	6.N.4.1, 6.N.4.2, 6.N.4.3
		Use modeling to interpret problems including money, measurement, geometry, and data.		6.N.4.4
Algebraic Reasoning & Algebra	Graph ordered pairs in all quadrants.	Represent relationships between varying positive quantities with rules, graphs, and tables.		6.A.1.1, 6.A.1.2
	Evaluate the value of a variable in expressions, equations, and inequalities.	Model or generate expressions, equations, and inequalities.		6.A.1.3, 6.A.2.1, 6.A.3.1
	Use number sense and properties of operations to solve and graph one-step equations on a number line.	Interpret the solution of a one-step equation.	Assess the reasonableness of the solution of a one-step equation.	6.A.3.2
Geometry & Measurement	Identify and display the effect of transformations.	Describe, apply, and predict transformations and use transformations to show congruence.		6.GM.1.1, 6.GM.1.2
	Identify lines of symmetry.	Describe lines of symmetry.		6.GM.1.3

Strand	Basic	Proficient	Advanced	Objective(s)
Geometry & Measurement	Determine the area of parallelograms, squares, and triangles.	Determine the area of polygons that can be decomposed into triangles and rectangles.	Develop the formulas for the area of parallelograms, squares, and triangles.	6.GM.2.1, 6.GM.2.2, 6.GM.2.3
	Identify angle relationships by name.	Use relationships between angles and the triangle sum theorem to solve problems.		6.GM.3.1, 6.GM.3.2
		Estimate weights and capacities. Estimate and solve problems requiring conversion of lengths.		6.GM.4.1, 6.GM.4.2
Data & Probability		Interpret the mean, median, and mode for a set of data.	Justify which measure of center would provide the most descriptive information for a set of data.	6.D.1.1, 6.D.1.2
	Represent possible outcomes using a probability continuum. Determine the sample space of simple experiments and identify possible outcomes.	Compare possible outcomes of simple experiments.	Analyze the differences between two outcomes of simple experiments.	6.D.2.1, 6.D.2.2, 6.D.2.3

OSTP Math Grade 7 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations		Compare and order rational numbers.		7.N.1.1
	Recognize equivalent representations of rational numbers.	Generate equivalent representations of rational numbers.		7.N.1.2
	Calculate the absolute value of a rational number.	Explain the absolute value of a rational number as the distance of that number from zero on a number line.	Apply the concept of absolute value to model and solve problems.	7.N.1.3
		Estimate solutions of problems involving rational numbers.	Assess the reasonableness of the solutions of problems with rational numbers.	7.N.2.1
	Multiply and divide integers.	Illustrate multiplication and division of integers using a variety of representations.		7.N.2.2, 7.N.2.3
	Solve problems involving rational numbers and exponents.	Model problems involving rational numbers and exponents.		7.N.2.4, 7.N.2.5
Algebraic Reasoning & Algebra	Identify a proportional relationship.	Identify the constant of proportionality from a graph.		7.A.1.1, 7.A.1.2
		Represent proportional relationships in a variety of ways and determine unit rates.	Translate from one representation of a proportional relationship to another.	7.A.2.1

Strand	Basic	Proficient	Advanced	Objective(s)
Algebraic Reasoning & Algebra		Solve problems involving proportional relationships.	Assess the reasonableness of solutions of problems involving proportional relationships.	7.A.2.2, 7.A.2.3, 7.A.2.4
	Solve equations.	Write equations.	Interpret equations and inequalities involving variables and rational numbers.	7.A.3.1
	Solve and graph inequalities.	Write inequalities.		7.A.3.2
	Evaluate expressions using the order of operations.	Generate and evaluate equivalent expressions.	Justify the steps when evaluating expressions.	7.A.4.1, 7.A.4.2
Geometry & Measurement	Develop the concepts of surface area and volume of rectangular prisms.	Develop the concepts of surface area and volume of rectangular prisms with non-whole number units. Calculate surface area of rectangular prisms.		7.GM.1.1, 7.GM.1.2., GM.1.3
	Calculate perimeter of composite figures.	Calculate area of trapezoids and composite figures.	Develop the formula for area of trapezoids.	7.GM.2.1, 7.GM.2.2
		Solve problems that require conversions of weights and capacities.		7.GM.3.1
	Recognize that pi can be approximated by rational numbers such as $\frac{22}{7}$ and 3.14. Calculate the circumference and area of circles.	Demonstrate an understanding of the proportional relationship between the diameter and circumference of a circle.	Make connections between circumference and area to solve problems involving circles.	7.GM.3.2, 7.GM.3.3
	Determine scale factors resulting from dilations.	Use scale factors to solve problems.		7.GM.4.1
		Describe similarity and compare figures for similarity.		7.GM.4.1
	Determine side lengths of similar triangles and rectangles.	Determine areas of similar triangles and rectangles.		7.GM.4.2

Strand	Basic	Proficient	Advanced	Objective(s)
Geometry & Measurement	Describe the effect of dilations, translations, and reflections.	Apply and graph the effect of dilations, translations, and reflections.	Apply and graph rotations. Analyze the effect of dilations and multiple transformations.	7.GM.4.3
Data & Probability			Design simple experiments and use data to draw conclusions and make predictions.	7.D.1.1
	Calculate measures of central tendency and spread.		Use measures of central tendency and spread to draw conclusions about data collected and make predictions.	7.D.1.1
		Display information on circle graphs and histograms.	Interpret information from circle graphs and histograms.	7.D.1.2
		Use box plots to identify relevant data.	Analyze box plots.	7.D.1.3
	Calculate theoretical probability.	Interpret theoretical probability and draw conclusions.	Predict relative frequencies based on theoretical probabilities.	7.D.2.1, 7.D.2.2, 7.D.2.3

OSTP Math Grade 8 - Range Performance Level Descriptors (Range PLDs)

Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations	Translate between standard form and scientific notation.	Multiply and divide numbers expressed in scientific notation.		PA.N.1.2, PA.N.1.3
	Locate, identify, compare, and order rational numbers on and off a number line.	Locate, identify, compare, and order irrational numbers on and off a number line.		PA.N.1.2, PA.N.1.4
	Identify square roots of perfect squares.	Locate square roots that are irrational numbers between two consecutive positive integers.		PA.N.1.4
		Apply the properties of integer exponents.	Develop the properties of integer exponents.	PA.N.1.1
Algebraic Reasoning & Algebra	Simplify and generate equivalent expressions.	Evaluate equivalent expressions. Evaluate expressions.	Justify equivalent expressions.	PA.A.3.1, PA.A.3.2
	Solve linear equations.	Represent situations using linear equations.	Interpret solutions of linear equations.	PA.A.4.1
		Represent, write, solve, and graph inequalities.		PA.A.4.2
	Identify linear relationships.	Describe linear relationships.	Analyze linear relationships.	PA.A.2.2
		Recognize that a function is a relationship between an independent variable and a dependent variable.		PA.A.1.1

Strand	Basic	Proficient	Advanced	Objective(s)
Algebraic Reasoning & Algebra	Identify linear functions from a graph.	Identify linear functions from an equation.		PA.A.1.3
	Identify linear relationships between two variables.	Describe linear relationships between two variables.	Analyze linear relationships between two variables.	PA.A.1.3
	Describe linear functions with two variables.	Represent and solve linear functions with two variables.	Analyze linear functions with two variables and interpret results.	PA.A.1.2, PA.A.2.1, PA.A.2.3, PA.A.2.5, PA.A.4.1, PA.A.4.2, PA.A.4.3
	Identify slope.	Identify intercepts.		PA.A.2.3
		Predict the effect on the graph of a linear function when the y-intercept is changed.	Predict the effect on the graph of a linear function when the slope is changed.	PA.A.2.4
Geometry & Measurement	Calculate the surface area of rectangular prisms.	Calculate the surface area and volume of right cylinders.	Justify the formulas for volume of rectangular prisms and right cylinders.	PA.GM.2.1, PA.GM.2.2, PA.GM.2.3, PA.GM.2.4
		Use and apply the Pythagorean theorem.	Justify the Pythagorean theorem.	PA.GM.1.1, PA.GM.1.2
Data & Probability			Describe the impact that inserting or deleting a data point has on the mean and the median of a data set.	PA.D.1.1
		Explain how outliers affect measures of center and spread.		PA.D.1.2
	Collect and display information on a scatter plot.	Identify the informal line of best fit from a given scatter plot.	Interpret a scatter plot, determine the rate of change, and use a line of best fit to make predictions.	PA.D.1.3
	Identify sample spaces, classify events as independent or dependent.	Calculate experimental probability, determine how samples are chosen, and generalize samples to populations.	Interpret and predict experimental probability.	PA.D.2.1, PA.D.2.2, PA.D.2.3

APPENDIX—B
ORDERED ITEM BOOKLET BLUEPRINTS

Table 1. OSTP ELA Grades 3-8 OIB Blueprint Percentages

Grade	Source	Standard 2	Standard 3	Standard 4	Standard 5	Standard 6
3	Target #	19-21	6-9	11-13	6-9	6-9
	OIB #	15	9	11	6	7
4	Target #	15-17	9-12	11-13	6-9	6-9
	OIB #	15	9	7	6	9
5	Target #	15-17	11-13	9-11	6-9	6-9
	OIB #	15	13	11	7	8
6	Target #	17-19	9-11	9-11	6-9	6-9
	OIB #	17	11	10	6	6
7	Target #	17-19	9-11	7-10	6-9	7-10
	OIB #	17	11	7	7	8
8	Target #	12-15	12-15	7-10	6-9	6-9
	OIB #	10	18	8	7	9

Table 2. OSTP Mathematics Grades 3-8 OIB Blueprint Percentages

Grade	Source	Number & Operations	Algebraic Reasoning & Algebra	Geometry and Measurement	Data & Probability
3	Target %	44-48	12-18	22-26	12-18
	OIB %	48	14	26	12
4	Target %	42-46	12-18	24-28	12-18
	OIB %	42	18	28	12
5	Target %	42-46	14-20	22-26	12-18
	OIB %	46	18	24	12
6	Target %	38-42	20-24	22-26	12-16
	OIB %	40	22	24	14
7	Target %	16-20	26-30	30-36	18-24
	OIB %	18	28	32	22
8	Target %	16-20	44-48	18-22	14-18
	OIB %	16	44	22	18

APPENDIX C

LOGISTIC REGRESSION CALCULATION

The proficient and advanced cut scores for the OSTP ELA and mathematics grades 3-8 tests were computed using the logistic regression method as follows:

$$\log \frac{P}{1-P} = \beta_0 + \beta_1 \theta$$

which is equivalent to:

$$P = \frac{\exp(\beta_0 + \beta_1 \theta)}{1 + \exp(\beta_0 + \beta_1 \theta)}$$

Where β_0 (intercept) and β_1 (slope) are two regression coefficients that need to be computed, theta (θ) is the RP67 value associated with each OIB page, and P is the probability of observing a performance level (level X or above) given theta. After fitting the model with data, the theta cut score is obtained by finding which score corresponds to a probability of 0.5 for being rated above the cut as follows:

$$\log \frac{0.5}{1-0.5} = 0 = \beta_0 + \beta_1 \theta$$

Solving the equation, the following is obtained:

$$\theta = -\frac{\beta_0}{\beta_1}$$

Additionally, the variance of the theta estimate will be computed as:

$$VAR(\theta) = \frac{\mu_{\beta_0}^2}{\mu_{\beta_1}^2} \left[\frac{\sigma_{\beta_0}^2}{\mu_{\beta_0}^2} - 2 \frac{Cov(\beta_0, \beta_1)}{\beta_0 \beta_1} + \frac{\sigma_{\beta_1}^2}{\mu_{\beta_1}^2} \right]$$

Therefore, the standard error of the estimate is given by:


$$SE(\theta) = \sqrt{VAR(\theta)}.$$

APPENDIX—D
STANDARD-SETTING TOOLKIT

This appendix contains sample screenshots of the Cognia Standard Setting Toolkit that panelists used for all standard setting activities during the meeting. Images provided include the (1) login screen, (2) readiness survey screen, (3) ordered item booklet view, and (4) item detail view.

Figure 1. Sample Login Screen

Panelists are provided with usernames and passwords to enable secure access to the toolkit.

 **Standard Setting Toolkit** [Home](#) [Register](#) [Login](#)

Log in


Email

Password

☐ Remember me?

[Log in](#)

Figure 2. Sample Readiness Survey


 **Standard Setting Toolkit** [Home](#) [Admin](#) sstksetup@cognia.org [Logout](#)

Questionnaire

Demo Subject Grade, step 1 - Readiness Survey

Position	Question	Response
1	I understand the goals of this meeting	-- ▾
2	I understand the task at hand	-- ▾
3	I am ready to proceed with the meeting activities	-- ▾

Figure 3. Sample Ordered Item Booklet View


Standard Setting Toolkit
[Home](#)
[Admin](#)

[sstksetup@cognia.org](#)
[Logout](#)

Item Review

Demo Subject Grade Step 2 Item Review

[Hide Documents](#)
[Standards](#)
[Range PLDs](#)
[Borderline PLDs](#)


























































Position	Asset ID	Description	Point Value	Relevant KSAs	Rationale or Notes	Item Descriptor Match Level	
1	ItemID01	Short Item Description 1	1			-- 	Detail
2	ItemID02	Short Item Description 2	1			-- 	Detail
3	ItemID03	Short Item Description 3	1			-- 	Detail
4	ItemID04	Short Item Description 4	1			-- 	Detail
5	ItemID05	Short Item Description 5	1			-- 	Detail
6	ItemID06	Short Item Description 6	1			-- 	Detail
7	ItemID07	Short Item Description 7	1			-- 	Detail
8	ItemID08	Short Item Description 8	1			-- 	Detail
9	ItemID09	Short Item Description 9	1			-- 	Detail
10	ItemID10	Short Item Description 10	1			-- 	Detail
11	ItemID11	Short Item Description 11	1			-- 	Detail
12	ItemID12	Short Item Description 12	1			-- 	Detail
13	ItemID13	Short Item Description 13	1			-- 	Detail
14	ItemID14	Short Item Description 14	1			-- 	Detail
15	ItemID15	Short Item Description 15	1			-- 	Detail
16	ItemID16	Short Item Description 16	1			-- 	Detail
17	ItemID17	Short Item Description 17	1			-- 	Detail
18	ItemID18	Short Item Description 18	1			-- 	Detail
19	ItemID19	Short Item Description 19	1			-- 	Detail

Figure 4. Sample Item Detail View

Standard Setting Toolkit
Home
Admin
Logout
sstksetup@cognia.org

Item Detail

Demo Subject Grade step 2 Item Review

Return to item review

Position 2
AssetId
ItemID02
Point Value 1

Item
Standards
Range PLDs
Borderline PLDs
Prev Item
Next Item

Item

Standard

Standard code here

Standard description detail here

Relevant KSAs

Rationale or Notes

Item Descriptor Match

-->

ItemImage

1 / 1

58%

+

-

⏮

⏭

Which of the following values belong in the \square to make this statement true?

$11 \times \frac{\square}{6}$ is greater than 11.

Select the three correct answers.

☐ A. 3
☐ B. 7
☐ C. 8
☐ D. 4

APPENDIX—E
FACILITATION POWERPOINT PRESENTATION



Panel activities over the next four days

- Welcome and introductions
- Meeting norms and process overview
- Experience the test activity
- Access to the Cognia Standard Setting Toolkit
- Familiarization with content standards and PLDs (higher grade)
- Training on the Item-Descriptor (ID) Matching Method
- Modeling and practice
- Three rounds of standard setting activities (higher grade)
- Familiarization with content standards and PLDs (lower grade)
- Three rounds of standard setting activities (lower grade)
- Final workshop evaluation survey



Welcome & introductions

- Facilitator introduction
 - Name, role at Cognia, role during standard setting
- Panelist introductions
 - Your name, district, what you teach
- Experience on assessment program committees
 - Item Reviews
 - Alignment Studies
 - Standard Setting
 - Others



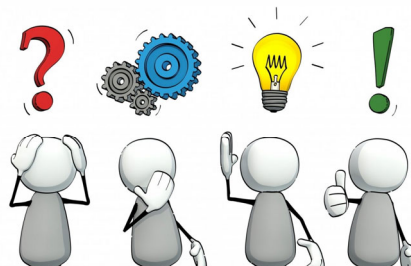
A Shift in Focus for this Week



OTHER WAYS YOU HAVE CONTRIBUTED

- Item writing, data review, content review and/or item review committees
- Review test items
- Purpose: Evaluate items for use on a test (potential problems with the items; suggest improvements)

THERE IS A
DIFFERENCE



THE WORK WE ARE DOING THIS WEEK

- Standard setting: Item-centered method with content-based judgment
- Look at test items
- Purpose: Identify the knowledge, skills, and abilities required to correctly answer the item



Meeting norms

- All conversations are confidential.
- Outside of this meeting, please **DO** talk about the general process we undertake, but **DO** disclose the specifics.
- Please **DO NOT**:
 - Use any personal devices in the room; you may step out at any time if needed.
 - Use the Chromebooks for anything other than the standard setting activities.
 - Take any of your notes or work with you when you leave the room.



Overview: Goals and expectations

Our shared goals

- Collect your recommendations on performance standards for the OSTP ELA or Math assessments that provide meaningful and actionable information

Your goals as panelists

- Learn concepts and procedures following the Item-Descriptor (ID) Matching Method
- Follow the procedures to complete the standard setting activities
- Make content-based judgments about test items
- Rely on your expertise about the content standards and student learning throughout the process



Breakout session: Schedule for day 1

Time	Activities



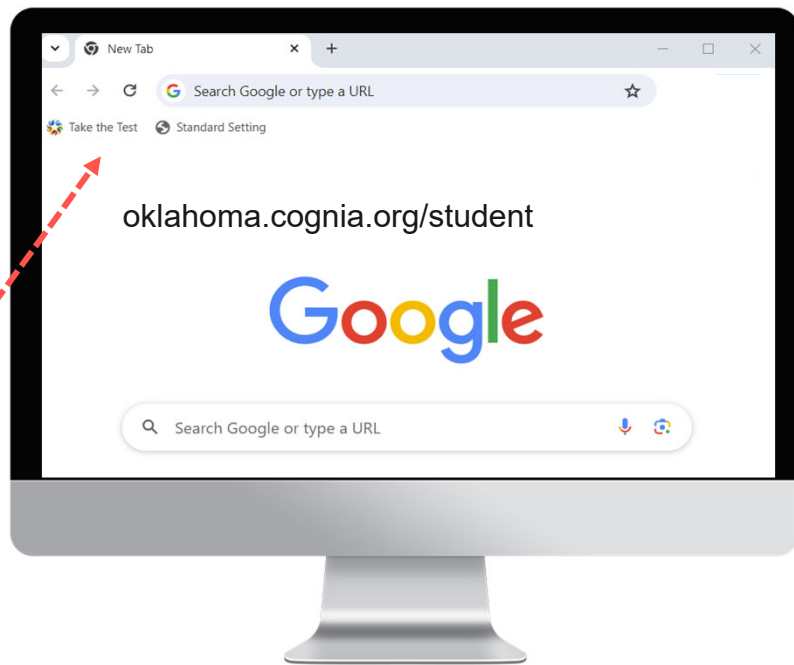
Experience the test activity

- You will experience the OSTP test in a format similar to the student experience.
- Purpose: Get familiar with the items as they appeared to students.
- Activity notes:
 - This session is scheduled for a duration of **45 mins**
 - Briefly examine the test items in the testing platform
 - Try not to linger on any one item
 - If you see any item sets, keep in mind that these sets will appear together in the testing platform but will not appear together when you work with them during the standard setting (more on this later)



Guidance: Take the test

1. Chromebook:
navigate to
Google Chrome
browser
2. Click on “Take the
Test” link - top left.
3. Use the log in
credentials
provided to
access the test.



Experience the test - discussion

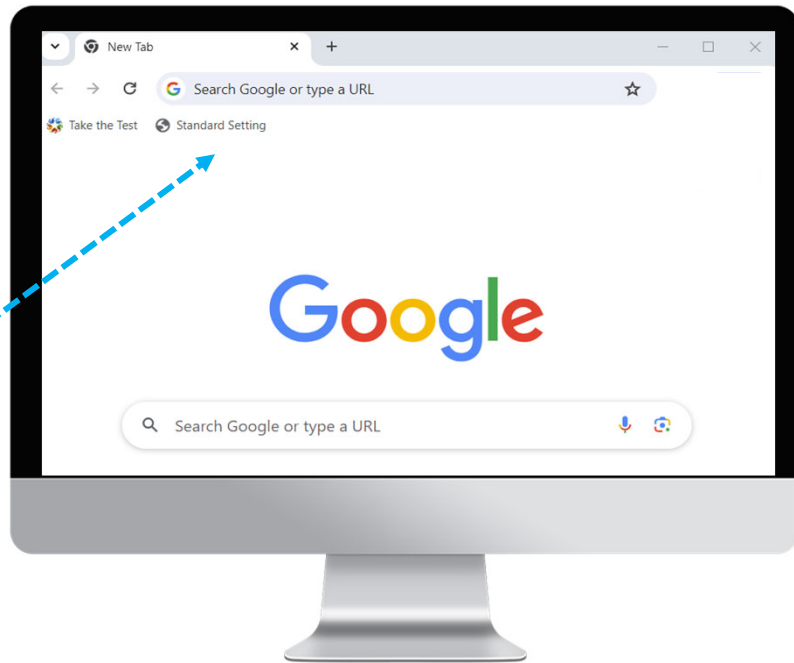


- Brief discussion
- Share thoughts/observations
based on your experience
with the test.



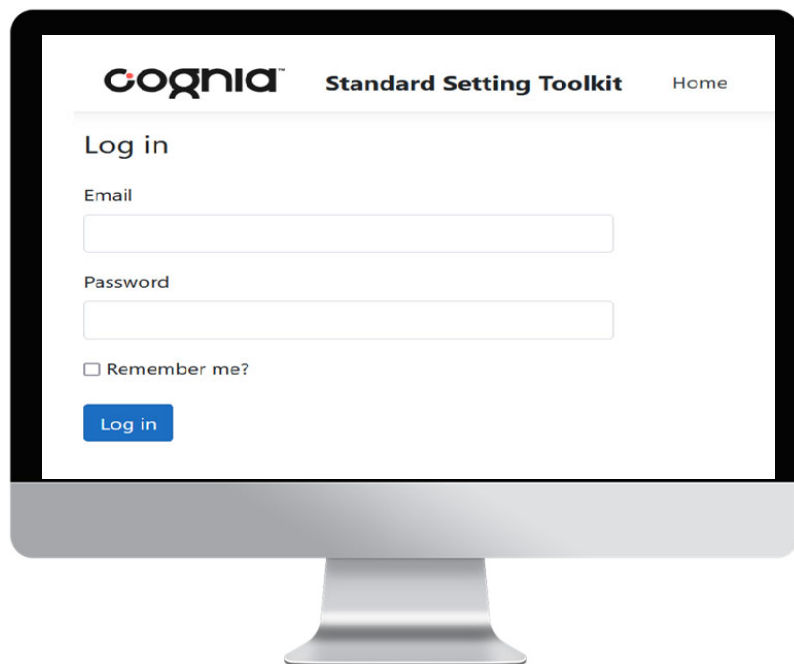
Guidance: Cognia Toolkit

1. Chromebook:
Navigate to
Google Chrome
browser
2. Click on
“Standard Setting”
link in the top left.



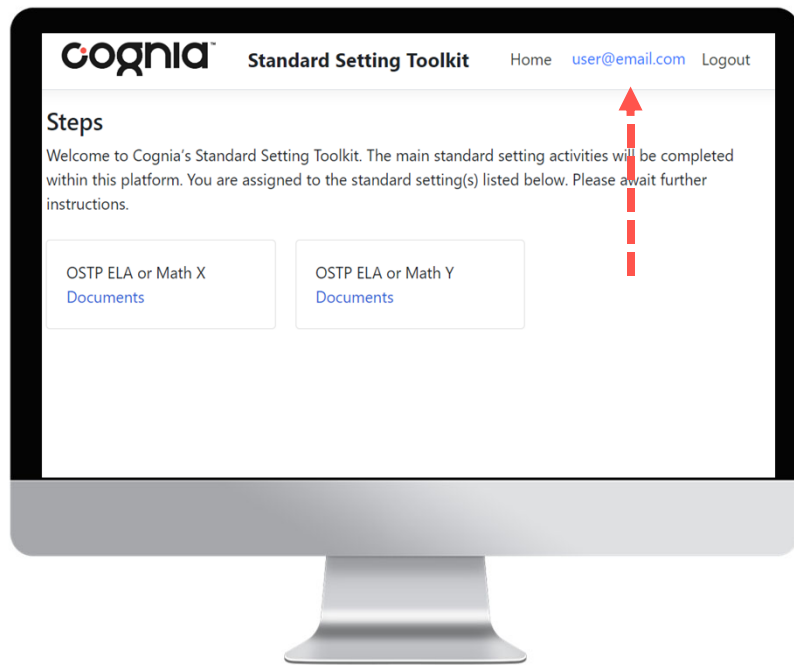
Cognia Toolkit

- Email
 - Registration email
 - All lowercase
- Initial Password
- After initial log in you will change your password



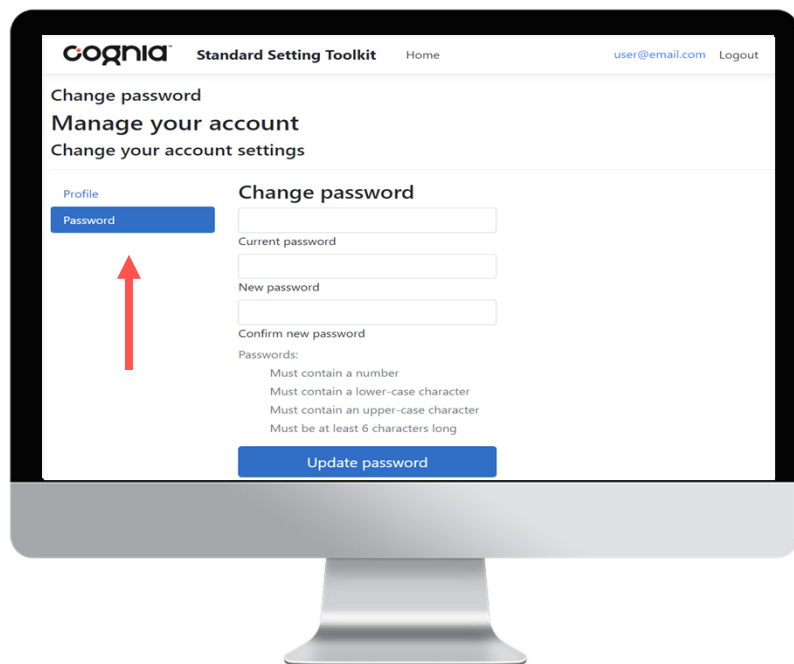
Change your password

- Click on your email - top right corner
- This will bring you to a profile page



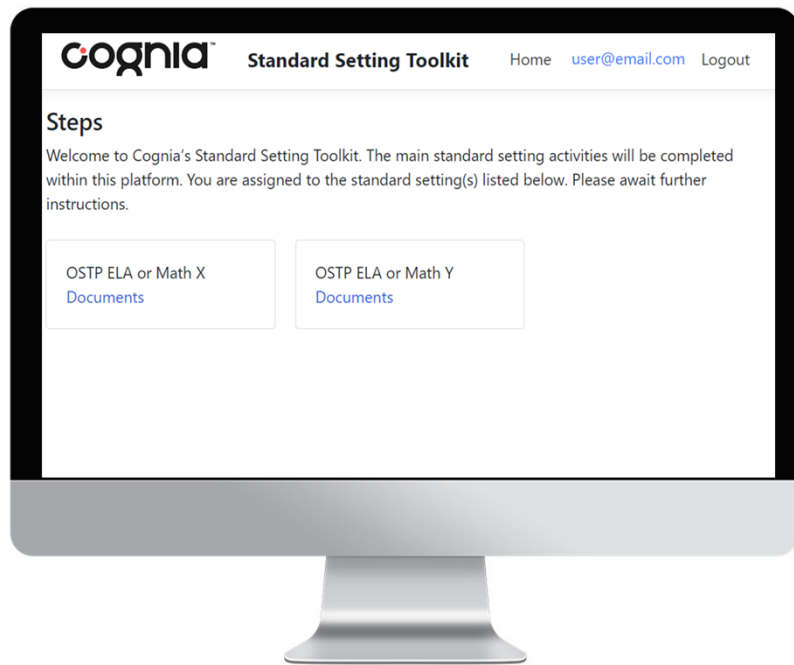
Change your password

- Click “Password” on the left menu
- Enter the initial password
- Enter new password
- Click “Update password”
- Log out
- Log back in with updated password



You should now be back on the following screen

Please confirm that you see the correct content area and two grades that you have been assigned



Review content standards & PLDs

- Review subject-specific content standards
- Obtain an understanding of the performance level descriptors (PLDs) in relation to content standards
 - This activity is critical because you will make judgments based on your understanding of PLDs.
 - The standards and PLD documents will be used throughout the workshop as you engage in the standard setting process.



Reminder: Performance Level Descriptors (PLDs)

- Provide a narrative account of the knowledge, skills, and abilities demonstrated by students in each level of achievement.
- Describe what students know and can do based on the Oklahoma Academic Standards.
- Inform stakeholders of how to interpret student test scores in relation to the Oklahoma Academic Standards.
- Are typically used for standard setting and score reporting.



Performance level descriptors (PLDs)

- Performance Levels
 - Below Basic
 - Basic
 - Proficient
 - Advanced
- Performance Level Descriptors (PLDs) represent intended interpretations of solid student achievement on the assessment for each level.
- Development of the PLDs began with the assumption that the grade-level content standards represent what students should know and be able to do at the end of a given grade level. Prior research on learning, cognition, and development in the subject areas, a variety of resources, and teaching experiences of content experts informed the development of definitions for solid achievement at each level.



Study and discuss performance level descriptors (PLDs)

- In-depth review/discussion of performance level descriptors (PLDs)
- Reach common understanding of what it means to be in each performance level.



Topics: Key concepts and processes

- The Item-Descriptor (ID) Matching method overview
- Ordered Item Booklet (OIB)
- ID Matching process
 - Standard setting judgment task
 - Nature of content-based judgment
 - Iterative 3-round process
- Modeling & Practice
 - Work with sample items
 - Learn how to navigate in the Toolkit



Item-Descriptor (ID) Matching Method for standard setting

Item-centered method

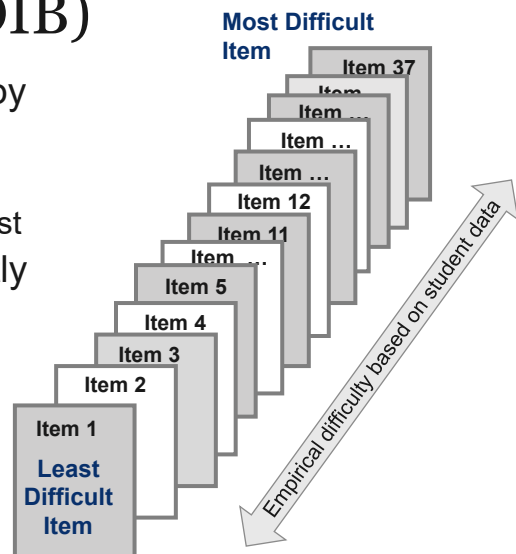
Content-based judgment

Individual judgments



















Ordered item booklet (OIB)

- The OIB contains test items ordered by difficulty.
- Each OIB page represents an item.
 - Easiest item first and the most difficult last
- The difference in difficulty is not exactly the same between each pair of neighboring items.
- Difficulty is based on data from the students who answered the items during prior administrations.



OIB in the Standard Setting Toolkit

<div> Hide Documents Standards Range PLDs Borderline PLDs </div>						
Position	Asset ID	Description	Point Value	KSAs & Reasoning	Notes	Item Descriptor Match Level
1	147542A	Item 1	1			-- Detail
2	636410	Item 2	1			-- Detail
3	147741A	Item 3	1			-- Detail
4	733131	Item 4	1			-- Detail
5	154758A	Item 5	1			-- Detail
6	733127	Item 6	1			-- Detail
7	479031	Item 7	1			-- Detail
8	154760A	Item 8	1			-- Detail

ID Matching process

For each item in the OIB:

1. Review the item and identify the KSAs

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

What does a student need to know or be able to do to correctly respond to this item?

2. Make an item-PLD alignment judgment

- Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced performance level descriptor (PLD).

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

ID-Matching process considerations



Useful

- Based on Content
- Links items to PLDs
- Refers to specific knowledge, skills, and abilities (KSAs)

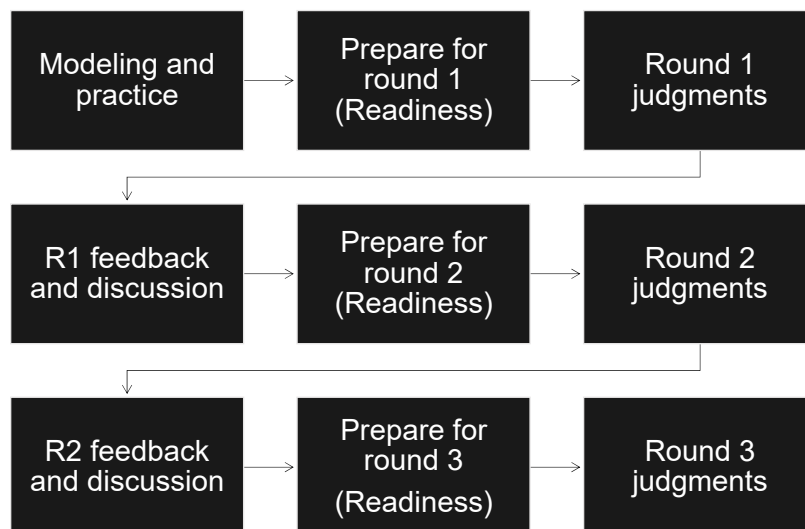


Not Useful

- Based on something other than the content (i.e., item quality)
- Too general
- Based on a specific student or class

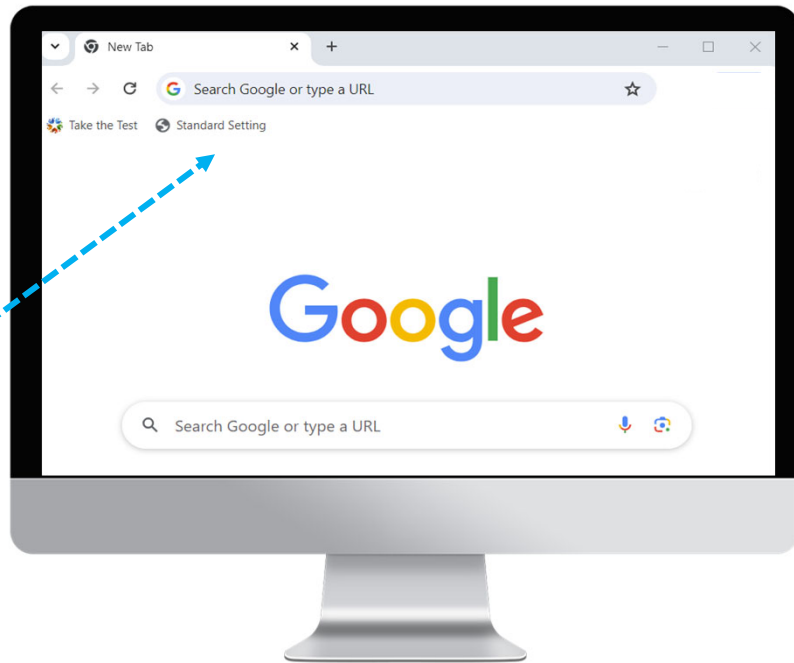


Overview: ID-Matching over 3 rounds



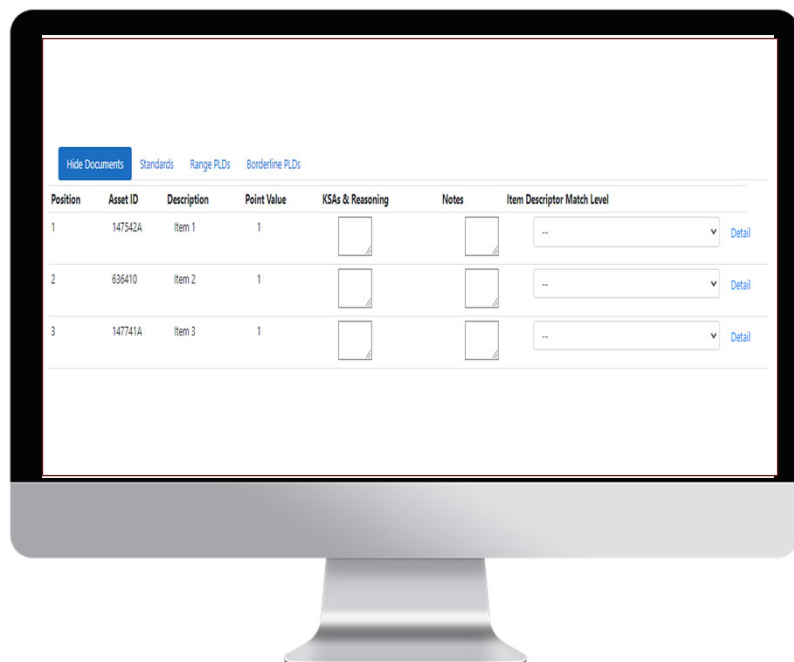
Guidance: Cognia Toolkit

1. Chromebook:
Navigate to
Google Chrome
browser
2. Click on
“Standard Setting”
link in the top left.



Practice round

- In the Toolkit, you will automatically be redirected to the practice round.
- You will see a list of sample items.
- Please make sure your screen shows the correct content area and grade



Modeling & practice of the ID-Matching judgmental task

We will begin by working with the first (top) item in the sample list.

1. Review the item and identify **KSAs**.

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

What does a student need to know or be able to do to correctly respond to this item?



Modeling & practice of the ID-Matching judgmental task

Continue working with the first (top) item in the sample list.

2. Match item to a PLD level

- Match the KSAs required by the item with the expectations described in either the **Basic**, **Proficient**, or **Advanced** performance level descriptor (PLD) for that standard.
- If not already done, be sure to add a note to the KSAs text box about the reasoning for the match.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?



Examples: KSAs & Reasoning

- Useful example:
 - The item requires students to connect fractions or decimals using models. Students are not just representing tenths or hundredths in one form, but moving between two different forms of a number.
- Not useful example:
 - The item matches the Proficient PLD and does not match the Basic PLD.



Reminder: ID-Matching process considerations



Useful

- Based on Content
- Links items to PLDs
- Refers to specific knowledge, skills, and abilities (KSAs)



Not Useful

- Based on something other than the content (i.e., item quality)
- Too general
- Based on a specific student or class



Practice round - Review

- Reviewed sample items and for each one:
 1. Identified the knowledge, skills, and abilities (KSAs) required to correctly respond to the item.
 2. Matched the item to either the Basic, Proficient, or Advanced PLD.
 - Included note about reasoning for PLD match in KSAs box where needed.
- Borderline considerations
 - Some items might be in the border between two adjacent PLDs.
 - Select the PLD that most closely matches the item.
 - Make notes for yourself next to these items to inform discussions later.
- Remaining questions or concerns?



Round 1 – Readiness

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and is ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.
 - Responses are reviewed in summary only



Position	Question	Response
1	Question 1	..
2	Question 2	..
3	Question 3	..
4	Question 5	..
5	Question 6	..
6	Question 7	..

Round 1 judgments

- You will now be redirected to Round 1
 - In the Toolkit you will see the full list of OIB items.
- Reminder – Your task for each item:
 1. Identify the KSAs
 2. Match the item to one of the PLDs
 - Use the “Notes” box for additional notes (*for example*: when an item seems to be in-between two PLDs)
- Item-PLD alignment is an individual activity. Please **DO NOT** discuss your work with your colleagues at this time.



Round 1 judgments

For each item in the OIB:

1. Review the item and identify KSAs.

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

What does a student need to know or be able to do to correctly respond to this item?

2. Make item-PLD alignment judgment.

- Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced PLD.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*





Breakout session – Agenda (day 2)

- Debrief day 1
- Complete round 1 judgments
- Lunch
- Discussion and preparation for round 2
- Begin round 2 judgments



Breakout session: Schedule for day 2

Time	Activities
08:30 AM – 09:15 AM	Debrief day 1 (Check-in on the process, challenges, etc.)
09:15 AM – 12:00 PM	Complete round 1
12:00 PM – 01:00 PM	Lunch break
01:00 PM – 02:30 PM	Discuss round 1 feedback/results; Introduce benchmarks; Prepare for round 2.
02:30 PM – 05:00 PM	Begin round 2
05:00 PM	Adjourn for the day



Debrief day 1

- Great job training, learning, being on
- Individuals are about $\frac{1}{4}$ to $\frac{1}{2}$ way through the items
- Feedback on Round 1 so far:
 - KSAs can be brief – 10-15 words max – but make sure language lines up
 - Be sure to look at all the PLD descriptors in the row
- Questions or thoughts from yesterday?



Round 1 judgments

For each item in the OIB:

1. Review the item and identify KSAs.

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

2. Make item-PLD alignment judgment.

- Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced PLD.

- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*

What does a student need to know or be able to do to correctly respond to this item?

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?



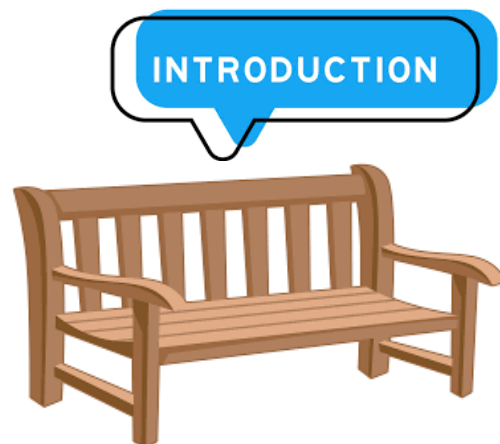
Feedback and Discussion

- The goal of the discussion is to hear perspectives from your fellow panelists
 - Additional information for your consideration
 - NOT meant to persuade or influence
- In the Toolkit, you will see your own data from Round 1
 - The only field you can use during this time is the “Notes” field.



Introduction to benchmarks

- Content-based information based on work from the Cognia/SDE content specialists
- Benchmarks serve as additional information for your consideration
- Will be presented as shaded rows in the OIB



Content-based benchmarks

- The shaded regions are calculated based on judgments from Cognia and SDE content specialists.
- This region represents a transition area where items between two performance levels are beginning to intersect.
- It is vital that we have the input of educators who teach to these standards and the Oklahoma student population.
- To that end, your results may very well differ from theirs.
- The content-based benchmarks provide additional information for your consideration but is not meant to constrain or persuade your judgments.



Round 2 – Readiness survey

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.

A screenshot of a questionnaire form displayed on a tablet. The form is titled "Questionnaire" and "OSTP ELA or Math X, step 2 - Round 1 Readiness". It contains a table with three columns: "Position", "Question", and "Response". The table has six rows, each with a position number, a question label, and a response dropdown menu.

Position	Question	Response
1	Question 1	..
2	Question 2	..
3	Question 3	..
4	Question 5	..
5	Question 6	..
6	Question 7	..

Round 2 judgments

- You will now be redirected to Round 2
 - In the toolkit, you will see the same list of items with your work from round 1 (notes and judgments)
 - You will also see the shaded regions for the content-based benchmarks
- **Reminder – Your task:**
 - Review items in the benchmark (shaded) regions, items discussed during round 1 feedback discussion, and items you were previously unsure about
 - Consider the KSAs, then decide to keep or change your initial PLD match
- Item-PLD alignment is an individual activity. Please **DO NOT** discuss your work with your colleagues at this time.



Round 2 judgments

- **Decide to retain/adjust your judgments:**
 - Review items we discussed, items in benchmark regions, and items you were previously unsure about.
 - Consider the KSAs and decide to keep or change your initial PLD match.

What does a student need to know or be able to do to correctly respond to this item?

- **Reminder:**
 1. Review the item and identify KSAs.
 2. Make item-PLD alignment judgment.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*





Breakout session – Agenda (day 3)

- Feedback/discussion of round 2 results
- Preparation for round 3
- Complete round 3 judgments
- Review standards and PLDs for the lower grade
- Prepare for and begin round 1 judgments



Debrief day 2

- Great job with following process!
- Focus on PLD interpretations and clarifications as we discuss round 2 results
- Questions or thoughts from yesterday?



Round 2 judgments

- **Decide to retain/adjust your judgments:**
 - Review items we discussed, items in benchmark regions, and items you were previously unsure about.
 - Consider the KSAs and decide to keep or change your initial PLD match.
- **Reminder:**
 1. Review the item and identify KSAs.
 2. Make item-PLD alignment judgment.
- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*

What does a student need to know or be able to do to correctly respond to this item?

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

Round 3 – Readiness survey

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.

Questionnaire
OSTP ELA or Math X, step 2 - Round 1 Readiness

Position	Question	Response
1	Question 1	.. ▾
2	Question 2	.. ▾
3	Question 3	.. ▾
4	Question 5	.. ▾
5	Question 6	.. ▾
6	Question 7	.. ▾

Round 3 judgments

- You will now be redirected to Round 3
 - In the toolkit, you will see the same list of items with your work from round 2 (notes and judgments)
- Reminder – Your task:
 - Review items discussed during round 2 feedback discussion, and items you were previously unsure about
 - Consider the KSAs, then decide to keep or change your initial PLD match
- Item-PLD alignment is an individual activity. Please **DO NOT** discuss your work with your colleagues at this time.



Round 3 judgments

- **Decide to retain/adjust your judgments:**
 - Review items we discussed, items in benchmark regions, and items you were previously unsure about.
 - Consider the KSAs and decide to keep or change your initial PLD match.

What does a student need to know or be able to do to correctly respond to this item?

- **Reminder:**
 1. Review the item and identify KSAs.
 2. Make item-PLD alignment judgment.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*

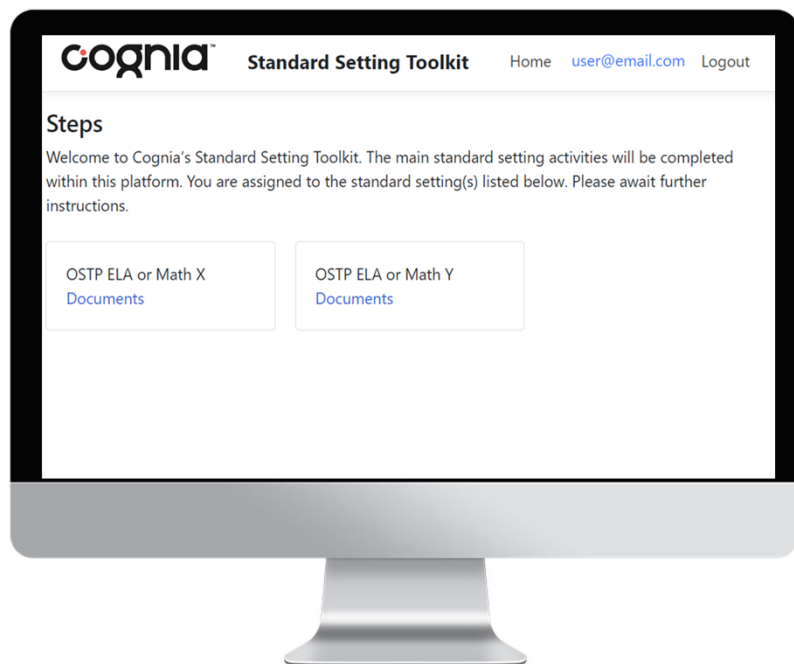


Review content standards & PLDs

- Review subject-specific content standards
- Obtain an understanding of the performance level descriptors (PLDs) in relation to content standards
 - This activity is critical because you will make judgments based on your understanding of PLDs.
 - The standards and PLD documents will be used throughout the workshop as you engage in the standard setting process.



Reminder:
Standards and
PLDs are linked
on the home page



Reminder: Performance Level Descriptors (PLDs)

- Performance Levels
 - Below Basic
 - Basic
 - Proficient
 - Advanced
- Performance level descriptors:
 - Describe what students know and can do based on the Oklahoma Academic Standards.
 - Represent intended interpretations of solid student achievement on the assessment for each level.
 - Inform stakeholders of how to interpret student test scores in relation to the Oklahoma Academic Standards.



Study and discuss performance level descriptors (PLDs)

- In-depth review/discussion of performance level descriptors (PLDs)
- Reach common understanding of what it means to be in each performance level.



Round 1 – Readiness survey

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.



Position	Question	Response
1	Question 1	...
2	Question 2	...
3	Question 3	...
4	Question 5	...
5	Question 6	...
6	Question 7	...

Round 1 judgments

- You will now be redirected to Round 1
 - In the Toolkit you will see the full list of OIB items.
- Reminder – Your task for each item:
 1. Identify the KSAs
 2. Match the item to one of the PLDs
 - Use the “Notes” box for additional notes (*for example*: when an item seems to be in-between two PLDs)
- Item-PLD alignment is an individual activity. Please **DO NOT** discuss your work with your colleagues at this time.



Round 1 judgments

For each item in the OIB:

1. Review the item and identify KSAs.

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

What does a student need to know or be able to do to correctly respond to this item?

2. Make item-PLD alignment judgment.

- Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced PLD.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

- ✓ *Write note about reasoning for your PLD match in the KSAs field*
- ✓ *Work independently*
- ✓ *Trust your expertise*



Content-based benchmarks

- The shaded regions are calculated based on judgments from other Cognia/SDE content specialists.
- This region represents a transition area where items between two performance levels are beginning to intersect.
- It is vital that we have the input of educators who teach to these standards and the OK student population.
- To that end, your results may very well differ from theirs.
- The content-based benchmarks provide additional information for your consideration but is not meant to constrain or persuade your judgments.

Round 2 – Readiness survey

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.



Questionnaire
OSTP ELA or Math X, step 2 - Round 1 Readiness

Position	Question	Response
1	Question 1	-- ▾
2	Question 2	-- ▾
3	Question 3	-- ▾
4	Question 5	-- ▾
5	Question 6	-- ▾
6	Question 7	-- ▾

Round 2 judgments

- You will now be redirected to Round 2
 - In the toolkit, you will see the same list of items with your work from round 1 (notes and judgments)
 - You will also see the shaded regions for the content-based benchmarks
- Reminder – Your task:
 - Review items in the benchmark (shaded) regions, items discussed during round 1 feedback discussion, and items you were previously unsure about
 - Consider the KSAs, then decide to keep or change your initial PLD match
- Item-PLD alignment is an individual activity. Please **DO NOT** discuss your work with your colleagues at this time.

Round 2 judgments

- **Decide to retain/adjust your judgments:**

- Review items we discussed, items in benchmark regions, and items you were previously unsure about.
- Consider the KSAs and decide to keep or change your initial PLD match.

What does a student need to know or be able to do to correctly respond to this item?

- **Reminder:**

1. Review the item and identify KSAs.
2. Make item-PLD alignment judgment.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

✓ *Write note about reasoning for your PLD match in the KSAs field*

✓ *Work independently*

✓ *Trust your expertise*



Breakout session – Agenda (day 4)

- Debrief day 3
- Round 2 feedback
- Discussion and preparation for round 3
- Complete round 3 judgments
- Wrap – final data
- Evaluation survey



Debrief day 3

- All panelists finished R2 judgments
- Focus on listening and considering all for R3 judgments – would expect some convergence of interpretations and judgments
- If on the fence between levels being used, can consider where in OIB the item is – “skills being used”
- Questions or thoughts from yesterday?



Round 3 – Readiness survey

- In a moment, you will be redirected in the Toolkit to a short survey.
- Goal: Determine if everyone understands the task at hand and ready to proceed.
- Read and answer each question.
- Once everyone has completed the survey, we will review responses and proceed accordingly.



Questionnaire

OSTP ELA or Math X, step 2 - Round 1 Readiness

Position	Question	Response
1	Question 1	..
2	Question 2	..
3	Question 3	..
4	Question 5	..
5	Question 6	..
6	Question 7	..

Round 3 judgments

- **Decide to retain/adjust your judgments:**

- Review items we discussed, items in benchmark regions, and items you were previously unsure about.
- Consider the KSAs and decide to keep or change your initial PLD match.

What does a student need to know or be able to do to correctly respond to this item?

- **Reminder:**

1. Review the item and identify KSAs.
2. Make item-PLD alignment judgment.

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

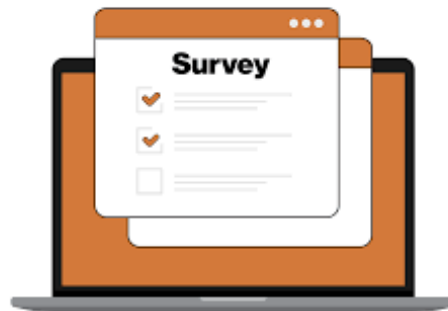
✓ *Write note about reasoning for your PLD match in the KSAs field*

✓ *Work independently*

✓ *Trust your expertise*

Final Workshop Evaluation Survey

- In a moment, you will be redirected in the Toolkit to the final workshop evaluation survey.
- Your responses serve as additional data for us to consider.
- Please do not leave until you have completed the survey.
- Note for those participating in articulation: You will reconvene tomorrow morning after breakfast.



APPENDIX—F
PANELIST INFORMATION

Table 1. OK OSTP ELA Grades 3-4 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Taylor	3	--	--
2	Glencoe Public Schools	2	44% Male, 55% Female	0.05% Hispanic, 11% AI, 0% Asian, 0.01% AA, 0% PI, 64% White, 17% Multiracial
3	Cleora	3	48% Male, 52% Female	0.07% Hispanic, 46% AI, 0.01% Asian, 0.01% AA, 0.01% PI, 45% White, 0.01% Multiracial
4	Mason	2	45% Male, 55% Female	0.02% Hispanic, 23% AI, 0% Asian, 0% AA, 0% PI, 46% White, 28% Multiracial
5	Geary	7	47% Male, 53% Female	14% Hispanic, 29% AI, 0% Asian, 0.02% AA, 0% PI, 45% White, .1% Multiracial
6	Deer Creek Public Schools	8	52% Male, 48% Female	13% Hispanic, 0.03% AI, .1% Asian, 0.08% AA, 0% PI, 57% White, 0.09% Multiracial
7	Collinsville School District	13	53% Male, 47% Female	.1% Hispanic, 12% AI, 0.05% Asian, 0.01% AA, 0% PI, .5% White, 21% Multiracial
8	Shawnee Public Schools	1	52% Male, 48% Female	14% Hispanic, 12% AI, 0% Asian, 0.05% AA, 0% PI, 44% White, 25% Multiracial
9	Keystone	5	54% Male, 46% Female	0.04% Hispanic, 12% AI, 0% Asian, 0% AA, 0% PI, 66% White, 18% Multiracial
10	Inola Public Schools	2	55% Male, 44% Female	0.06% Hispanic, 25% AI, 0.05% Asian, 0.01% AA, 0% PI, 48% White, 16% Multiracial
11	Glenpool Public Schools	3	51% Male, 49% Female	11% Hispanic, 16% AI, .1% Asian, 0.03% AA, 0% PI, 41% White, 19% Multiracial

Table 2. OK OSTP ELA Grades 5-6 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Santa Fe South Public Charter	3	--	--
2	Vian	3	53% Male, 47% Female	0.04% Hispanic, 45% AI, 0.01% Asian, 0.03% AA, 0% PI, 32% White, 15% Multiracial
3	Pryor Public Schools	2.5	.5% Male, .5% Female	0.07% Hispanic, 26% AI, 0.01% Asian, 0% AA, 0% PI, 42% White, 24% Multiracial
4	Deer Creek	15	52% Male, 48% Female	13% Hispanic, 0.03% AI, .1% Asian, 0.08% AA, 0% PI, 57% White, 0.09% Multiracial
5	Guthrie public schools	26	52% Male, 48% Female	19% Hispanic, 0.03% AI, 0% Asian, 0.06% AA, 0% PI, 58% White, 12% Multiracial
6	Paden	5+	56% Male, 44% Female	0.06% Hispanic, .2% AI, 0.02% Asian, 0.04% AA, 0% PI, 51% White, 16% Multiracial
7	Tulsa Public Schools	15	51% Male, 49% Female	38% Hispanic, 0.04% AI, 0.02% Asian, 22% AA, 0.01% PI, 21% White, 11% Multiracial
8	Edmond Schools	1	52% Male, 48% Female	13% Hispanic, 0.02% AI, 0.05% Asian, 11% AA, 0% PI, 57% White, 12% Multiracial
9	Hilldale Public Schools	19	51% Male, 49% Female	.1% Hispanic, .3% AI, 0.01% Asian, 0.02% AA, 0% PI, 39% White, 17% Multiracial
10	Putnam City Schools	3	51% Male, 49% Female	39% Hispanic, 0.02% AI, 0.04% Asian, 24% AA, 0% PI, 21% White, 11% Multiracial

Table 3. OK OSTP ELA Grades 7-8 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Oklahoma City Public Schools	8	51% Male, 49% Female	57% Hispanic, 0.02% AI, 0.02% Asian, .2% AA, 0% PI, 11% White, 0.08% Multiracial
2	Santa Fe South Schools	3	--	--
3	Oklahoma City	2	51% Male, 49% Female	57% Hispanic, 0.02% AI, 0.02% Asian, .2% AA, 0% PI, 11% White, 0.08% Multiracial
4	Santa Fe South Schools	9	--	--
5	Bristow Public Schools	9	52% Male, 48% Female	0.04% Hispanic, .2% AI, 0.01% Asian, 0.06% AA, 0% PI, 56% White, 14% Multiracial
6	Dove Schools	16	52% Male, 48% Female	63% Hispanic, 0.02% AI, 0.02% Asian, 12% AA, 0% PI, 14% White, 0.06% Multiracial
7	Broken Arrow Public Schools	4	51% Male, 49% Female	19% Hispanic, 0.07% AI, 0.04% Asian, 0.07% AA, 0% PI, 49% White, 14% Multiracial
8	Okeene Public Schools	30+	56% Male, 44% Female	.2% Hispanic, 0.05% AI, 0% Asian, 0.01% AA, 0% PI, 74% White, 0.01% Multiracial
9	John Rex Charter School	5	.5% Male, .5% Female	29% Hispanic, 0.02% AI, 0.04% Asian, 17% AA, 0% PI, 35% White, 14% Multiracial
10	Elk City	4	51% Male, 49% Female	24% Hispanic, 0.03% AI, 0.01% Asian, 0.04% AA, 0% PI, .6% White, 0.08% Multiracial

Table 4. OK OSTP Mathematics Grades 3-4 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Lawton	12	51% Male, 49% Female	24% Hispanic, 0.05% AI, 0.01% Asian, .2% AA, 0.01% PI, 32% White, 17% Multiracial
2	Deer Creek School District	16	52% Male, 48% Female	13% Hispanic, 0.03% AI, .1% Asian, 0.08% AA, 0% PI, 57% White, 0.09% Multiracial
3	Coweta Public Schools	25	52% Male, 48% Female	0.08% Hispanic, 23% AI, 0.01% Asian, 0.04% AA, 0% PI, 55% White, 0.07% Multiracial
4	Glencoe Public Schools	6	44% Male, 55% Female	0.05% Hispanic, 11% AI, 0% Asian, 0.01% AA, 0% PI, 64% White, 17% Multiracial
5	Putnam City Schools	1	51% Male, 49% Female	39% Hispanic, 0.02% AI, 0.04% Asian, 24% AA, 0% PI, 21% White, 11% Multiracial
6	Bartlesville Public Schools	1	52% Male, 48% Female	13% Hispanic, .1% AI, 0.02% Asian, 0.03% AA, 0% PI, 52% White, .2% Multiracial
7	Bartlesville public schools	1	52% Male, 48% Female	13% Hispanic, .1% AI, 0.02% Asian, 0.03% AA, 0% PI, 52% White, .2% Multiracial
8	Bridge Creek	24	52% Male, 48% Female	14% Hispanic, 0.05% AI, 0% Asian, 0.01% AA, 0% PI, 66% White, 14% Multiracial
9	Keystone	6	54% Male, 46% Female	0.04% Hispanic, 12% AI, 0% Asian, 0% AA, 0% PI, 66% White, 18% Multiracial
10	Moore Public Schools	15	51% Male, 49% Female	23% Hispanic, 0.04% AI, 0.05% Asian, 0.08% AA, 0% PI, 43% White, 17% Multiracial
11	Bartlesville Public Schools	3	52% Male, 48% Female	13% Hispanic, .1% AI, 0.02% Asian, 0.03% AA, 0% PI, 52% White, .2% Multiracial

Table 5. OK OSTP Mathematics Grades 5-6 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Hilldale Public Schools	33	51% Male, 49% Female	.1% Hispanic, .3% AI, 0.01% Asian, 0.02% AA, 0% PI, 39% White, 17% Multiracial
2	Union Public school	4	.5% Male, .5% Female	41% Hispanic, 0.04% AI, 0.07% Asian, 15% AA, 0% PI, 23% White, .1% Multiracial
3	Moore Public Schools	20	51% Male, 49% Female	23% Hispanic, 0.04% AI, 0.05% Asian, 0.08% AA, 0% PI, 43% White, 17% Multiracial
4	Chelsea	20	53% Male, 47% Female	0.06% Hispanic, 34% AI, 0.02% Asian, 0.01% AA, 0% PI, 34% White, 23% Multiracial
5	Walters	1	51% Male, 49% Female	11% Hispanic, 0.09% AI, 0.01% Asian, 0.01% AA, 0% PI, .6% White, 18% Multiracial
6	Stillwater	3	52% Male, 48% Female	13% Hispanic, 0.05% AI, 0.04% Asian, 0.06% AA, 0% PI, 58% White, 13% Multiracial
7	Washington Public School	2	52% Male, 48% Female	0.06% Hispanic, 11% AI, 0% Asian, 0.01% AA, 0% PI, 79% White, 0.02% Multiracial
8	Weatherford Public Schools	30	53% Male, 47% Female	23% Hispanic, 0.06% AI, 0.01% Asian, 0% AA, 0% PI, 59% White, 11% Multiracial
9	Shawnee Public Schools	16	52% Male, 48% Female	14% Hispanic, 12% AI, 0% Asian, 0.05% AA, 0% PI, 44% White, 25% Multiracial
10	Owasso	20	52% Male, 48% Female	15% Hispanic, 0.07% AI, 0.06% Asian, 0.04% AA, 0% PI, 53% White, 16% Multiracial
11	Oklahoma City Public Schools	22	51% Male, 49% Female	57% Hispanic, 0.02% AI, 0.02% Asian, .2% AA, 0% PI, 11% White, 0.08% Multiracial

Table 6. OK OSTP Mathematics Grades 7-8 Standard Setting Panel Participant List

Panelist #	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Putnam City Schools	6	51% Male, 49% Female	39% Hispanic, 0.02% AI, 0.04% Asian, 24% AA, 0% PI, 21% White, 11% Multiracial
2	Central High	11	49% Male, 51% Female	11% Hispanic, 0.05% AI, 0% Asian, 0.01% AA, 0% PI, 74% White, 0.09% Multiracial
3	Tulsa Public Schools	2	51% Male, 49% Female	38% Hispanic, 0.04% AI, 0.02% Asian, 22% AA, 0.01% PI, 21% White, 11% Multiracial
4	Epic Charter School	16	49% Male, 51% Female	15% Hispanic, 0.06% AI, 0.01% Asian, 0.07% AA, 0% PI, 51% White, 21% Multiracial
5	Ada	7	51% Male, 49% Female	15% Hispanic, 21% AI, 0.01% Asian, 0.02% AA, 0% PI, 37% White, 24% Multiracial
6	Mustang	8	51% Male, 49% Female	19% Hispanic, 0.03% AI, 0.04% Asian, 0.06% AA, 0% PI, 54% White, 13% Multiracial
7	Vinita Public Schools	22	51% Male, 49% Female	0.05% Hispanic, 26% AI, 0.04% Asian, 0% AA, 0% PI, 41% White, 23% Multiracial
8	Stigler	23	54% Male, 46% Female	0.08% Hispanic, 36% AI, 0.01% Asian, 0% AA, 0% PI, .5% White, 0.05% Multiracial
9	Stilwell	33	51% Male, 49% Female	18% Hispanic, 47% AI, 0.02% Asian, 0% AA, 0% PI, 18% White, 16% Multiracial
10	Broken Arrow Public Schools	5	51% Male, 49% Female	19% Hispanic, 0.07% AI, 0.04% Asian, 0.07% AA, 0% PI, 49% White, 14% Multiracial
11	Ada City School	8	51% Male, 49% Female	15% Hispanic, 21% AI, 0.01% Asian, 0.02% AA, 0% PI, 37% White, 24% Multiracial
12	Stillwater	13	52% Male, 48% Female	13% Hispanic, 0.05% AI, 0.04% Asian, 0.06% AA, 0% PI, 58% White, 13% Multiracial

Table 7. OK OSTP ELA Articulation Panel Participant List

Panelist #	Standard Setting Panel	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	ELA 3-4	Keystone	5	54% Male, 46% Female	0.04% Hispanic, 12% AI, 0% Asian, 0% AA, 0% PI, 66% White, 18% Multiracial
2	ELA 3-4	Inola Public Schools	2	55% Male, 44% Female	0.06% Hispanic, 25% AI, 0.05% Asian, 0.01% AA, 0% PI, 48% White, 16% Multiracial
3	ELA 3-4	Glenpool Public Schools	3	51% Male, 49% Female	11% Hispanic, 16% AI, .1% Asian, 0.03% AA, 0% PI, 41% White, 19% Multiracial
4	ELA 5-6	Paden	5+	56% Male, 44% Female	0.06% Hispanic, .2% AI, 0.02% Asian, 0.04% AA, 0% PI, 51% White, 16% Multiracial
5	ELA 5-6	Tulsa Public Schools	15	51% Male, 49% Female	38% Hispanic, 0.04% AI, 0.02% Asian, 22% AA, 0.01% PI, 21% White, 11% Multiracial
6	ELA 5-6	Edmond Schools	1	52% Male, 48% Female	13% Hispanic, 0.02% AI, 0.05% Asian, 11% AA, 0% PI, 57% White, 12% Multiracial
7	ELA 5-6	Hilldale Public Schools	19	51% Male, 49% Female	.1% Hispanic, .3% AI, 0.01% Asian, 0.02% AA, 0% PI, 39% White, 17% Multiracial
8	ELA 5-6	Putnam City Schools	3	51% Male, 49% Female	39% Hispanic, 0.02% AI, 0.04% Asian, 24% AA, 0% PI, 21% White, 11% Multiracial
9	ELA 7-8	John Rex Charter School	5	.5% Male, .5% Female	29% Hispanic, 0.02% AI, 0.04% Asian, 17% AA, 0% PI, 35% White, 14% Multiracial
10	ELA 7-8	Elk City	4	51% Male, 49% Female	24% Hispanic, 0.03% AI, 0.01% Asian, 0.04% AA, 0% PI, .6% White, 0.08% Multiracial

Table 8. OK OSTP Mathematics Articulation Panel Participant List

Panelist #	Standard Setting Panel	District	Years Teaching Experience	District Gender Breakdown	District Ethnicity Breakdown
1	Mathematics 3-4	Bridge Creek	24	52% Male, 48% Female	14% Hispanic, 0.05% AI, 0% Asian, 0.01% AA, 0% PI, 66% White, 14% Multiracial
2	Mathematics 3-4	Keystone	6	54% Male, 46% Female	0.04% Hispanic, 12% AI, 0% Asian, 0% AA, 0% PI, 66% White, 18% Multiracial
3	Mathematics 3-4	Moore Public Schools	15	51% Male, 49% Female	23% Hispanic, 0.04% AI, 0.05% Asian, 0.08% AA, 0% PI, 43% White, 17% Multiracial
4	Mathematics 3-4	Bartlesville Public Schools	3	52% Male, 48% Female	13% Hispanic, .1% AI, 0.02% Asian, 0.03% AA, 0% PI, 52% White, .2% Multiracial
5	Mathematics 5-6	Weatherford Public Schools	30	53% Male, 47% Female	23% Hispanic, 0.06% AI, 0.01% Asian, 0% AA, 0% PI, 59% White, 11% Multiracial
6	Mathematics 5-6	Shawnee Public Schools	16	52% Male, 48% Female	14% Hispanic, 12% AI, 0% Asian, 0.05% AA, 0% PI, 44% White, 25% Multiracial
7	Mathematics 5-6	Owasso	20	52% Male, 48% Female	15% Hispanic, 0.07% AI, 0.06% Asian, 0.04% AA, 0% PI, 53% White, 16% Multiracial
8	Mathematics 5-6	Oklahoma City Public Schools	22	51% Male, 49% Female	57% Hispanic, 0.02% AI, 0.02% Asian, .2% AA, 0% PI, 11% White, 0.08% Multiracial
9	Mathematics 7-8	Stilwell	33	51% Male, 49% Female	18% Hispanic, 47% AI, 0.02% Asian, 0% AA, 0% PI, 18% White, 16% Multiracial
10	Mathematics 7-8	Broken Arrow Public Schools	5	51% Male, 49% Female	19% Hispanic, 0.07% AI, 0.04% Asian, 0.07% AA, 0% PI, 49% White, 14% Multiracial
11	Mathematics 7-8	Ada City School	8	51% Male, 49% Female	15% Hispanic, 21% AI, 0.01% Asian, 0.02% AA, 0% PI, 37% White, 24% Multiracial
12	Mathematics 7-8	Stillwater	13	52% Male, 48% Female	13% Hispanic, 0.05% AI, 0.04% Asian, 0.06% AA, 0% PI, 58% White, 13% Multiracial

APPENDIX—G
MEETING AGENDA

Oklahoma OSTP Standard Setting

Meeting Agenda | June 17–21, 2024 | ELA/Mathematics Grades 3–8

Day 1: Monday, June 17

Time	Agenda Item	Activities
07:30 – 08:30	Breakfast	Registration & Check In
08:30 – 10:00	Orientation Session: Welcome & Overview	OSDE & Cognia introductions; Overview of meeting goals, OSTP ELA/Mathematics assessments, standard setting, and the ID Matching method.
10:00 – 10:15	Break & transition to breakout rooms	
10:15 – 12:00	Breakout sessions: Welcome & Overview	Facilitator and panelist introductions, meeting norms, and experience the test
12:00 – 01:00	Lunch	
01:00 – 02:30	Familiarization with OSTP assessment for grades 4, 6, or 8 as assigned.	Review & discuss standards and Performance Level Descriptors (PLDs)
02:30 – 03:15	Key concepts/processes, training & practice	Training on ID Matching method and the ordered item booklet (OIB)
03:15 – 03:30	Break	
03:30 – 04:15	Key concepts/processes, training & practice	Practice: Facilitator models ID-Matching judgmental task; Panelists practice and discussion; Prepare for Round 1
04:15 – 05:00	Round 1 Judgements	Begin round 1 (grades 4, 6, or 8 as assigned).
05:00	Adjourn for the day	

Day 2: Tuesday, June 18

Time	Agenda Item	Activities
07:30 – 08:30	Breakfast	After breakfast, convene in breakout rooms
08:30 – 09:15	Debrief Day 1	Check-in on the process, challenges, etc.
09:15 – 12:00	Complete Round 1	Complete round 1 (grades 4, 6, 8 as assigned).
10:00	Break	*Panelists take breaks as needed while working
12:00 – 01:00	Lunch	
01:00 – 02:30	Discussion and preparation for Round 2	Discuss round 1 feedback/results; Introduce benchmarks; Prepare for round 2.
02:30 – 05:00	Begin Round 2	Begin round 2 (grades 4, 6, or 8 as assigned).
03:15*	Break*	*Panelists take breaks as needed while working
05:00	Adjourn for the day	

Day 3: Wednesday, June 19

Time	Agenda Item	Activities
07:30 – 08:30	Breakfast	After breakfast, convene in breakout rooms
08:30 – 09:00	Debrief Day 2	Check-in on the process, challenges, etc.
09:00 – 10:00	Complete Round 2	Complete round 2 (grades 4, 6, or 8 as assigned). Panelists take breaks as needed.
10:00 – 10:15	Break	
10:15 – 11:00	Discussion & preparation for Round 3	Discuss round 2 feedback/results; Prepare for round 3.
11:00 – 12:00	Complete Round 3	Complete round 3 (grades 4, 6, or 8 as assigned).
12:00 – 01:00	Lunch	
01:00 – 02:30	Familiarization with OSTP assessment for grades 3, 5, or 7 as assigned.	Review & discuss standards and Performance Level Descriptors (PLDs)
02:30 – 05:00	Round 1 Judgements	Begin round 1 (grades 3, 5, or 7 as assigned).
03:15*	Break*	*Panelists take breaks as needed while working
05:00	Adjourn for the day	

Day 4: Thursday, June 20

Time	Agenda Item	Activities
07:30 – 08:30	Breakfast	After breakfast, convene in breakout rooms
08:30 – 09:00	Debrief Day 3	Check-in on the process, challenges, etc.
09:00 – 10:45	Round 1 Judgements (continuation)	Complete round 1 (grades 3, 5, or 7 as assigned).
10:00*	Break*	*Panelists take breaks as needed while working
10:45 – 12:00	Discussion & Preparation for Round 2	Discuss round 1 feedback/results; Introduce benchmarks; Prepare for round 2.
12:00 – 01:00	Lunch	
01:00 – 02:30	Round 2 Judgements	Complete round 2 (grades 3, 5, or 7 as assigned).
02:30 – 03:30	Discussion & preparation for Round 3	Discuss round 2 feedback/results; Prepare for round 3.
03:15*	Break*	*Panelists take breaks as needed while working
03:30 – 04:30	Round 3 Judgements	Complete round 3 (grades 3, 5, or 7 as assigned).
04:30 – 05:00	Wrap up and evaluation Survey	Review results for both grades, and complete final evaluation survey
05:00	*Adjourn	

**Adjourn for standard setting panelists. Panelists selected to stay for the Articulation meeting will reconvene in the morning.*

Vertical Articulation Meeting

Day 5: Friday, June 21

Time	Agenda Item	Activities
07:30 – 08:30	Breakfast	
08:30 – 10:00	Vertical Articulation	Key concepts/processes and training; complete readiness survey; start articulation process
10:00 – 10:15	Break	
10:15 – 12:00	Vertical articulation	Continuation
12:00 – 12:30	Wrap up and Evaluation Survey	
12:30	Adjourn	To go lunch

Terminology Reference

During the standard-setting meeting, acronyms or terms will be introduced and defined as it becomes relevant. A list of the most used acronyms and terms, along with brief descriptions, is presented below for quick reference.

Acronym / Term	Brief Description
Cut Score	The minimum test score a student must earn to be considered at a specific performance level. Three cut scores result in four levels of performance.
ID Matching	Item-Descriptor Matching: An item-centered, content-based method for standard setting
KSAs	Knowledge, Skills, and Abilities.
OAS	Oklahoma Academic Standards
OIB	Ordered Item Booklet: A set of test items ordered by item difficulty (content and grade specific).
OSDE	Oklahoma State Department of Education
OSTP	Oklahoma School Testing Program
Performance Levels	Reflect the specific knowledge and skills that a student should be able to demonstrate based on their performance on the test. OSTP has four performance levels: Below basic, basic, proficient, and advanced.
PLDs	Performance Level Descriptors: A narrative account of the knowledge, skills, and abilities demonstrated by students in each level of performance. Describe what students know and can do based on the Oklahoma Academic Standards. (Content and grade specific)

APPENDIX—H
NON-DISCLOSURE AGREEMENT

Nondisclosure Agreement
Oklahoma State Testing Program
Standard Setting
June 17-21, 2024

The undersigned is an employee, contractor, assessment committee member, or person otherwise authorized to view secure state assessment materials. The undersigned hereby agrees to be bound by the terms of this agreement restricting the disclosure of said materials.

It is essential to the integrity of this item development project and testing program that all test items remain secure. To maintain this security, only authorized persons are permitted to view the test questions. With the exception of materials released by the Oklahoma State Department of Education for informational purposes, all test questions (draft or final) in hardcopy or electronic format and associated materials must be regarded as secure documents. As a result, such materials may not be reproduced, electronically transmitted, discussed, used in classroom instruction, or in any way released or distributed to unauthorized persons. All materials including items and item drafts must be returned at the end of the meeting.



I understand that I am responsible for test materials security. By breaching test materials security as described here, I am breaching professional testing ethics and may be subject to additional penalties under law.

Name: _____

Signature: _____

Date: _____

APPENDIX—I
ORIENTATION POWERPOINT PRESENTATION




OSTP ELA/Math Grades 3-8

Standard Setting Orientation
June 17 – 21, 2024


© 2024 Cognia, Inc.

1



Orientation Session - Agenda

- Introduction of the Standard Setting Team
- OSDE: Welcome
- Standard Setting Goals and Outcomes
- Overview of the OSTP ELA/Math Assessments
 - Test Design
 - Performance Level Descriptors
- Overview of Key Concepts and Procedures
- Transition to Breakout Rooms



2

Standard Setting Team

Oklahoma SDE Members

- Catherine Boomer, Program Director, State Assessments
- Samantha Sheppard, Project Manager, Science
- Caroline Misner, Project Manager, OAAP
- Alyssa Tyra, Project Manager, ELA Assessments
- Corinne Beasler, Project Manager, Math Assessments
- Sharon Morgan, Program Director, Standards & Learning
- Jason Stephenson- Project Manager, Secondary ELA
- Deann Jones- Project Director, RSA
- Rori Hodges, Specialist, Early Childhood



3

Standard Setting Team - Cognia

Program Management

- Elizabeth Garcia
- Sharman Lyons (Events team)

Content Specialists

- Breanne Moore Math
- Mary Kate Clauson ELA

Psychometricians

- Sandra Sweeney
- Frank Padellaro
- Qi Qin

Facilitation Team

- Karen Whisler Math 3-4
- Katie Schmidt Math 5-6
- Jill Stepanek Math 7-8
- Jessica Keymer ELA 3-4
- Lisa Jones Kennedy ELA 5-6
- Rebecca Young ELA 7-8



4

Standard Setting Team – Outside Observers

- Erika Landl, Center for Assessment, OSTP Technical Advisory Committee Member
- Maria Elena Oliveri, Purdue University, OSTP Technical Advisory Committee Member
- Eric Jones, Administrative Programs Manager, Office of Educational Quality & Accountability



5

Housekeeping

- Reimbursement form:
 - Fill out completely
 - For those staying overnight provide itemized receipts for dinner
- W9 form:
 - Anyone receiving a stipend of \$600 or more must fill out a W9 form. If you do fill out and return, your reimbursement will not be processed.
 - Please complete the W9 form today and give to your facilitator to turn in at the end of the day. This will speed up the process of your reimbursement.



6

Assessment History

- In 2016, the Oklahoma Legislature directed the State Board of Education to evaluate Oklahoma's current state assessment system and make recommendations for its future.
- As a result, the Oklahoma State Department of Education
 - Held regional meetings across the state to determine stakeholder concerns
 - Convened the Oklahoma Assessment & Accountability Task Force to develop recommendations
 - Followed federal requirements and rules as described in ESSA.

7 |



7

Goal for Oklahoma Schools

- Focus on college- and career-readiness:
 - *College and career ready means that students graduate from high school prepared to enter and succeed in postsecondary opportunities whether college or career.*
- One measurement of college- and career readiness is the Oklahoma School Testing Program.

8 |



8

Oklahoma Statute on Performance Levels

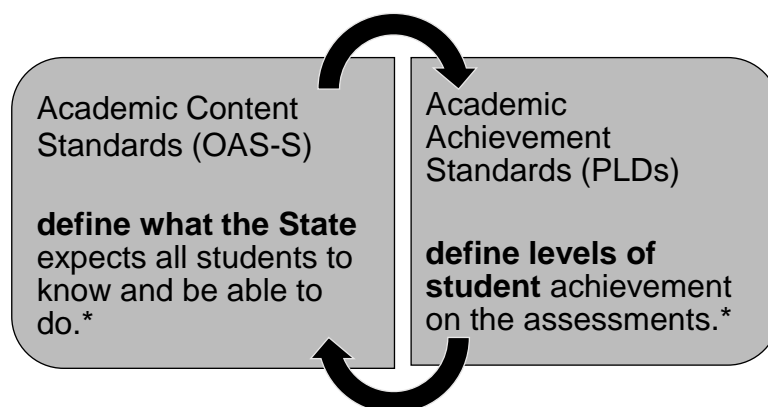
- OSTP Performance is divided into performance levels.
- The Performance levels shall be set by a method that indicates students are ready for the next grade, course, or level of education, as applicable.
- The Commission for Educational Quality and Accountability (CEQA) shall determine and adopt a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.
- §70-1210.541

9 |



9

Content Standards and PLDs



**U.S. Department of Education Peer Review of State Assessment Systems Non-Regulatory Guidance for States, September 25, 2015*

10 |



10



Standard Setting Goals

Our shared goals

- Use your judgments to help provide performance standards recommendations for the OSTP ELA/Math assessments that provide meaningful and actionable information

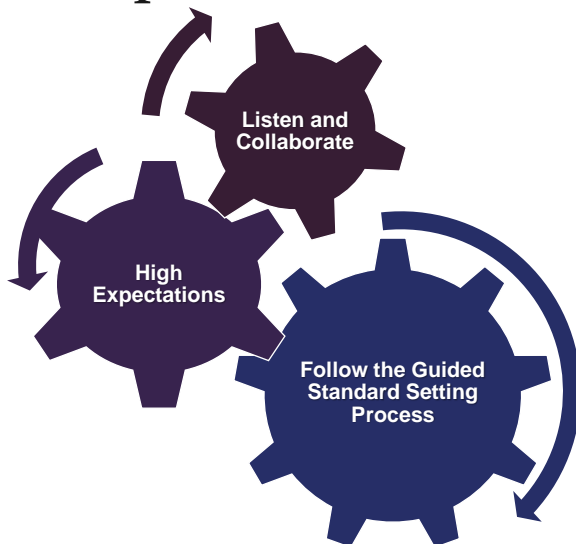
Your goals as panelists

- Learn concepts and procedures following the Item-Descriptor (ID) Matching standard setting method
- Follow the procedures to complete the standard setting activities
- Rely on your expertise about the content standards, student learning, and students throughout the process



11

Expectations of all Panelists



- Security is of the utmost importance
 - You can discuss the process in general terms
- You may NOT
 - Share details about the items or specific details about the process (e.g., cuts that were recommended)
 - Use your phones or personal devices while in the room
 - Use the Chromebooks for anything other than standard setting activities



12

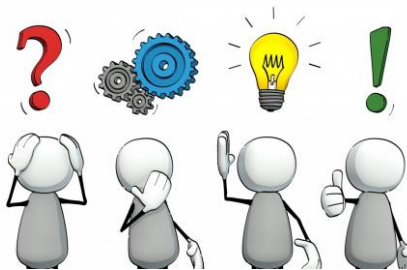
A Shift in Focus for this Week



OTHER WAYS YOU HAVE CONTRIBUTED

- Item writing, data review, content review and/or item review committees
- Review test items
- Purpose: Evaluate items for use on a test (potential problems with the items; suggest improvements)

THERE IS A
DIFFERENCE



THE WORK WE ARE DOING THIS WEEK

- Standard setting: Item-centered method with content-based judgement
- Look at test items
- Purpose: Identify the knowledge, skills, and abilities required to correctly answer the item



13

Purpose of Standard Setting

- Allows Oklahoma State Department of Education (OSDE) to have educator expertise inform *performance standards* for the OSTP ELA/Math assessments:
- Opportunity for educator input on *cut scores* used to define *performance levels*
- To ensure recommendations are consistent with expectations stated in the *Performance Level Descriptors*



14

Performance Levels

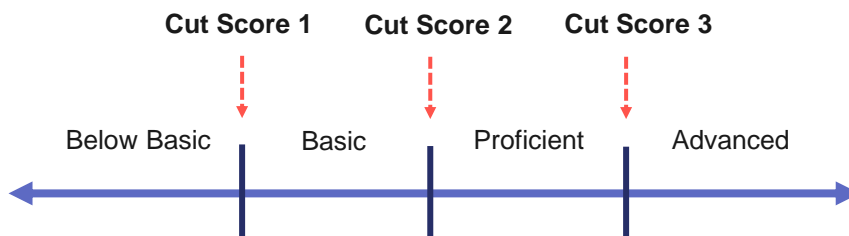
- Performance Levels reflect the specific knowledge and skills that a student should be able to demonstrate based on their performance on the test.
- The Oklahoma School Testing Program (OSTP) has four performance levels.



15

Cut Scores

- A cut score is the minimum test score a student must earn to be considered at a specific performance level.
- Three cut scores result in four levels of performance.



16

Cut Score Considerations

- We don't rely on percentages.
 - They are arbitrary and don't consider the content.
- We use content-based judgment.
 - Content links assessment items, performance level descriptors (PLDs), and the Oklahoma Academic Standards (OAS).



17



Math Test Design

- Each math test has 50 Operational items and 10 Field Test items.
- The 50 operational items must match the blueprint which is broken down by the four math strands, which correspond to the four math reporting categories.

Grade	Number & Operations	Algebraic Reasoning & Algebra	Geometry & Measurement	Data & Probability
3	44 – 48%	12 – 18%	22 – 26%	12 – 18%
4	42 – 46%	12 – 18%	24 – 28%	12 – 18%
5	42 – 46%	14 – 20%	22 – 26%	12 – 18%
6	38 – 42%	20 – 24%	22 – 26%	12 – 16%
7	16 – 20%	26 – 30%	30 – 36%	18 – 24%
8	16 – 20%	44 – 48%	18 – 22%	14 – 18%



18

Math Depth of Knowledge (DOK)

SUBJECT	SUMMARY DEFINITIONS OF DEPTH OF KNOWLEDGE (WEBB'S DOK™)			
	WEBB'S DOK LEVEL 1	WEBB'S DOK LEVEL 2	WEBB'S DOK LEVEL 3	WEBB'S DOK LEVEL 4
Mathematics	<p>Requires students to recall or observe facts, definitions, and terms. Includes simple one-step procedures. Includes computing simple algorithms (e.g., sum, quotient).</p> <p>Examples:</p> <ul style="list-style-type: none"> Recall or recognize a fact, term, or property. Represent in words, pictures, or symbols a math object or relationship Perform a routine procedure, such as measuring At higher grades, solve a quadratic equation or a system of two linear equations with two unknowns 	<p>Requires students to make decisions on how to approach a problem. Requires students to compare, classify, organize, estimate, or order data. Often involves procedures with two or more steps.</p> <p>Examples:</p> <ul style="list-style-type: none"> Specify and explain relationships between facts, terms, properties, or operations Select procedure according to criteria and perform it Use concepts to solve routine multiple-step problems. 	<p>Requires reasoning, planning, or use of evidence to solve a problem or algorithm. May involve an activity with more than one possible answer. Requires conjecture or restructuring of problems. Involves drawing conclusions from observations, citing evidence and developing logical arguments for concepts. Uses concepts to solve non-routine problems.</p> <p>Examples:</p> <ul style="list-style-type: none"> Formulate original problem, given situation Formulate mathematical model for complex situation Produce a sound and valid mathematical argument Devise an original proof Critique a mathematical argument 	<p>Requires complexity at least at the level of DOK 3 but also an extended time to complete the task. A project that requires extended time but repetitive or lower-DOK tasks is not at Level 4. Requires complex reasoning, planning, developing, and thinking. May require students to make several connections and apply one approach among many to solve the problem. May involve complex restructuring of data, establishing and evaluating criteria to solve problems.</p> <p>Examples:</p> <ul style="list-style-type: none"> Apply a mathematical model to illuminate a problem, situation Conduct a project that specifies a problem, identifies solution paths, solves the problem, and reports results Design a mathematical model to inform and solve a practical or abstract situation

19

Math DOK Blueprint

- The 50 operational items must match the blueprint which is broken down by the three DOK levels.

Grade	DOK 1	DOK 2	DOK 3
3	40 – 50%	45 – 55%	5 – 10%
4	20 – 30%	65 – 75%	5 – 15%
5	20 – 30%	65 – 75%	5 – 15%
6	15 – 25%	65 – 75%	10 – 20%
7	15 – 25%	65 – 75%	10 – 20%
8	10 – 20%	65 – 75%	15 – 25%

20



ELA Test Design

- Each ELA test has 50 Operational items and 10 Field Test items.
- The 50 operational items must match the blueprint which is broken down by the five assessed ELA standards, which correspond to the five ELA reporting categories.

Grade	Reading & Writing Process	Critical Reading & Writing	Vocabulary	Language	Research
3	38 – 42 %	12 – 18 %	22 – 26 %	12 – 18 %	12 – 18 %
4	30 – 34 %	18 – 22 %	22 – 26 %	12 – 18 %	12 – 18 %
5	30 – 34 %	22 – 26 %*	18 – 22 %	12 – 18 %	12 – 18 %
6	34 – 38 %	18 – 22 %	18 – 22 %	12 – 18 %	12 – 18 %
7	34 – 38 %	18 – 22 %	14 – 20 %	12 – 18 %	14 – 20 %
8	24 – 30 %	24 – 30 %*	14 – 20 %	12 – 18 %	12 – 18 %



21



ELA Stimulus

- Stimuli consist of authentic literature or are commissioned specifically for OAS.
- They represent topics and genres appropriate for each grade.
- Qualitative and quantitative measures

Grade	Word Count*	Authentic Literary Selections	Expository Selections
3	200 - 600	3 – 6	3 – 5
4	200 – 600	4 – 6	3 – 5
5	300 – 700	4 – 6	4 – 6
6	300 – 700	4 – 6	4 – 6
7	500 – 900	4 – 6	4 – 6
8	500 – 900	4 – 6	4 – 6



22



ELA Depth of Knowledge (DOK)

SUBJECT	SUMMARY DEFINITIONS OF DEPTH OF KNOWLEDGE (DOK)			
	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
English Language Arts	<p>Requires students to recall, observe, question, or represent facts, simple skills, or abilities. Requires only surface understanding of text, often verbatim recall.</p> <p>Examples:</p> <ul style="list-style-type: none"> Support ideas by reference to verbatim (or only slightly paraphrased) details in text Use a dictionary to find meanings of words Recognize figurative language in a passage Identify correct spelling or meaning of words 	<p>Requires processing beyond recall and observation. Requires both comprehension and subsequent processing of text or portions of text. Involves ordering, classifying text as well as identifying patterns, relationships, and main points.</p> <p>Examples:</p> <ul style="list-style-type: none"> Use context to identify unfamiliar words Predict a logical outcome Identify and summarize main points Apply knowledge of conventions of standard American English Compose accurate summaries of the major events in a narrative 	<p>Requires students to go beyond text. Requires students to explain, generalize, and connect ideas. Involves deep inferencing, prediction, elaboration, and summary. Requires students to support positions using prior knowledge and evidence and to manipulate themes across passages.</p> <p>Examples:</p> <ul style="list-style-type: none"> Determine effect of author's purpose on text elements Summarize information from multiple sources Critically analyze literature Compose focused, organized, coherent, purposeful prose Evaluate the internal logic or credibility of a message 	<p>Requires complexity at least at the level of DOK 3 but also an extended time to complete the task, such as conducting a research project over many weeks. A project that requires extended time but repetitive or lower-DOK tasks is not at Level 4. May require generating hypotheses and performing complex analyses and connections among texts.</p> <p>Examples:</p> <ul style="list-style-type: none"> Analyze and synthesize information from multiple sources Examine and explain alternative perspectives across sources Describe and illustrate common themes across a variety of texts Create compositions that synthesize, analyze, and evaluate

Revised 2014



WebbAlign (c) 2016. All Rights Reserved. WebbAlign offers alignment studies and professional development on Webb's Depth of Knowledge. Please contact us at contracts@wceps.org or 877-249-4211 for more information.

ELA ➡ Depth of Knowledge (DOK)

- The 50 operational items must match the blueprint which is broken down by the three DOK levels.

Grade	DOK 1	DOK 2	DOK 3
3	15-30%	65-80%	5-10%
4	10-20%	65-75%	5-15%
5	5-15%	70-85%	5-20%
6	5-15%	70-85%	10-20%
7	5-15%	70-85%	10-20%
8	5-10%	60-75%	20-30%



Math and ELA Item Types

Math Item Types

- Multiple Choice
- Cluster Multiple Choice Items with a Shared Stimulus
- Technology Enhanced Items (TEIs)
- Paper Equivalent Items for TEIs

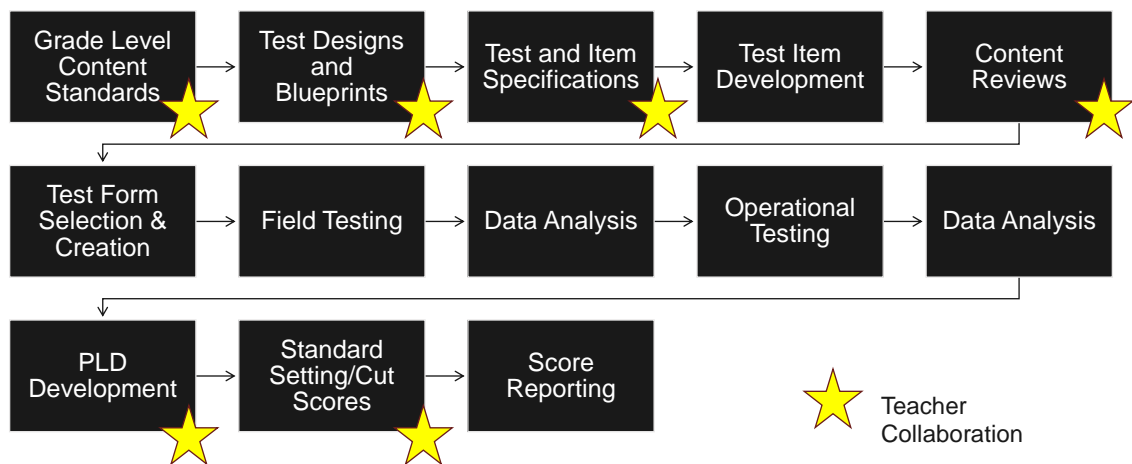
ELA Item Types

- Multiple Choice
- Cluster Multiple Choice Items with a Shared Stimulus
- Technology Enhanced Items (TEIs)
- Paper Equivalent Items for TEIs
- Constructed Response
- Writing Prompt



25

OK Test Development Cycle



26



Background on PLD development

- New standards were adopted by OSDE. As a result, the PLDs needed to be updated so that they accurately reflect what students know and can do at each performance level.
- OSDE and Cognia staff collaborated on the development of new PLDs using the updated standards as a foundation.



29

Background on the PLD Development

- Teacher committees reviewed and discussed draft PLDs. After this discussion, OSDE finalized the PLDs.
- This week, the new PLDs will be used to complete the standard setting activities that will result in cut score recommendations for the OSTP ELA and Math assessments.



30

Language for Math & ELA Policy PLDs

Below Basic	Basic	Proficient	Advanced
Students have not performed at least at the Basic level.	Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:



31

Math PLD Organization

- Math PLDs are arranged by:
 - Grade level
 - Strand (Numbers and Operations, Algebraic Reasoning & Algebra, Geometry & Measurement, and Data & Probability)
 - PLD Level (Basic, Proficient, and Advanced)
 - Objective



32

Math PLDs for Grade 5

OSTP Math Grade 5 - Range Performance Level Descriptors (Range PLDs)				
Strand	Basic	Proficient	Advanced	Objective(s)
	OK Policy PLD Basic: Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:	OK Policy PLD Proficient: Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:	OK Policy PLD Advanced: Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level. Students scoring at the Advanced level typically:	
Numbers & Operations	Represent decimal fractions with a model.			5.N.1.1
	Recognize and generate equivalent decimals, fractions, and mixed numbers and represent whole numbers.	Compare and order fractions. Compare and order decimals.	Order a mix of decimals, fractions, mixed numbers, and whole numbers.	5.N.1.2, 5.N.1.3, 5.N.1.4
	Solve division, multiplication, addition, and subtraction problems.	Estimate and solve division problems with the remainder represented as a fraction, decimal, or whole number.	Interpret the remainder of division problems within the context of the problem.	5.N.2.1, 5.N.2.2, 5.N.2.3, 5.N.2.4
	Add and subtract decimals and fractions with like denominators.	Estimate, illustrate, add, and subtract fractions and mixed numbers.	Order a mix of decimals, fractions, mixed numbers, and whole numbers.	5.N.3.1, 5.N.3.2, 5.N.3.3, 5.N.3.4
Algebraic Reasoning & Algebra	Describe patterns of change. Identify the origin and axes in relation to the coordinates.	Graph patterns of change as ordered pairs on a coordinate plane. Use a rule or table to represent ordered pairs.	Make predictions and generalizations about patterns of change.	5.A.1.1, 5.A.1.2



33

ELA PLD Organization

- ELA PLDs are arranged by:
 - Grade level
 - Standard (Reading & Writing Process, Critical Reading & Writing, Vocabulary, Language and Research)
 - PLD Level (Basic, Proficient, and Advanced)
 - Objectives



34

ELA PLDs for Grade 8

OSTP ELA Grade 8 - Range Performance Level Descriptors (Range PLDs)

Objective	Basic	Proficient	Advanced
	OK Policy PLD Basic: <i>Students demonstrate partial mastery of the essential knowledge and skills appropriate to their grade level. Students scoring at the Basic level typically:</i>	OK Policy PLD Proficient: <i>Students demonstrate mastery over appropriate grade-level subject matter and readiness for the next grade level. Students scoring at the Proficient level typically:</i>	OK Policy PLD Advanced: <i>Students demonstrate superior performance on challenging subject matter. In addition to demonstrating a broad and in-depth understanding and application of all skills at the Proficient level, students scoring at the Advanced level typically:</i>
Reading & Writing Process			
8.2.R.1	Summarize an alphabetic or multimodal text to demonstrate comprehension of a text.	Summarize alphabetic and/or multimodal texts about similar topics to demonstrate comprehension within and between texts.	Summarize alphabetic and/or multimodal texts about similar topics to demonstrate comprehension within and between texts; evaluate summaries.
8.2.R.2	Identify details in fiction, poetry, and nonfiction texts to distinguish genres.	Analyze details in fiction, poetry, and nonfiction texts to identify characteristics of genres.	Analyze details in fiction, poetry, and nonfiction texts to identify characteristics of genres and provide supporting evidence.
8.2.R.3	Paraphrase a paragraph in their own words to demonstrate comprehension.	Paraphrase a portion of passage in their own words to demonstrate comprehension.	
8.2.W.1	Identify a prewriting strategy (e.g., develop ideas and plan).	Prewrite (e.g., develop ideas and plan).	Create and use a prewriting strategy.
	Minimally plan/organize ideas.	Organize and develop ideas to compose a first draft.	Organize and develop ideas related to a thesis to



35

Overview of Item-Descriptor (ID) Matching Method

Item-
centered
Method

Content-
based
Judgment

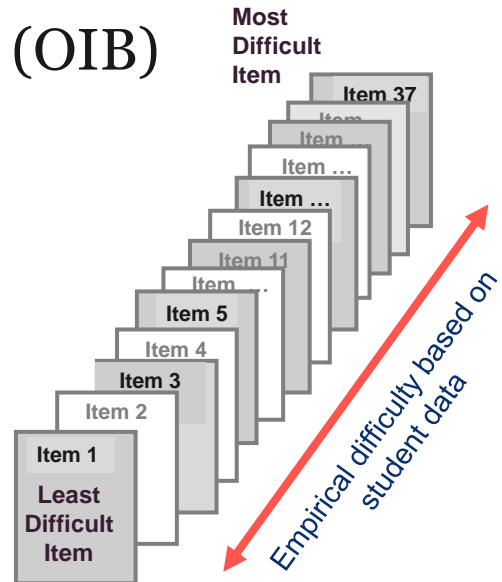
Iterative
Process



36

Ordered Item Booklet* (OIB)

- A set of test items
- One item per 'page'
- Items ascend by difficulty
 - Easiest item appears first
 - Most difficult item appears last
- Order is based on empirical item difficulties
 - Not the order in which they appear for students during the test



37

Overview of ID Matching Method

- Panelists review each item in the OIB.
 - Identify the knowledge, skills, and abilities (KSAs) required to answer the item correctly.
- For each item, make the following judgment:
 - Match the knowledge, skills, and abilities (KSAs) required by the item with the expectations described in either the Basic, Proficient, or Advanced performance level descriptor (PLD).
- Judgments are made independently
- Iterative process
 - Across three rounds (for each grade)



38

Content-based Judgments



Useful

- Based on Content
- Links items to PLDs
- Refers to specific knowledge, skills, and abilities (KSAs)



Not Useful

- Based on something other than the content
- Too general
- Based on a specific student or class



39

Content-Based Benchmarks - Overview

- Benchmarks based on Cognia and OSDE content team judgments
 - Benchmarks will be presented to you at the beginning of Round 2.
- Benchmarks serve as additional information for you to consider as you engage in the 2nd and 3rd rounds of the standard setting process.

→ More detailed information/training to come later today



40

APPENDIX—J
READINESS SURVEYS & ROUND BY ROUND RESULTS

Standard Setting Readiness Surveys

The following three tables show the survey questions and associated response options administered to panelists prior to each judgment round, which panelists used to indicate their readiness to proceed with the judgment tasks for the upcoming round.

Readiness Survey—Round 1

Question	Response Options
I understand the goals of the standard setting meeting.	Yes No
I understand the procedures we are using to set standards.	Yes No
I understand the differences between the performance levels.	Yes No
I understand how to make item-PLD alignment judgements.	Yes No
The quality of the item is important to consider when making item-PLD alignment judgments.	Agree Unsure Disagree
How important is it to consider a typical student's ability while engaging in the standard setting activities?	Not important Unsure Very important
I understand how to use the Cognia Standard Setting Toolkit.	Yes No
I am ready to proceed with the standard setting process.	Yes No

Readiness Survey—Round 2

Question	Yes	No
I understand the round 1 feedback.		
I understand that I should use the round 1 feedback as information, not persuasion, for me to consider as I make my judgements in round 2.		
I understand what the content-based benchmarks represent.		
I understand that I can use the content-based benchmarks as additional information, not persuasion, for me to consider as I make my judgements in round 2.		
I understand that I should consider the insights of my colleagues as information, but not persuasion, as I make my own independent judgments in round 2.		
I am ready to proceed with Round 2 of the standard setting process.		

Readiness Survey—Round 3

Question	Yes	No
I understand the round 2 feedback.		
I understand that I should use the round 2 feedback as information, not persuasion, for me to consider as I make my judgements in round 3.		
I understand that I should consider the insights of my colleagues as information, but not persuasion, as I make my own independent judgments in round 3.		
I am ready to proceed with Round 3 of the standard setting process		

Standard Setting Round by Round Results

The following series of figures represent the results presented to panelists after each judgment round and were used to facilitate discussions. These results were presented as frequency graphs with the ordered item booklet (OIB) page numbers on the x-axis and the number of panelists on the y-axis. The stacked bars represented the number of panelists that selected the basic (yellow), proficient (green), or advanced (blue) performance level for each item in the OIB. Since these results were calculated and presented after each judgment round, there were three figures (corresponding to rounds 1, 2, and 3, respectively) for each grade within each content area.

Figure 1. ELA Grade 3 Round 1 - Frequency of Panelist Judgments by Performance Level

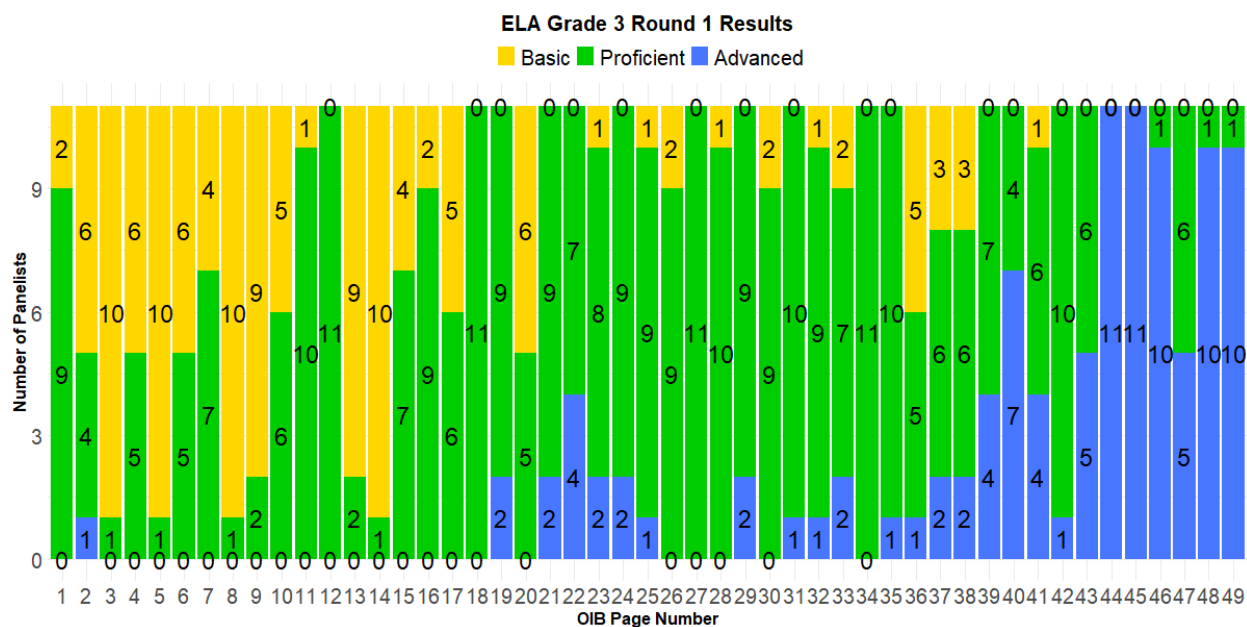


Figure 2. ELA Grade 3 Round 2 - Frequency of Panelist Judgments by Performance Level

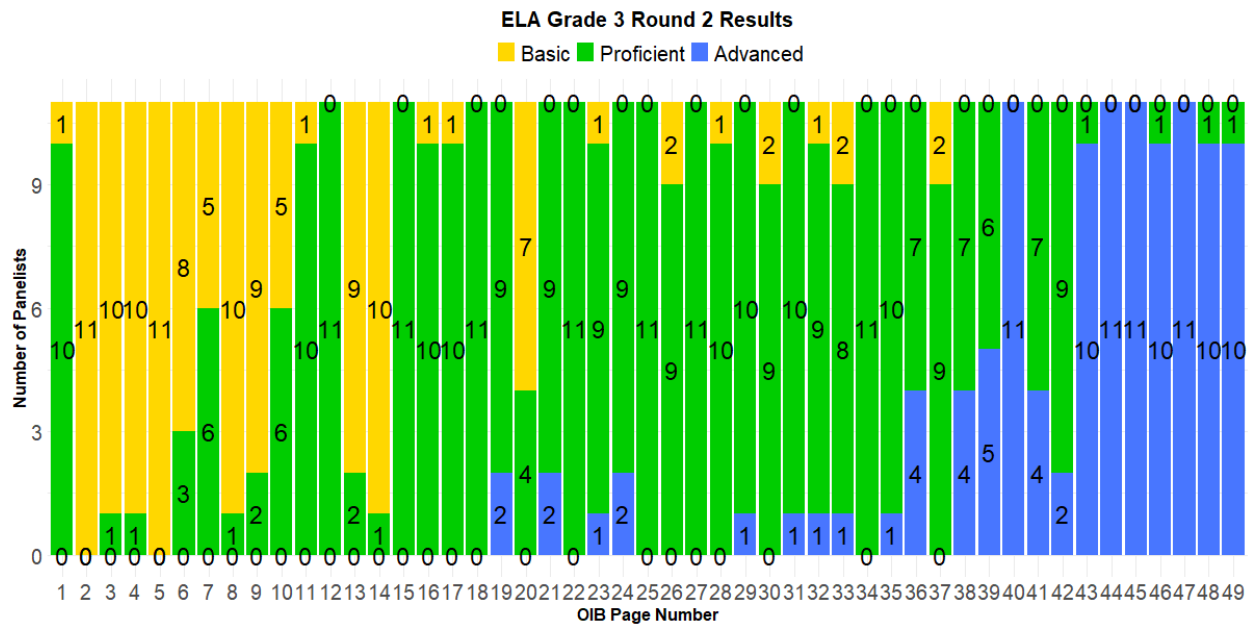


Figure 3. ELA Grade 3 Round 3 - Frequency of Panelist Judgments by Performance Level

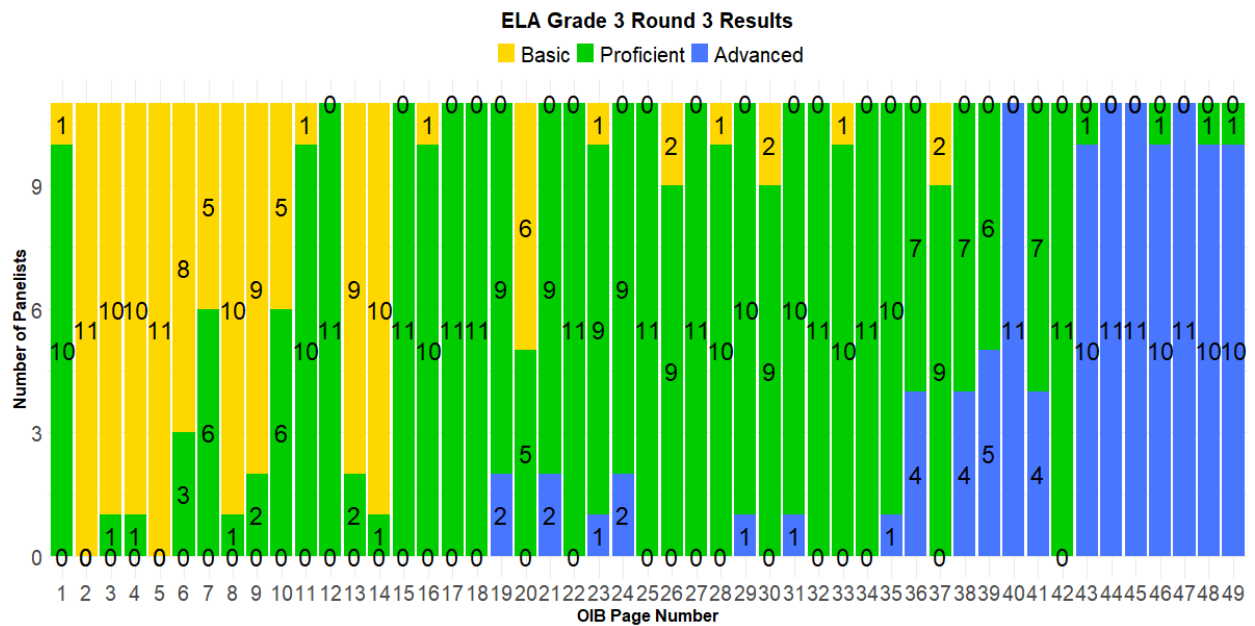


Figure 4. ELA Grade 4 Round 1 - Frequency of Panelist Judgments by Performance Level

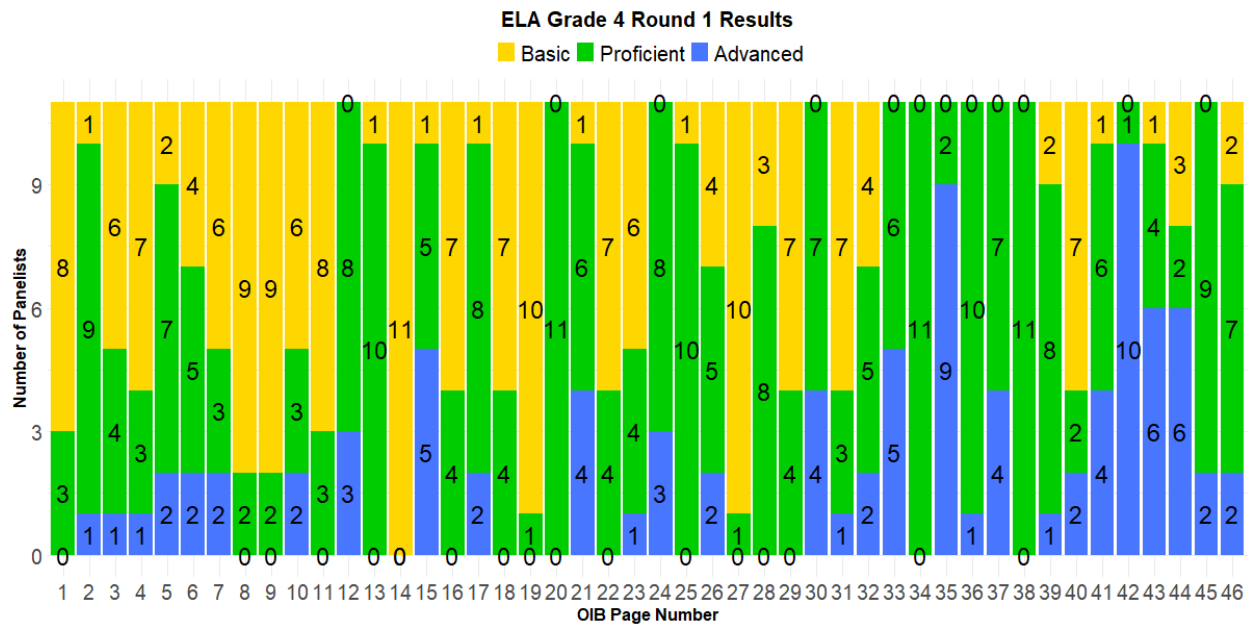


Figure 5. ELA Grade 4 Round 2 - Frequency of Panelist Judgments by Performance Level

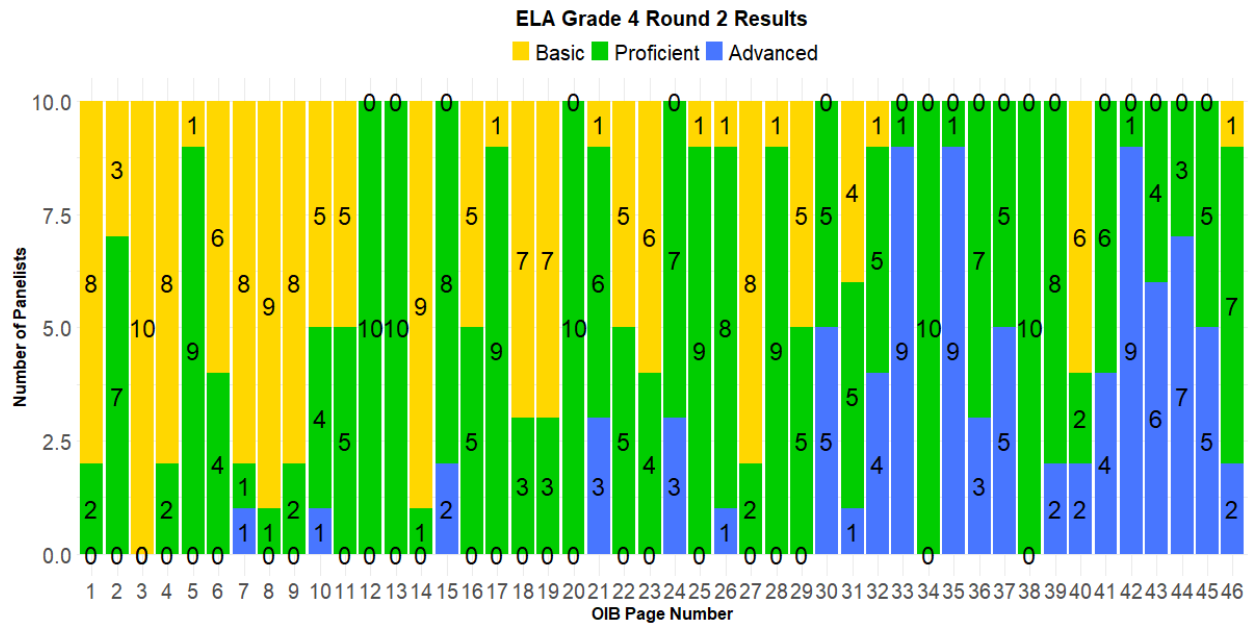


Figure 6. ELA Grade 4 Round 3 - Frequency of Panelist Judgments by Performance Level

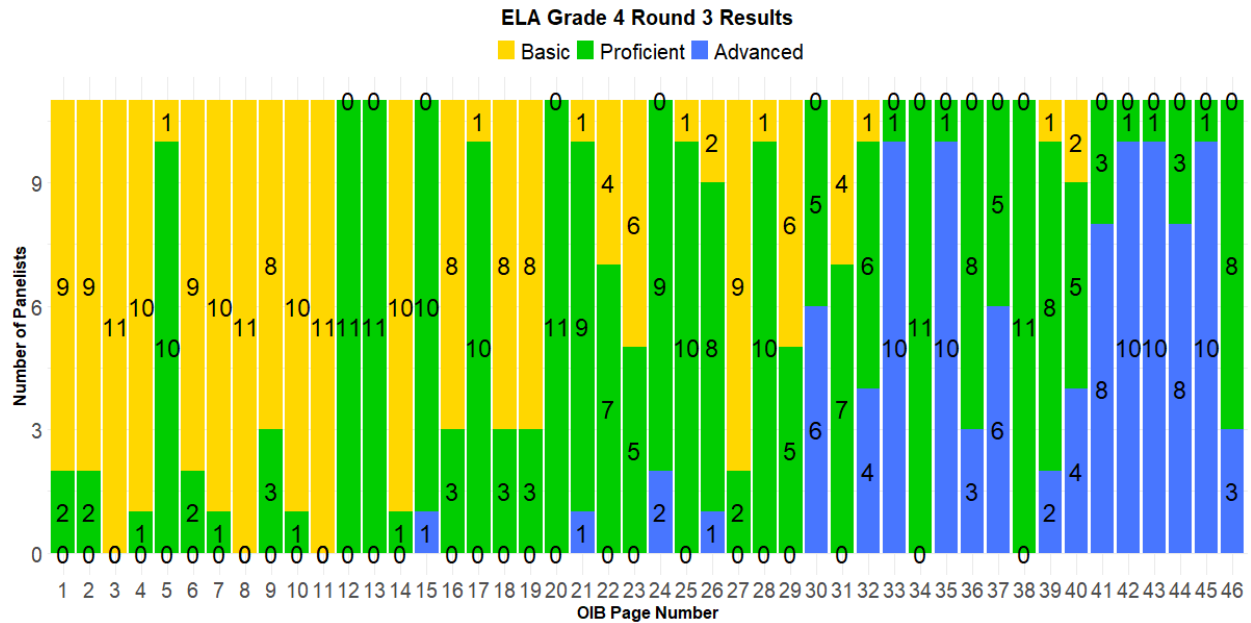


Figure 7. ELA Grade 5 Round 1 - Frequency of Panelist Judgments by Performance Level

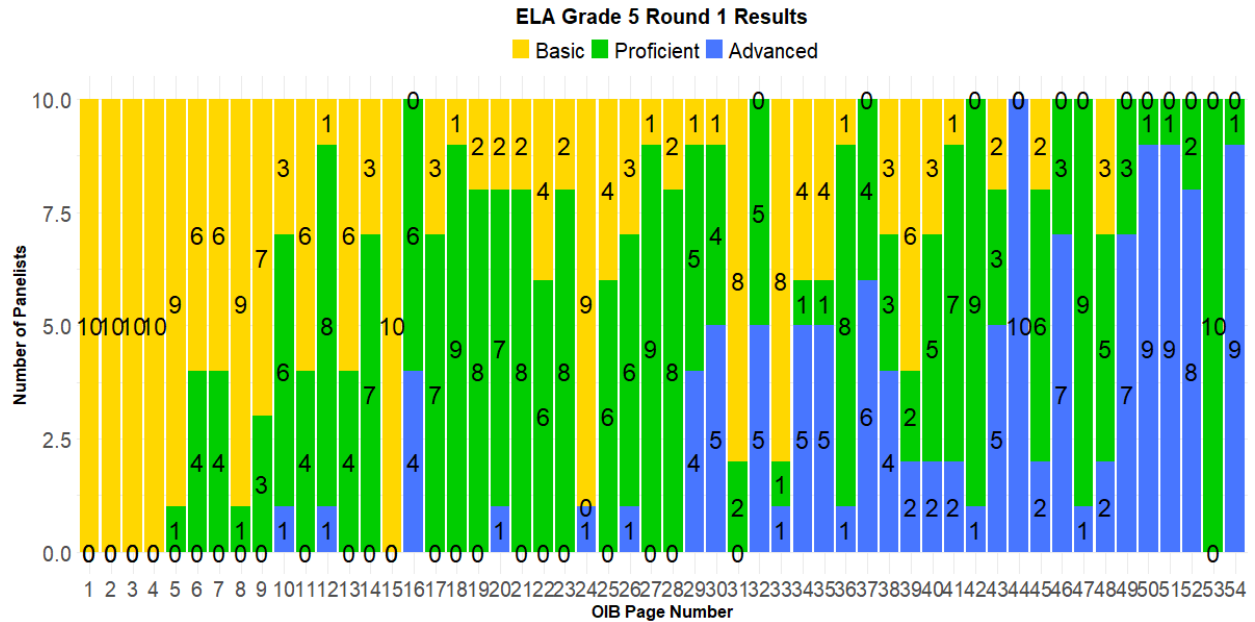


Figure 8. ELA Grade 5 Round 2 - Frequency of Panelist Judgments by Performance Level

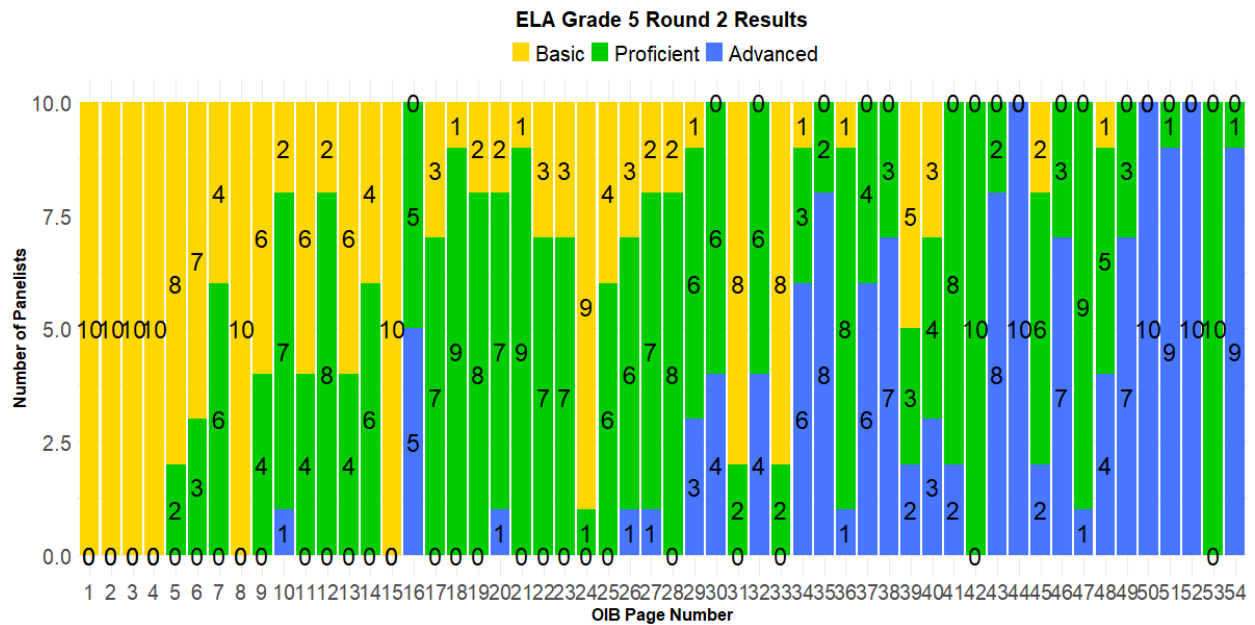


Figure 9. ELA Grade 5 Round 3 - Frequency of Panelist Judgments by Performance Level

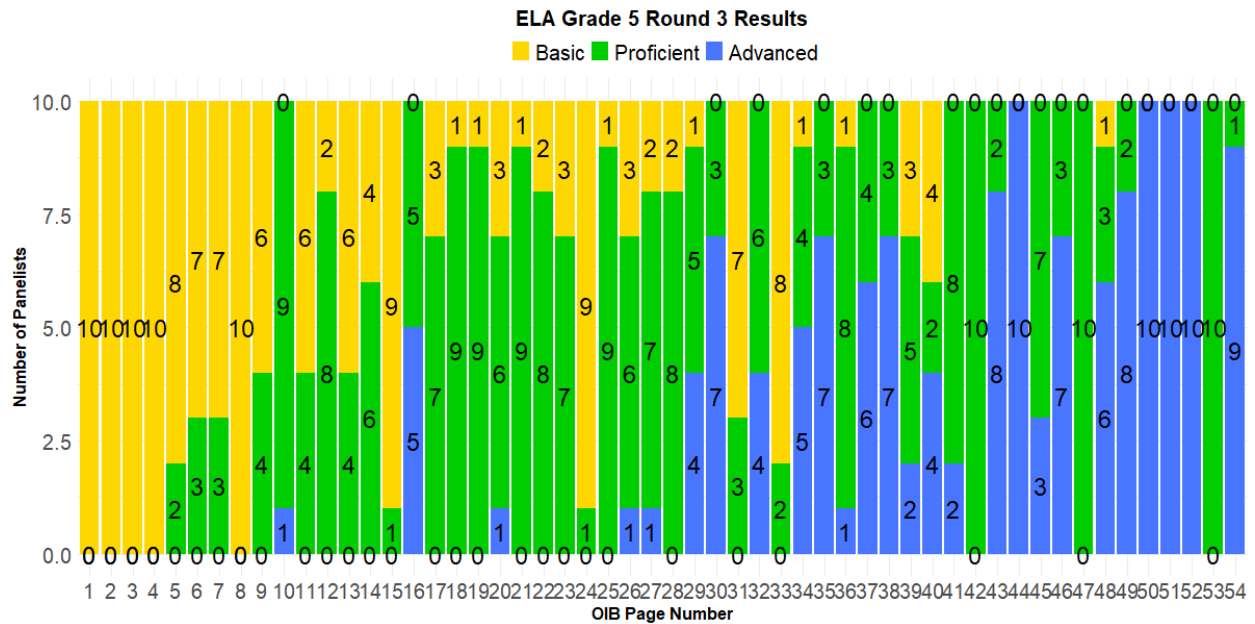


Figure 10. ELA Grade 6 Round 1 - Frequency of Panelist Judgments by Performance Level

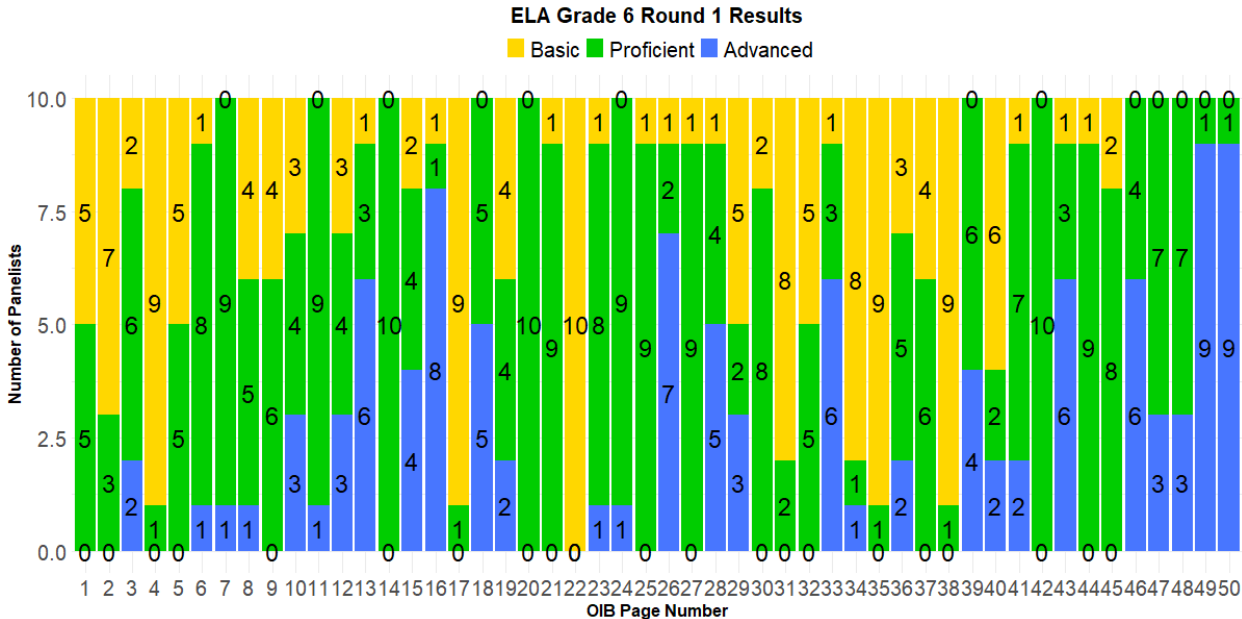


Figure 11. ELA Grade 6 Round 2 - Frequency of Panelist Judgments by Performance Level

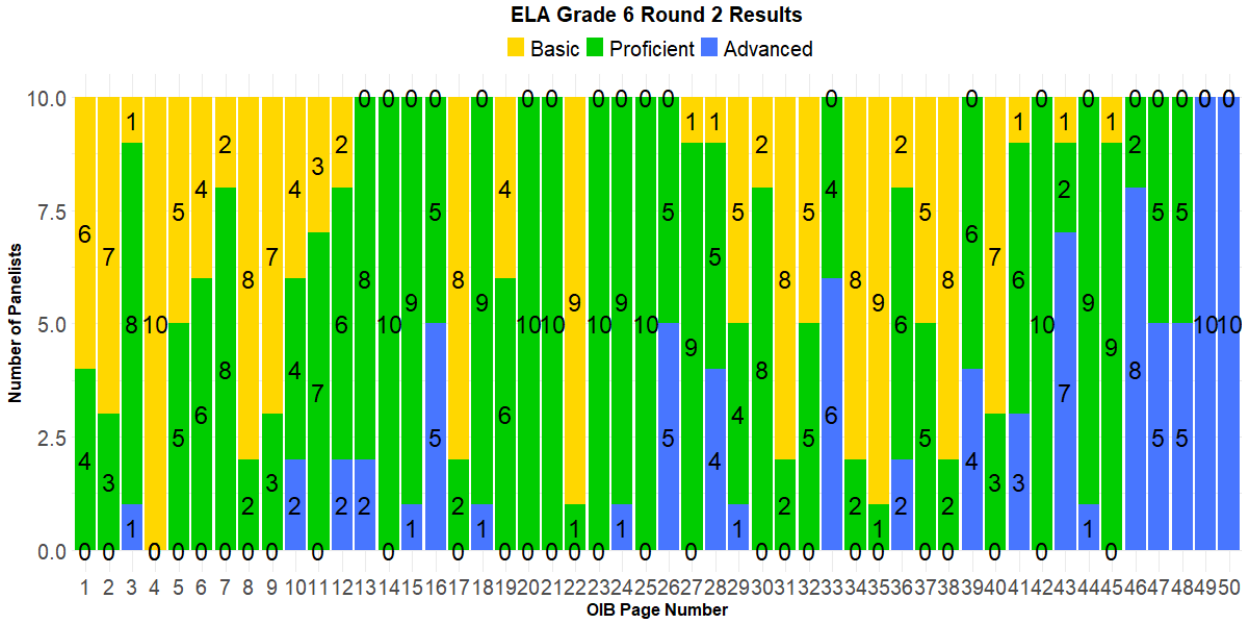


Figure 12. ELA Grade 6 Round 3 - Frequency of Panelist Judgments by Performance Level

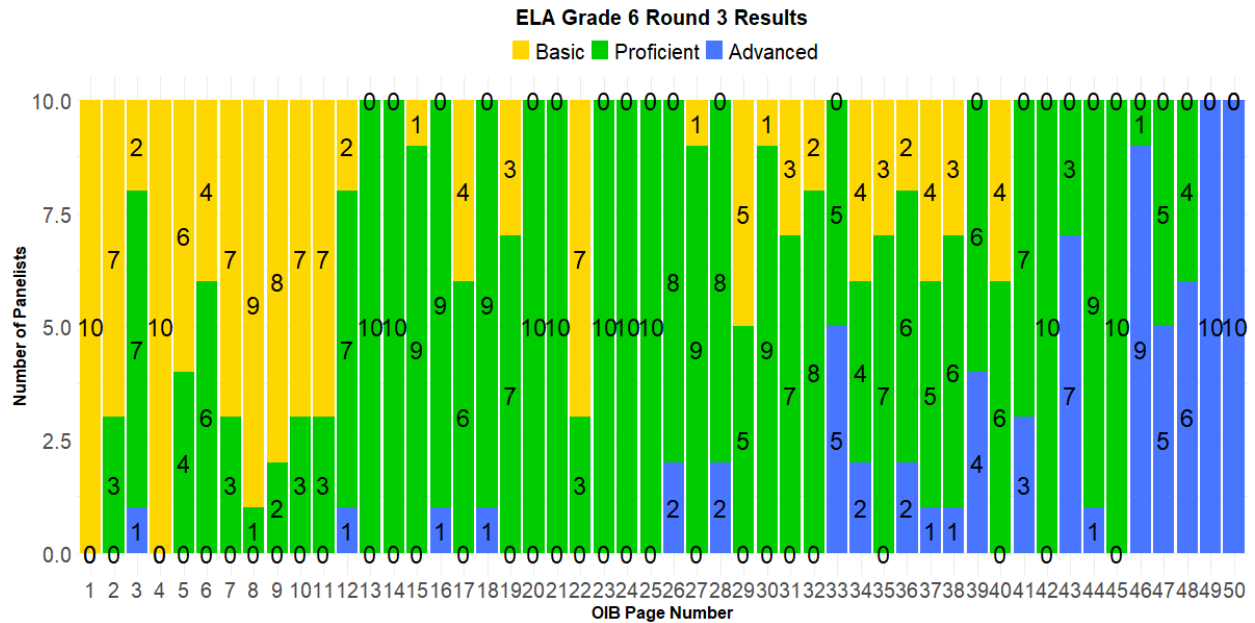


Figure 13. ELA Grade 7 Round 1 - Frequency of Panelist Judgments by Performance Level

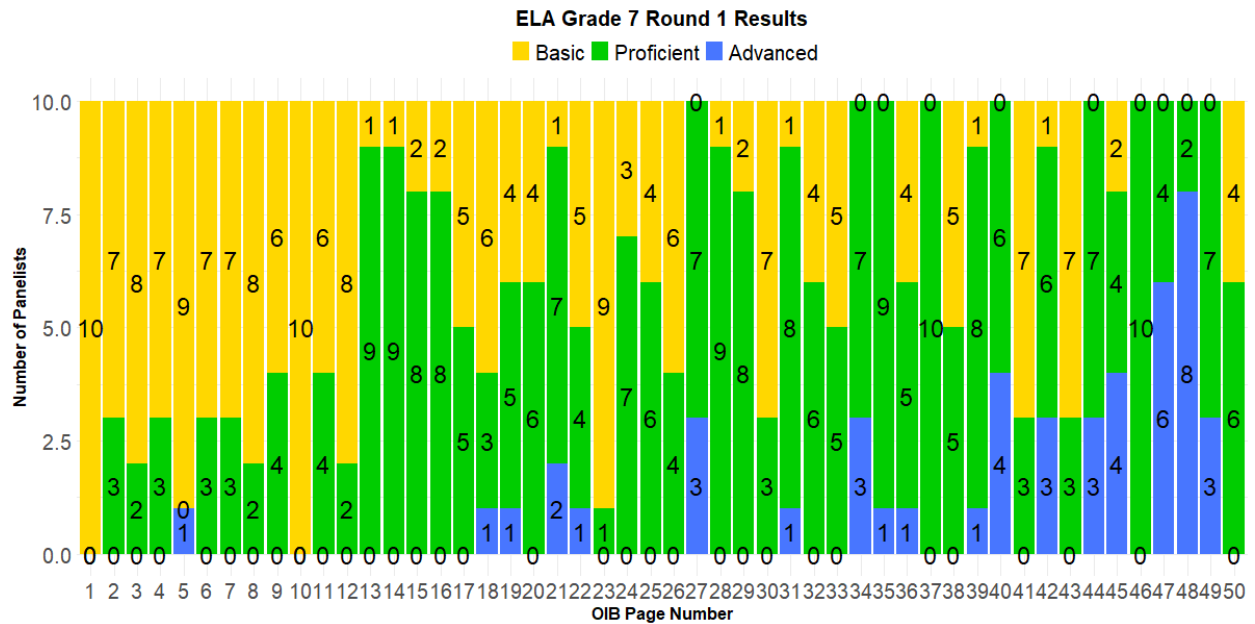


Figure 14. ELA Grade 7 Round 2 - Frequency of Panelist Judgments by Performance Level

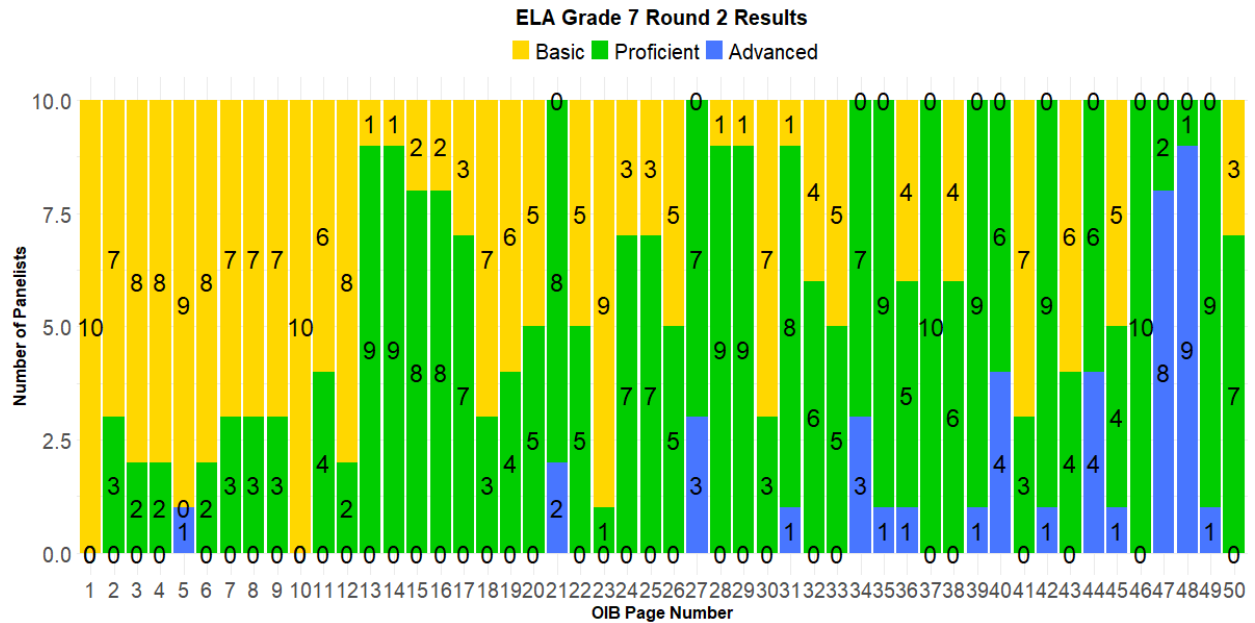


Figure 15. ELA Grade 7 Round 3 - Frequency of Panelist Judgments by Performance Level

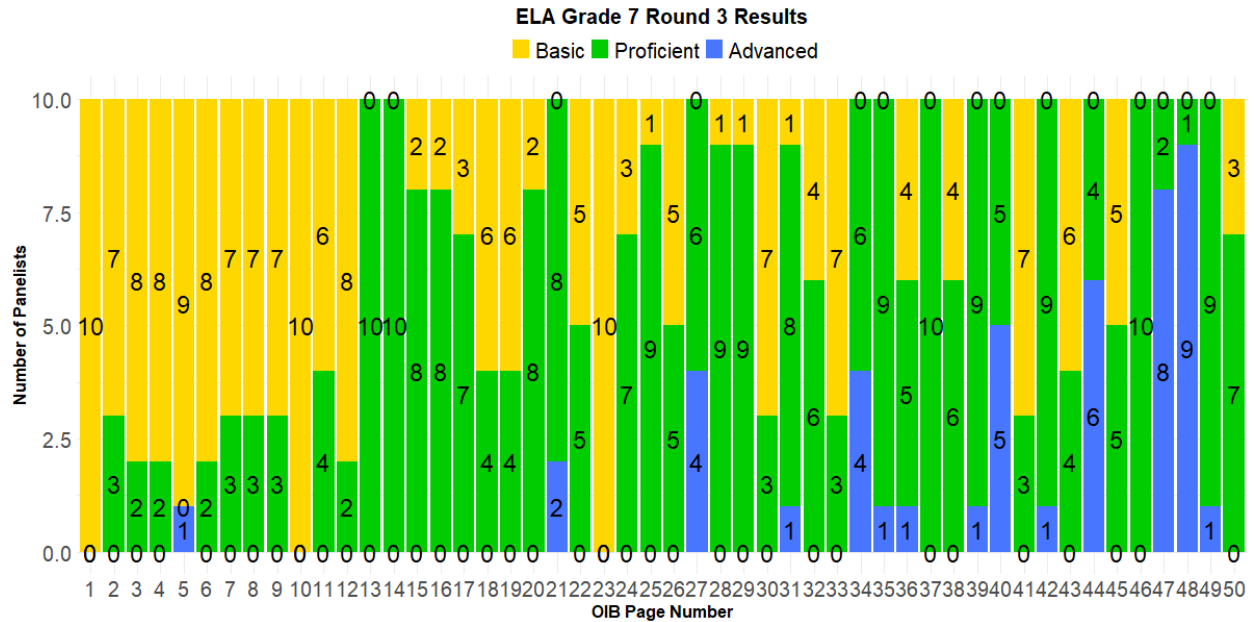


Figure 16. ELA Grade 8 Round 1 - Frequency of Panelist Judgments by Performance Level

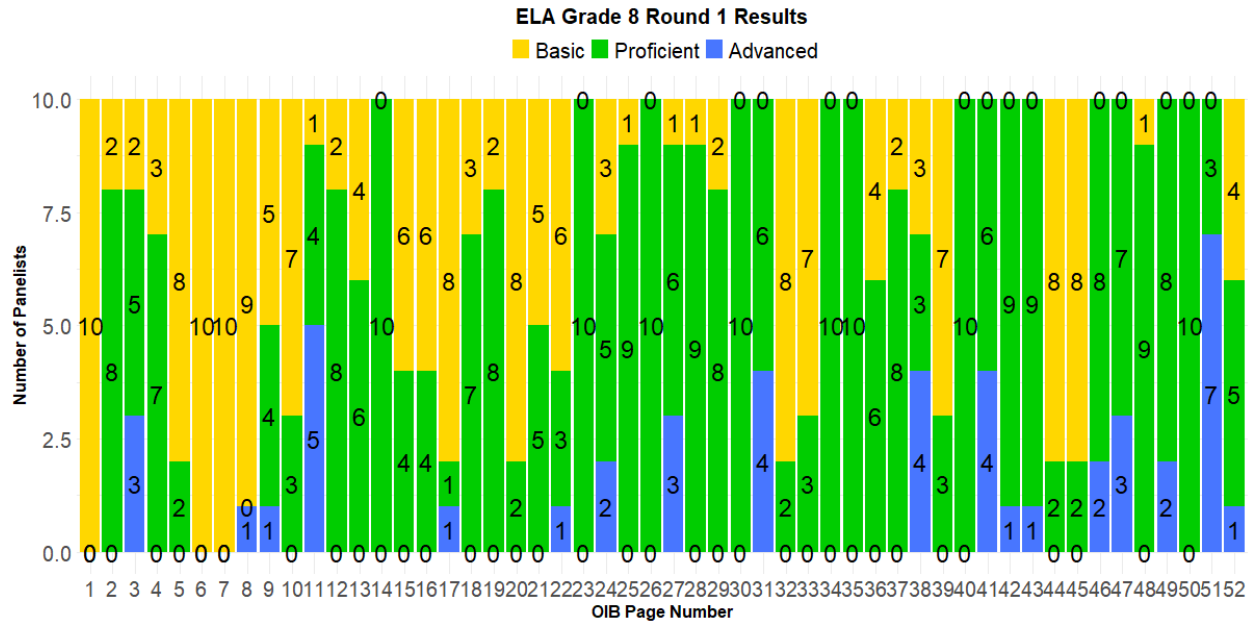


Figure 17. ELA Grade 8 Round 2 - Frequency of Panelist Judgments by Performance Level

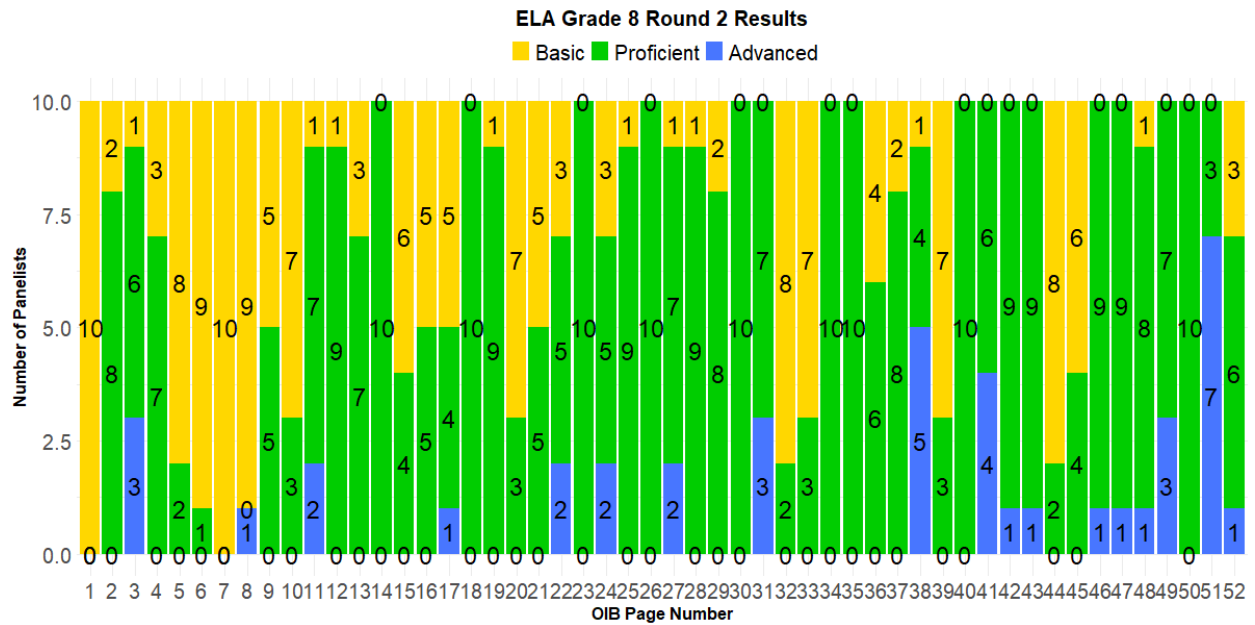


Figure 18. ELA Grade 8 Round 3 - Frequency of Panelist Judgments by Performance Level

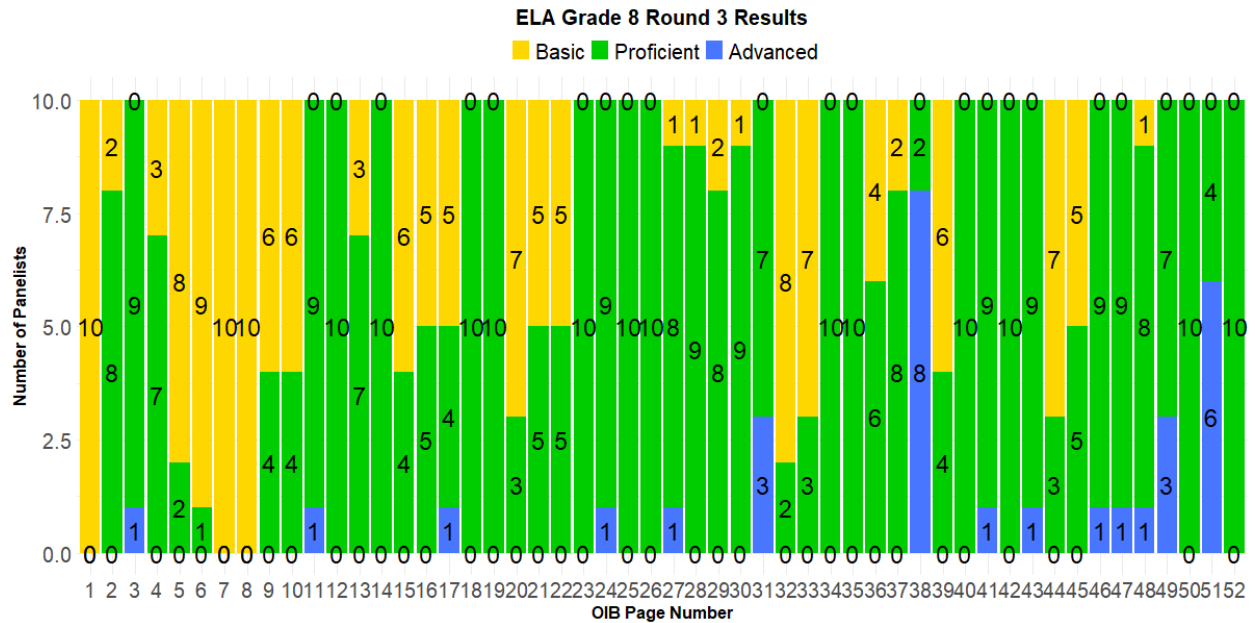


Figure 19. Mathematics Grade 3 Round 1 - Frequency of Panelist Judgments by Performance Level

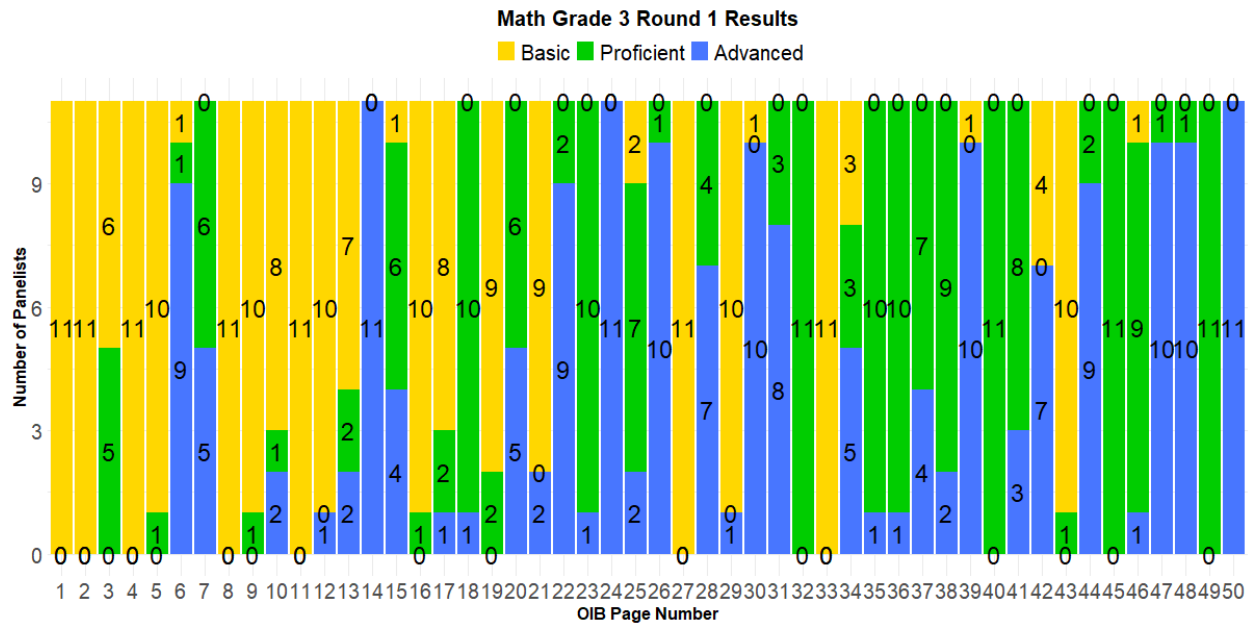


Figure 20. Mathematics Grade 3 Round 2 - Frequency of Panelist Judgments by Performance Level

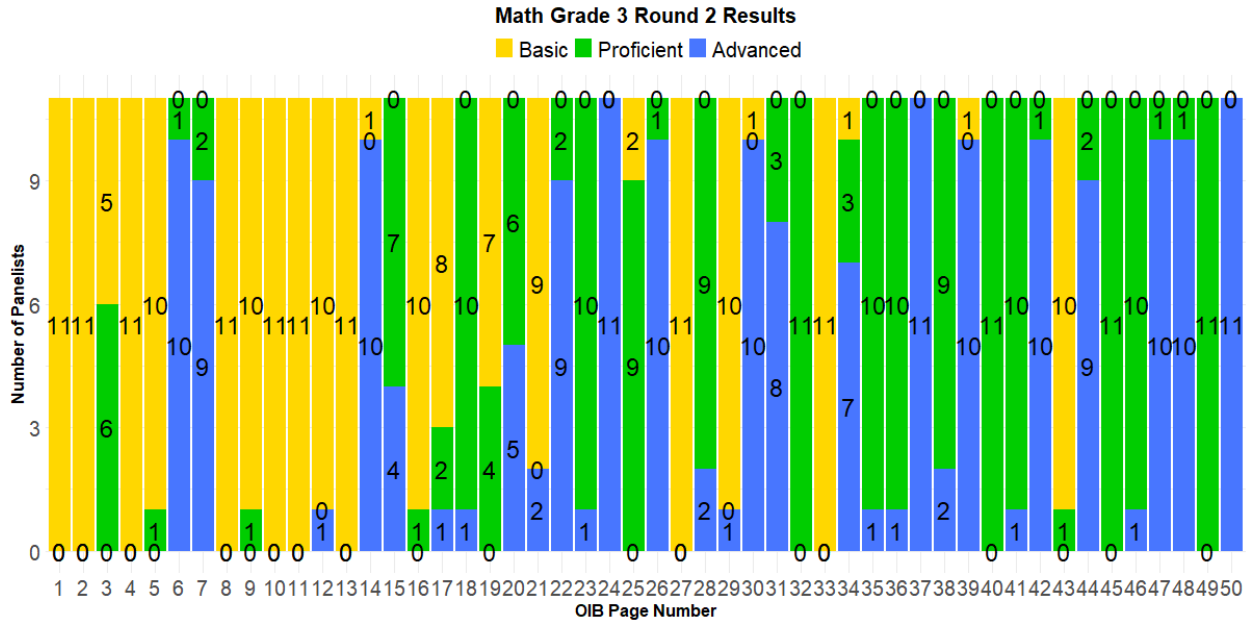


Figure 21. Mathematics Grade 3 Round 3 - Frequency of Panelist Judgments by Performance Level

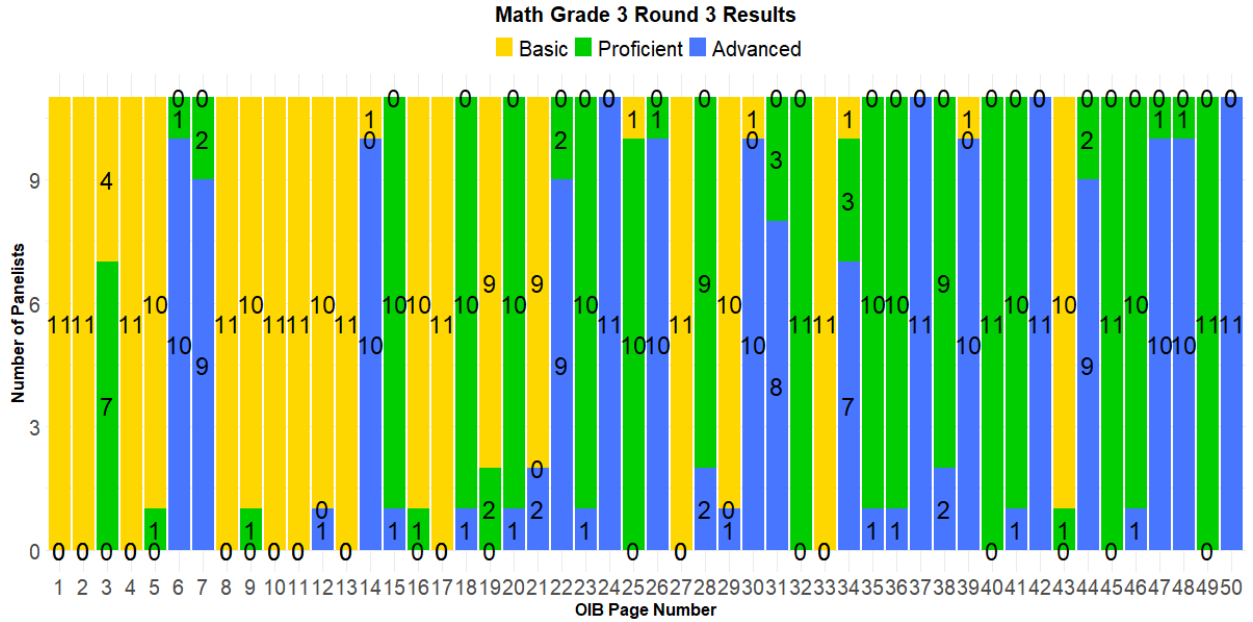


Figure 22. Mathematics Grade 4 Round 1 - Frequency of Panelist Judgments by Performance Level

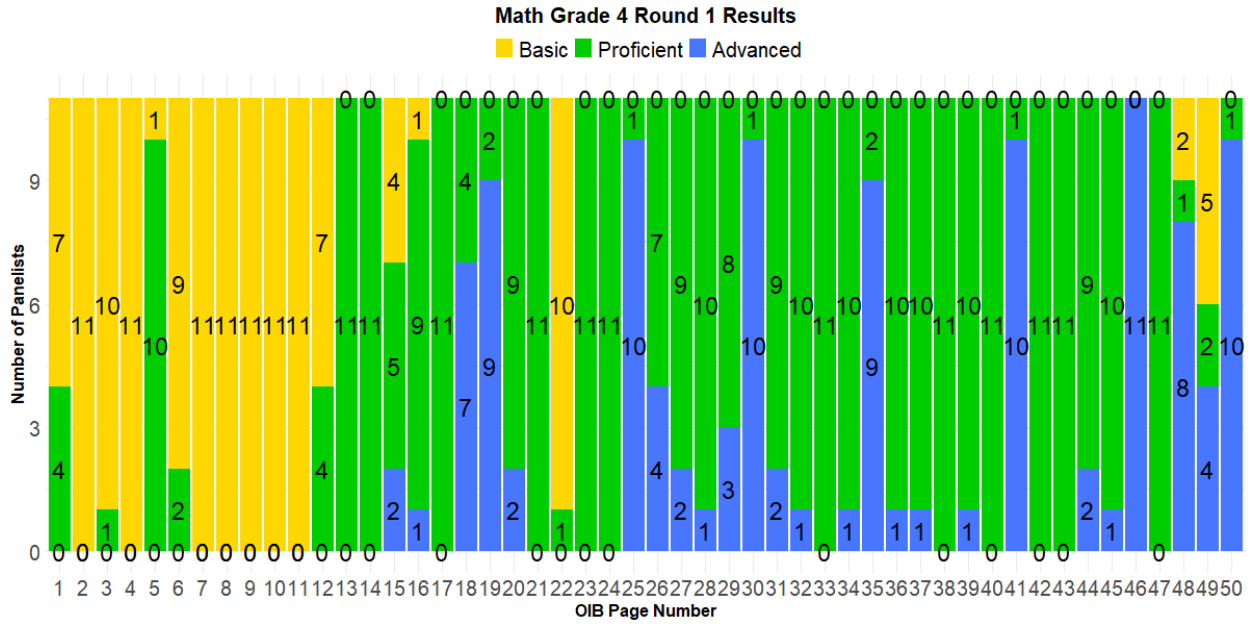


Figure 23. Mathematics Grade 4 Round 2 - Frequency of Panelist Judgments by Performance Level

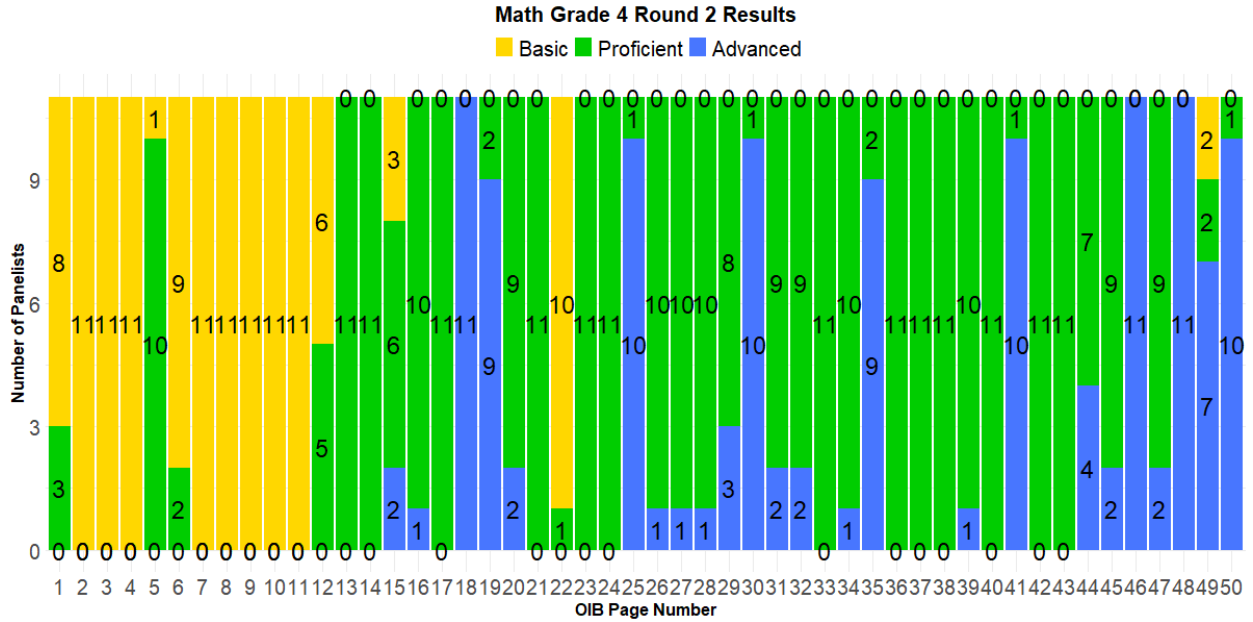


Figure 24. Mathematics Grade 4 Round 3 - Frequency of Panelist Judgments by Performance Level

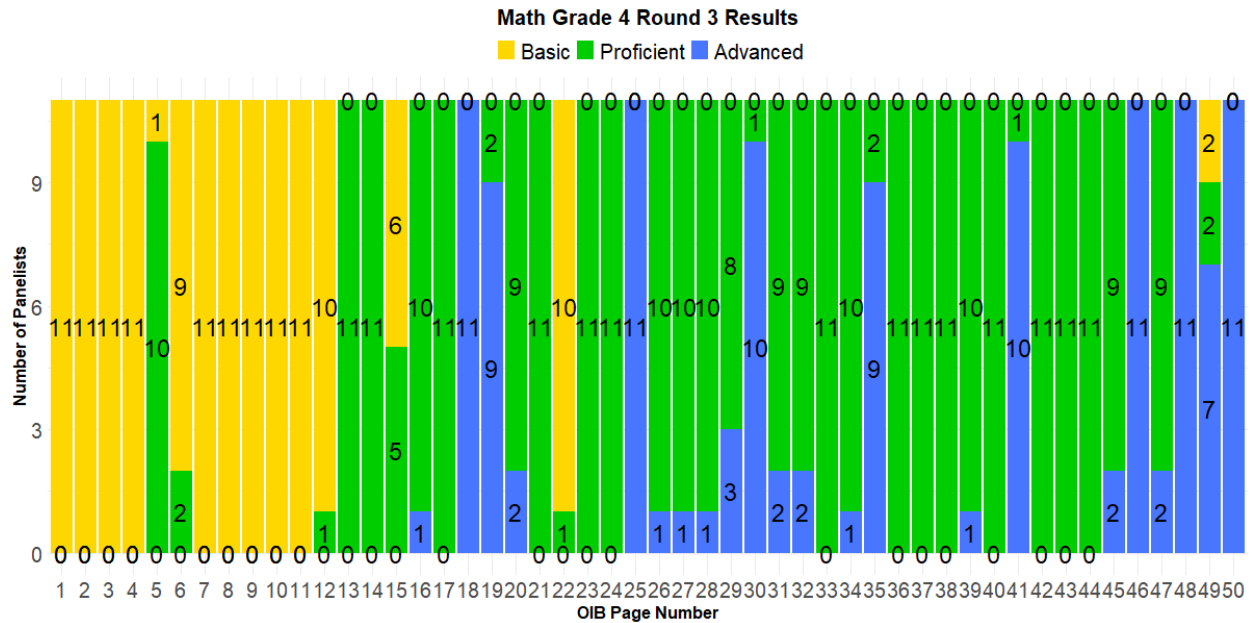


Figure 25. Mathematics Grade 5 Round 1 - Frequency of Panelist Judgments by Performance Level

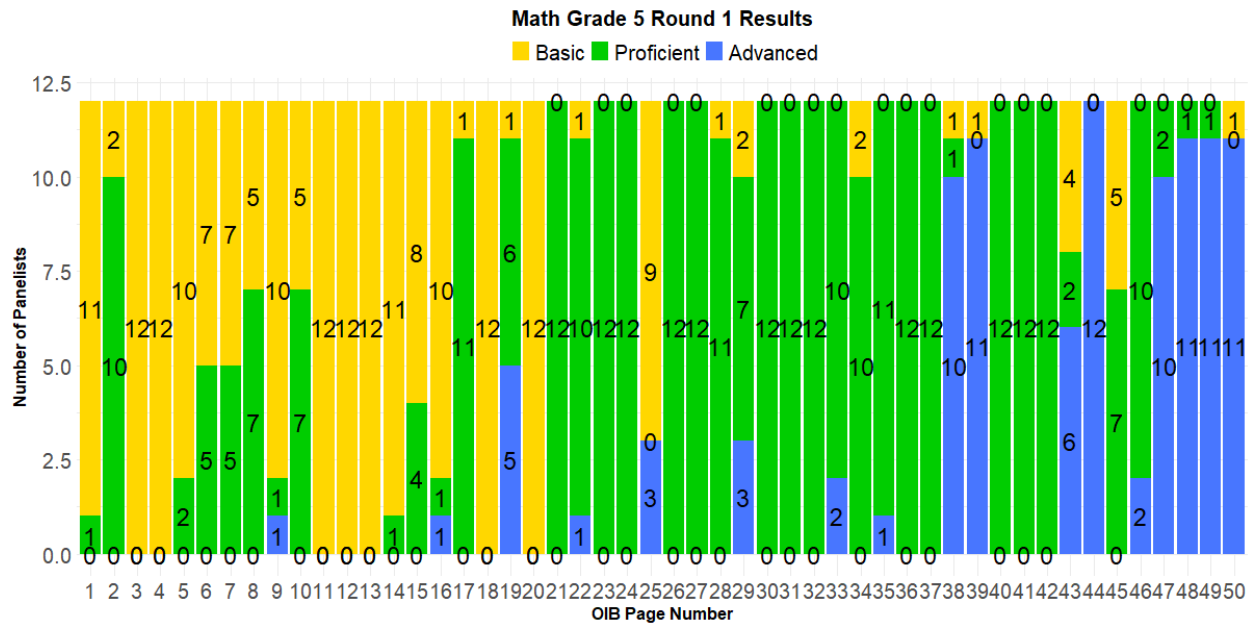


Figure 26. Mathematics Grade 5 Round 2 - Frequency of Panelist Judgments by Performance Level

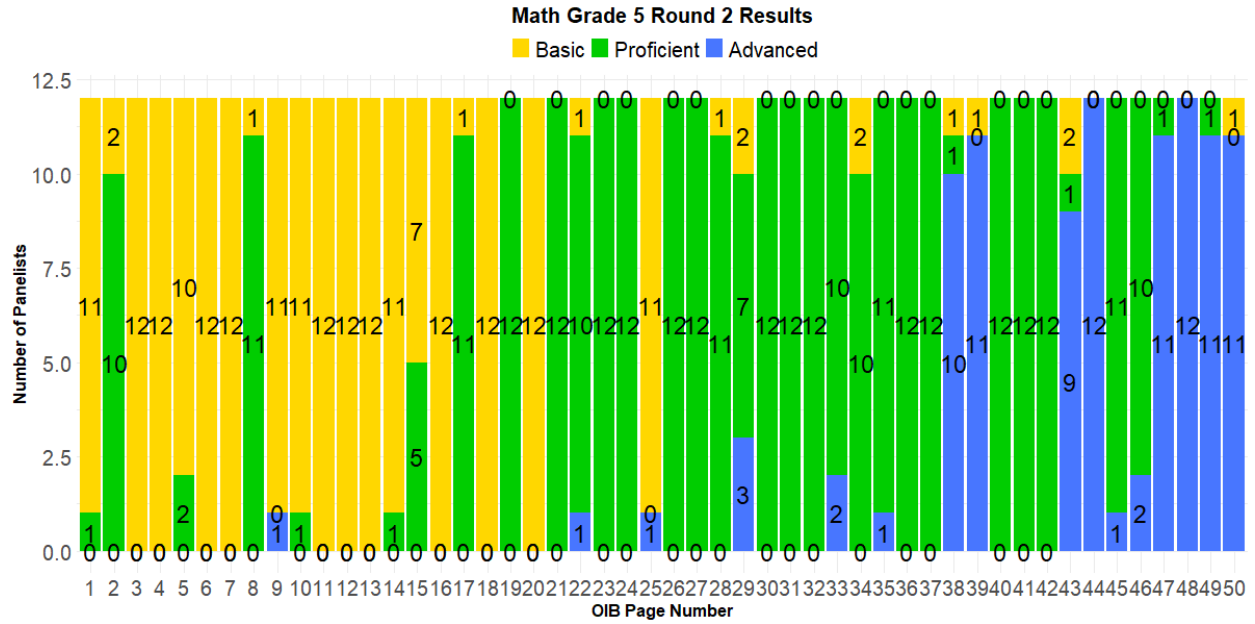


Figure 27. Mathematics Grade 5 Round 3 - Frequency of Panelist Judgments by Performance Level

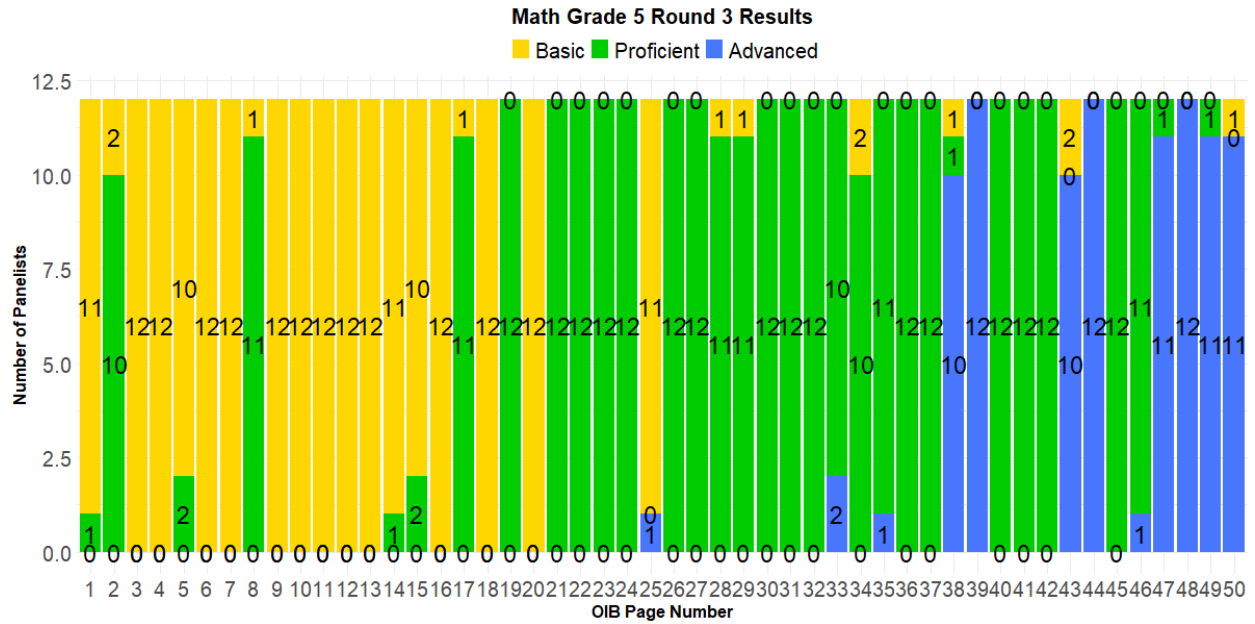


Figure 28. Mathematics Grade 6 Round 1 - Frequency of Panelist Judgments by Performance Level

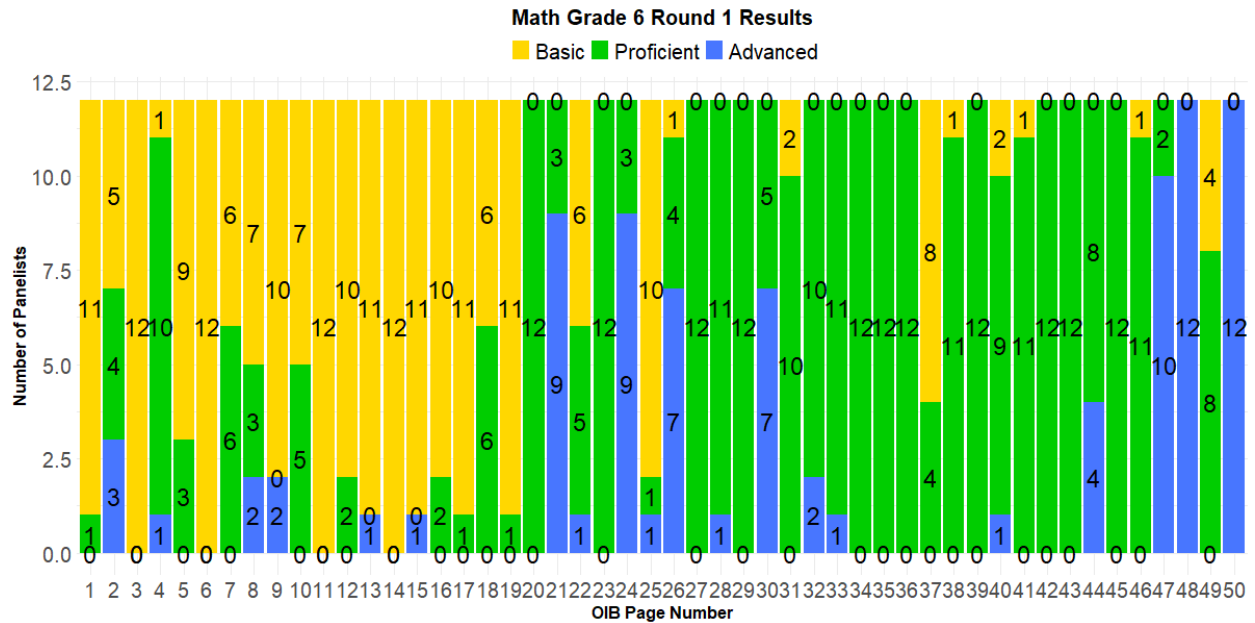


Figure 29. Mathematics Grade 6 Round 2 - Frequency of Panelist Judgments by Performance Level

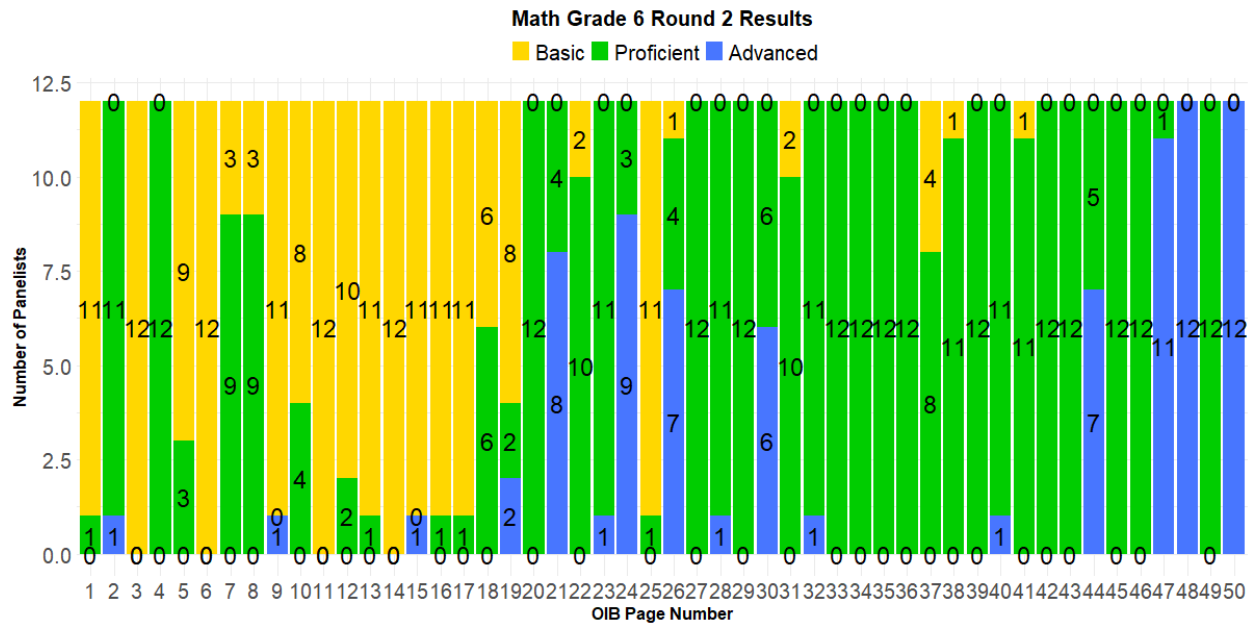


Figure 30. Mathematics Grade 6 Round 3 - Frequency of Panelist Judgments by Performance Level

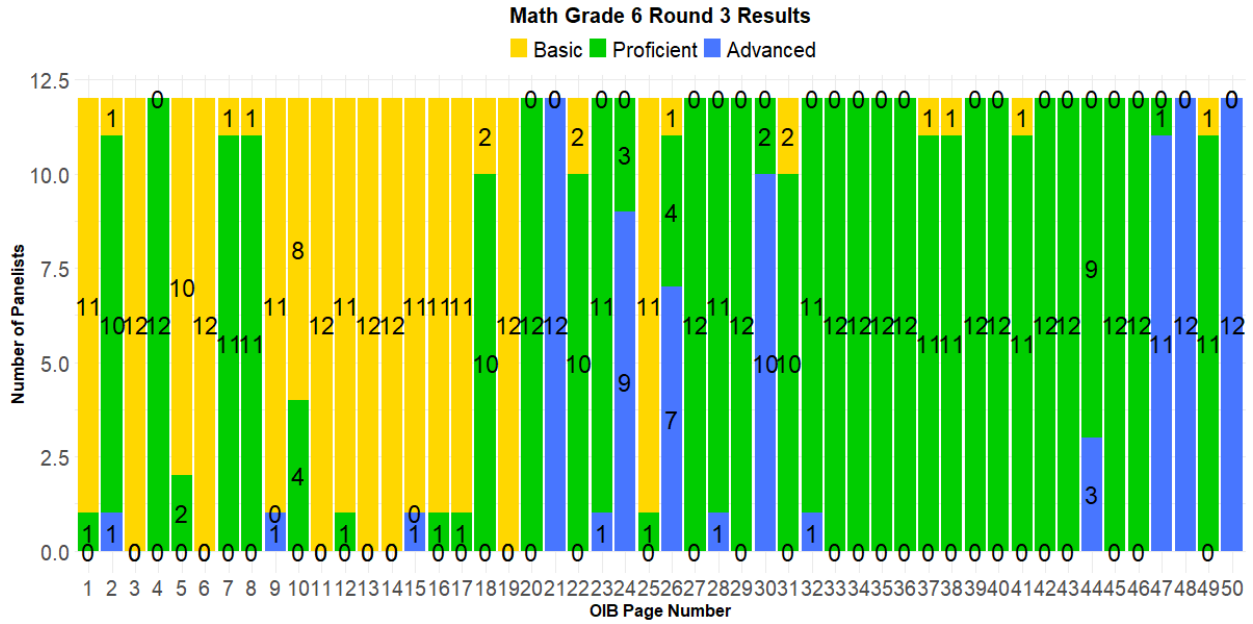


Figure 31. Mathematics Grade 7 Round 1 - Frequency of Panelist Judgments by Performance Level

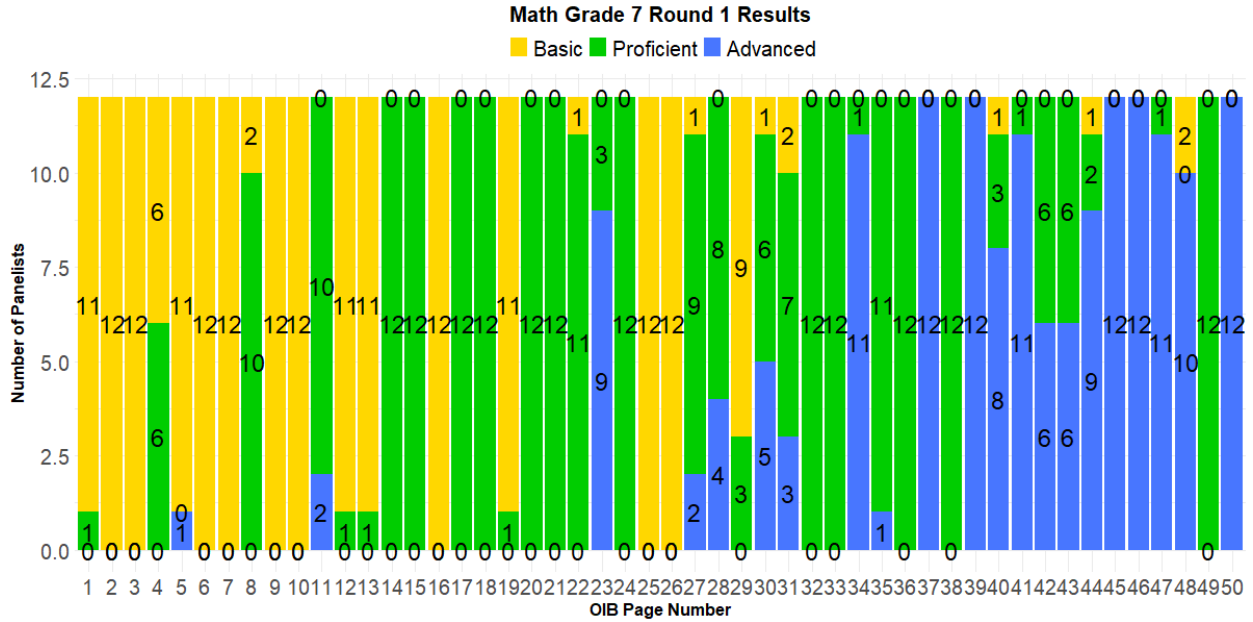


Figure 32. Mathematics Grade 7 Round 2 - Frequency of Panelist Judgments by Performance Level

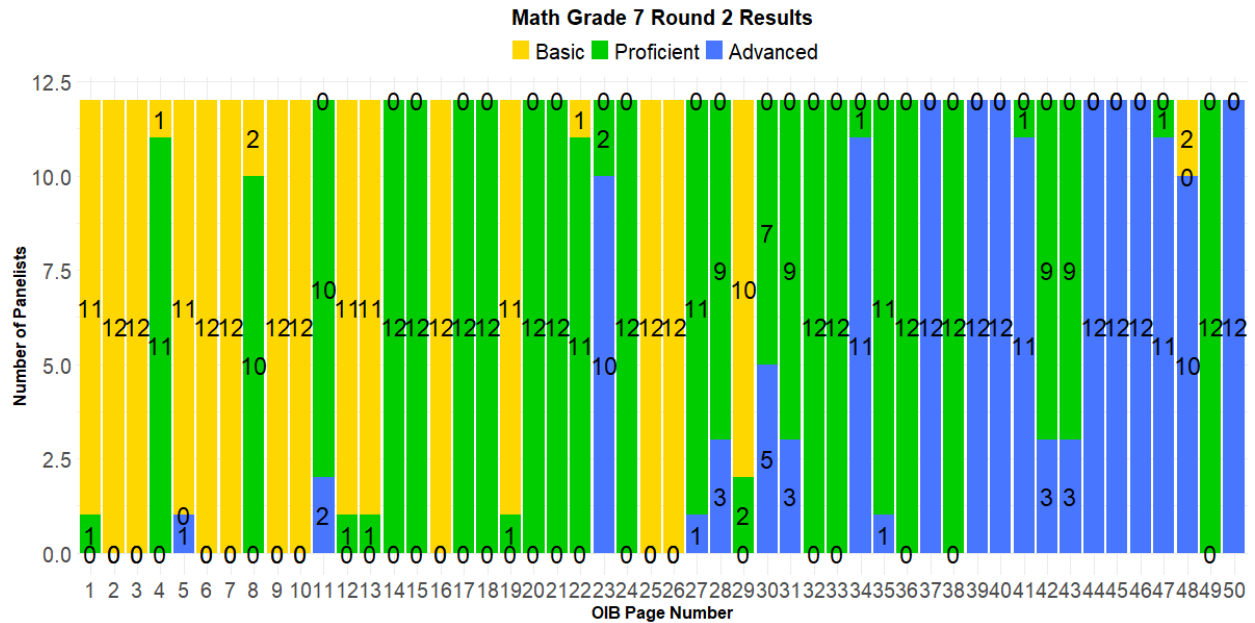


Figure 33. Mathematics Grade 7 Round 3 - Frequency of Panelist Judgments by Performance Level

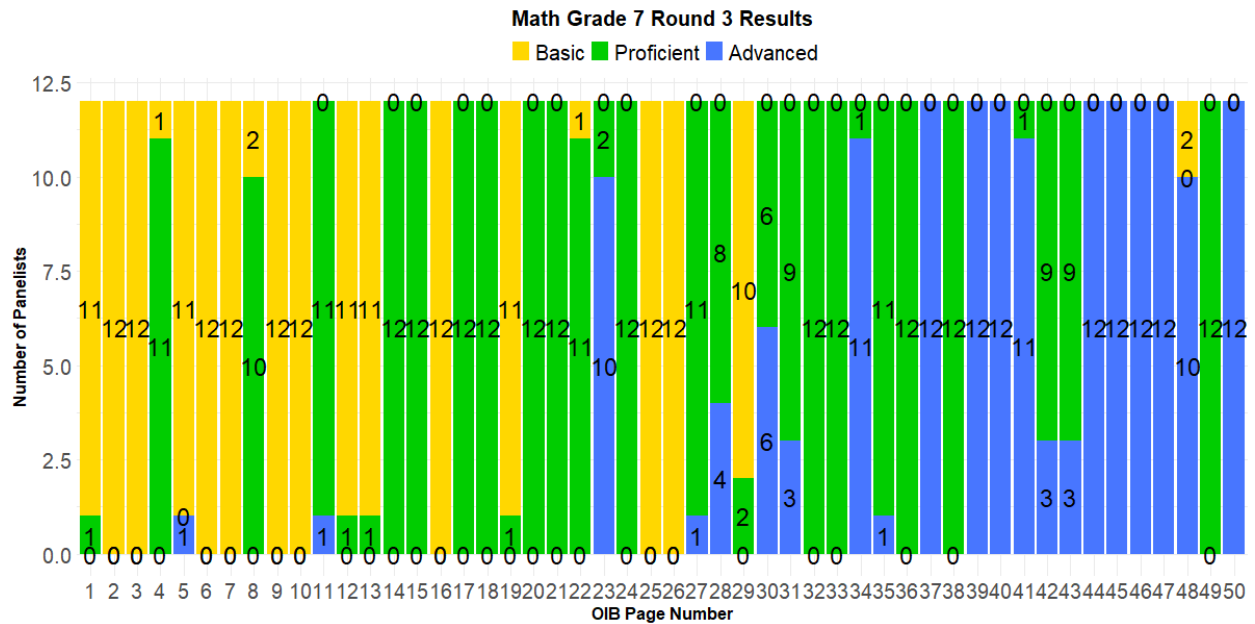


Figure 34. Mathematics Grade 8 Round 1 - Frequency of Panelist Judgments by Performance Level

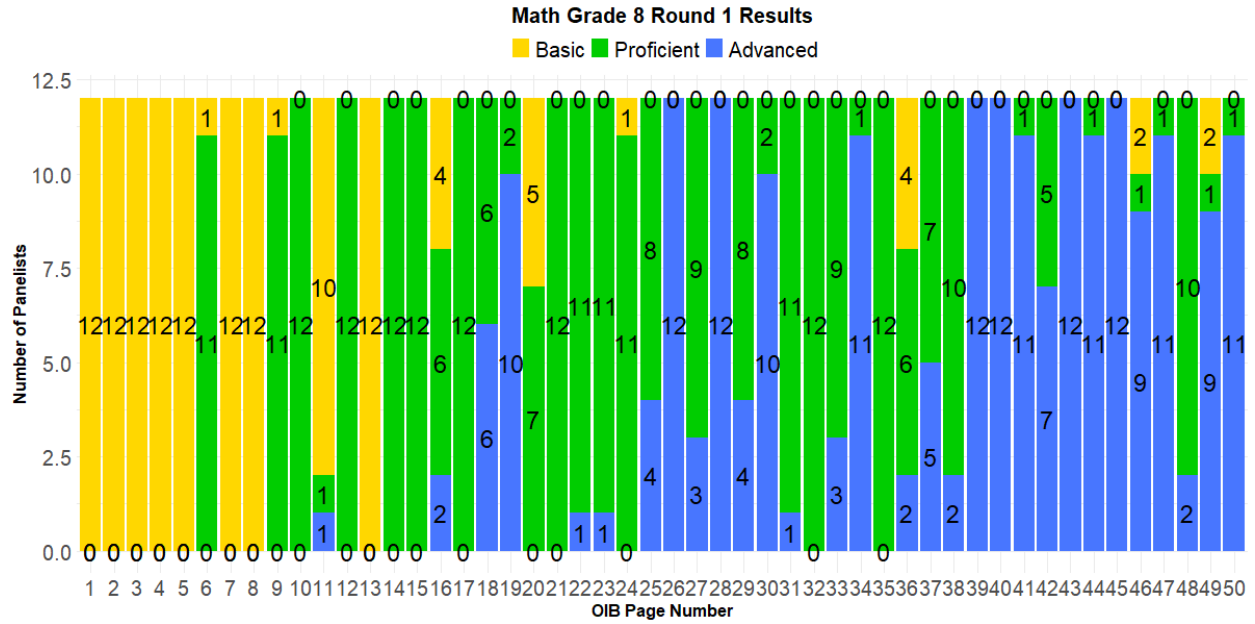


Figure 35. Mathematics Grade 8 Round 2 - Frequency of Panelist Judgments by Performance Level

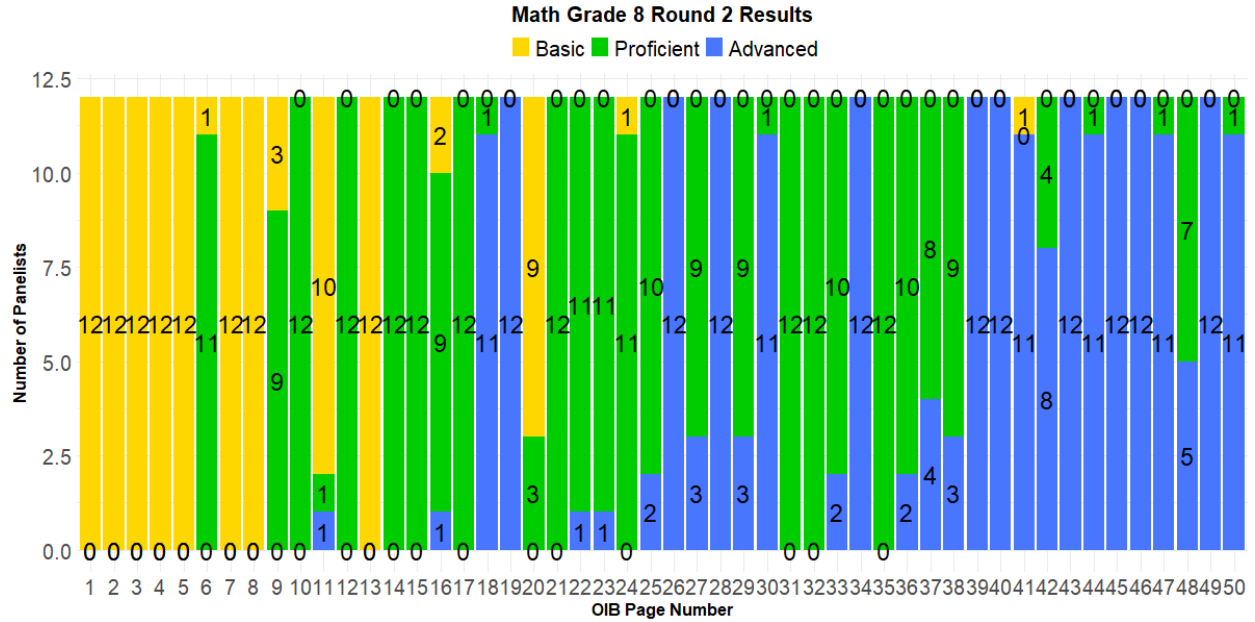
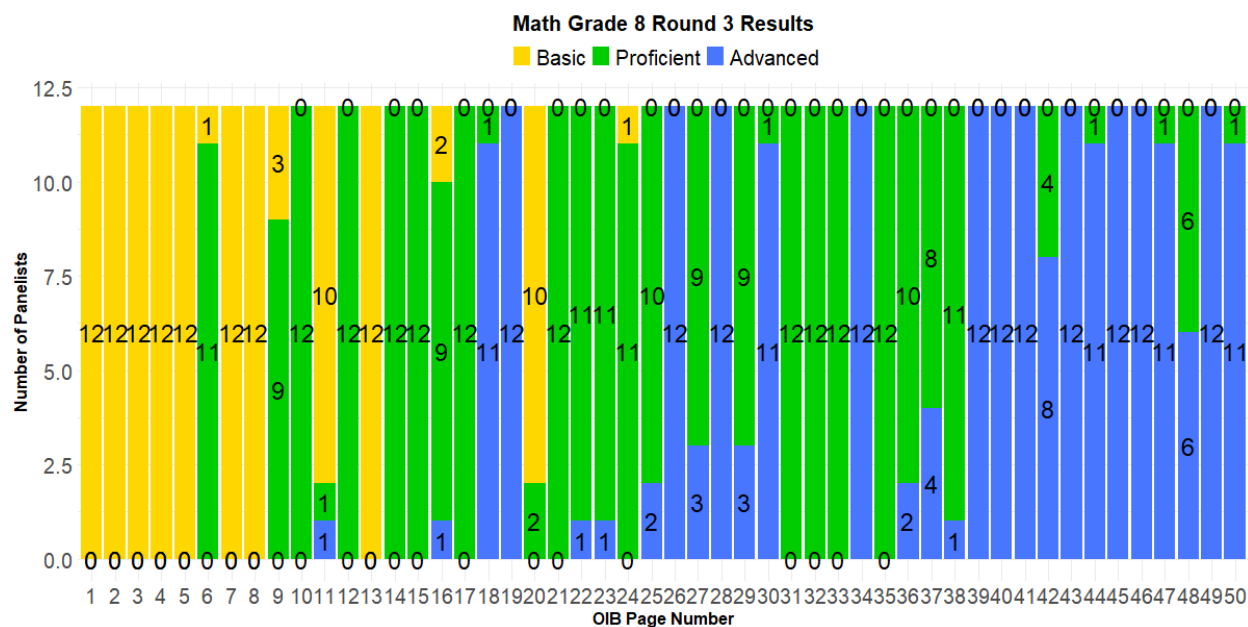


Figure 36. Mathematics Grade 8 Round 3 - Frequency of Panelist Judgments by Performance Level



APPENDIX—K
STANDARD-SETTING EVALUATION SURVEY & RESULTS

Table 1. ELA Panel Grades 3 & 4 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	1	10
2	I understood the procedures we followed to set standards.	0	0	0	1	10
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	0	11
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	1	10
5	I am confident about my understanding of this standard setting process.	0	0	0	2	9
6	The workshop facilitator explained things clearly to us.	0	0	0	0	11
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	0	11
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	1	10
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	0	11
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	1	10
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	0	6	5
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	0	0	0	4	7
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item response demands to PLDs.	0	0	0	4	7
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	1	10
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	2	9
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	1	10
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	0	1	10
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	0	1	10
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for ELA GRADE 3, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	8	3	0	0	0
20	Based on the impact data results for ELA GRADE 3, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	1	3	7	0	0
21	Based on the impact data results for ELA GRADE 3, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	9	2	0	0
22	Based on the impact data results for ELA GRADE 3, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	9	2	0	0
23	Based on the impact data results for ELA GRADE 4, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	10	1	0	0	0
24	Based on the impact data results for ELA GRADE 4, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	1	1	9	0	0
25	Based on the impact data results for ELA GRADE 4, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	7	4	0	0
26	Based on the impact data results for ELA GRADE 4, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	9	2	0	0

Table 2. ELA Panel Grades 3 & 4 – Text Responses for Open-ended Questions

Question #	Question Text	Response
27	Please indicate any parts of the standard setting training and process that we should improve.	I think that maybe there could be a better understanding of what the final goal is in the beginning before round 1. I wasn't fully understanding the final goal until after round 1.
		There were several PLDs that were too closely aligned that made it tricky to decipher which PLD to decide on. In our group we went back and forth between several in both 3rd and 4th grade ELA. I would recommend more clearly stating some of those PLDs to separate them more. For example, some of the PLDs only differed by "identify" vs. "find." If the PLDs stay as is, I would recommend adding the "Assessment Words" sheet to the PDF file of the PLDs for teachers to reference. I would also clarify what the difference is between "identify" and "find."
		Standards should be clearer on the few standards that require an opinion.
		This was my first time doing this and it was very well planned, and the instructor was a great help answering any questions that arose.
		Before beginning the workshop, I felt small and unqualified to be here. After the training, I can confidently say I felt equipped with the tools needed to get the job done. In the summarizing standard on 3rd grade, it does not state "summary" in the PLD Advanced. After discussion, we feel like it's probably implied but maybe we could look at that PLD again and possibly add in the summary expectation.
		The first day is really long and overwhelming. I feel that some of it could be condensed down a bit and the room facilitators could explain the process in the room so we can go at our own pace.
		I only have one suggestion, and it is for seating placement. In the meeting we had a table of four. My chair placement had my back to the projector. I would recommend considering that for any future training sessions. I had to turn around to see the projector.
		The amount of down time, the waiting around tiring.
		I think a flowchart, or a pyramid diagram or some sort of visual aid would be helpful in knowing how to go about making decisions on items that don't perfectly align with the PLD. Do we place more weight on staying as close to the exact wording on the PLD? Do we consider text complexity/answer choices? Do we consider what we believe most students in the grade level are capable of doing/understanding?
		Having to discuss your own opinions about each standard was highly intimidating. People are not understanding even if they are told that it is ok to disagree. Teachers in particular are hard to carry out a discussion platform with because everyone thinks they are right and are not very understanding when someone doesn't agree. I don't know how to make it less intimidating but that would be my recommendation for the next standard setting process.
		Round 1 is long and tedious with needing to figure out KSAs and PLDs for all items. I am not sure how it would work with time, but perhaps splitting round 1 work into smaller chunks/sections would help with item fatigue. Some of the later items in the OIB require more thought (either due to item complexity or trying to comprehend why students found these specific items the most difficult) and after dissecting the other questions apart to determine KSAs and PLD, some of those later OIB items did not get the focus or attention they deserved in round 1.
		28
Most of the PLD's were clear and easy to tell the difference between levels.		
I believe being able to discuss with peers after each round was very helpful.		
The training was awesome. The discussion in between rounds was very valuable.		
The discussions after the rounds were very informative and I enjoyed listening to other teacher's thoughts and ideas. Our facilitator Jessica, was very informative and it was nice to work with her.		
I thought it was mapped out well. We stayed on task and followed the schedule pretty closely. I like having an agenda to follow.		
I loved the process and learning about how this works. I loved getting the opportunity to be part of this and learn. I feel that my input along with other teachers input is valuable.		
I appreciated working to make my own judgements first and then having two opportunities to discuss items.		
The training was beneficial I felt the way things were explained and the documents that we provided for me to use helped me to understand and fulfill the process to my best of abilities.		
continued		

Question #	Question Text	Response
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	The ability to debrief with fellow colleagues between rounds really helped me understand the way others viewed specific items on interpreted the PLDs.
29	Please note any other feedback you would like us to consider.	<p>Thank you for being very generous hosts. I have never eaten so much in my life. You spoiled us!</p> <p>Treating yourself to a job well done!</p> <p>I really enjoyed the opportunity to help set the standards.</p> <p>This was an incredible learning experience! I will be honest. I signed up for this because I saw "stipend" and "travel accommodations" in the email. I did not have a clue what to expect. After my 4 days here though, I can honestly say I am so happy I came. It was really cool to see a piece of the puzzle behind the scenes and be a part of it. In addition to that, I truly believe using the PLDs this week will have me using them regularly in the classroom and really help me understand discrepancies in some of the complexity of learning materials in the classroom.</p> <p>I would love to participate in these types of meetings, data gathering more often. It has helped me as a teacher with my knowledge and understanding of the standards and has given me ideas that I will be using in my classroom this year.</p> <p>The 3rd grade PLDs were more clear on distinguishing between the proficiency levels compared to 4th grade. It was easier to align test items to the PLDs with the 3rd grade set. I am not sure if this is something to consider before PLDs are approved.</p>

Table 3. ELA Panel Grades 5 & 6 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	2	8
2	I understood the procedures we followed to set standards.	0	0	0	3	7
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	2	8
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	5	5
5	I am confident about my understanding of this standard setting process.	0	0	0	4	6
6	The workshop facilitator explained things clearly to us.	0	0	0	1	9
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	2	8
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	2	8
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	1	9
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	3	7
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	1	3	6
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	0	0	0	2	8
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item response demands to PLDs.	0	0	0	4	6
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	1	9
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	1	9
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	2	8
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	0	4	6
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	0	3	7
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for ELA GRADE 5, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	3	6	0	1	0

continued

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
20	Based on the impact data results for ELA GRADE 5, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	2	4	4	0	0
21	Based on the impact data results for ELA GRADE 5, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	7	3	0	0
22	Based on the impact data results for ELA GRADE 5, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	3	7	0	0	0
23	Based on the impact data results for ELA GRADE 6, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	8	1	0	1	0
24	Based on the impact data results for ELA GRADE 6, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	1	3	6	0	0
25	Based on the impact data results for ELA GRADE 6, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	8	2	0	0
26	Based on the impact data results for ELA GRADE 6, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	2	8	0	0

Table 4. ELA Panel Grades 5 & 6 – Text Responses for Open-ended Questions

Question #	Question Text	Response
27	Please indicate any parts of the standard setting training and process that we should improve.	I would have appreciated being assigned to the specific grade-level band in which I was familiar. I was moved up the first day of the workshop, which added an additional layer of stress in being unfamiliar with both grades of which I participated.
		I think that a little bit more time should be allotted to DAY 1 of the standard setting process. I felt a little bit "rushed through" learning all of the new vocabulary terms & their meaning. I did not feel adequately prepared to begin "Round One" on the first day. There was ALOT of new information to mentally process and retain before "Round One."
		I liked this step in the process, I wish the PLD writing had as much training as this had and had a vertical articulation as well. I feel like the PLDs are unnecessarily flawed and inconsistent. I think there is a lot of room for improvement there.
		I really enjoyed this process otherwise. I loved the discussions and I felt like it is a solid process.
		I would have liked to have a conversation about our answers with my table as well as the room
		maybe a little more time explaining the initial process on day 1
		I liked how it was broken down. I think discussions allowed us to revisit the PLD alignment. The part I would change would be only visiting questions with a wide range of discrepancy.
		Table groups should be shuffled daily to provide for alternative perspectives in the small table conversations and discussion s that inevitably crop up between rounds.
		I think that we shouldn't have known about the OIB questions being in order until after the first round and the colored bands for data until the last round. Sometimes I felt pressured to make my judgements align with expectations. I would like time to discuss more of the questions. I know time is an issue, but I feel it would be helpful.
		Some PLDs were almost identical to others and resulted in lengthy discussions. Other wording could have been used so the differences were more apparent.
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	I thought the three rounds and discussions were adequate. It gave my group plenty of opportunity to discuss and rethink our choices, and I felt my final decisions were on target.
		I do feel that our workshop facilitator did a great job helping us prepare for tasks and keeping panelists on task.
		The people, the amount of time it took, the focus on training, and the inclusion of round discussions.
		The training and discussions
		discussion
continued		

Question #	Question Text	Response
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	more understanding as we went through the process. The facilitator was amazing and helpful. gave us great knowledge
		I thought the rounds work really well.
		The general format (individual, analysis, discussion, repeat) was very effective. It allowed me to clarify items where needed and provided other viewpoints for items I had felt confident about. It also allowed grade-level experts to clarify items for those who did not teach that grade.
		I thought the process worked very well. Our facilitator did an amazing job of keeping us moving along and explaining everything. I liked the size of the group and the ease with which we were able to communicate and collaborate. I felt that the process was very supportive.
		Everything worked well except as noted in #27
29	Please note any other feedback you would like us to consider.	I enjoyed the facility and thought the staff did an excellent job hosting us. I also thought it was a fairly smooth 4 days of work. Everyone on the Cognia and OSDE teams worked hard and in tandem to ensure we had everything we needed to do our week efficiently/effectively.
		Thank you for the opportunity to contribute knowledge and teaching experience to standard setting scores cuts this school year. It is my hope that our panelist group helps student learning to improve in some way with this exercise.
		It was fun and insightful
		I enjoyed it and would love to attend more!
		I felt the meeting really helped me familiarize myself with the standards of 5th grade.
		I am concerned that some people are participating in too many steps of the process. One individual in my group will have participated in 3 different elements of this process. Since these are very small groups, I worry that this could cause some bias. While some overlapping participation is likely beneficial (particularly for vertical articulation), I am concerned about having some dominant voices heard too much. Other than that, I feel that this was a very enjoyable, interesting, and valuable experience.
		I enjoyed being a part of this process. I feel like it was very helpful. I would like to have updates on how the process is going as it moves forward (mostly because I am just curious). I am a bit worried about how the OSDE will use the data - (to prove that public school isn't working) and I would like to know that the data isn't being overly manipulated.
		You did an excellent job by involving and listening to teachers who are at the frontline of this education war!

Table 5. ELA Panel Grades 7 & 8 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	1	9
2	I understood the procedures we followed to set standards.	0	0	0	3	7
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	1	9
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	3	7
5	I am confident about my understanding of this standard setting process.	0	0	0	2	8
6	The workshop facilitator explained things clearly to us.	0	0	0	2	8
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	2	8
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	2	8
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	1	9
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	1	9
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	0	5	5
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	1	0	0	2	7
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item response demands to PLDs.	0	0	0	4	6
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	2	8
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	2	8

continued

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	2	8
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	0	1	9
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	0	1	9
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for ELA GRADE 7, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	5	4	0	1	0
20	Based on the impact data results for ELA GRADE 7, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	0	4	5	1	0
21	Based on the impact data results for ELA GRADE 7, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	10	0	0	0
22	Based on the impact data results for ELA GRADE 7, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	8	2	0	0
23	Based on the impact data results for ELA GRADE 8, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	2	8	0	0	0
24	Based on the impact data results for ELA GRADE 8, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	0	9	1	0	0
25	Based on the impact data results for ELA GRADE 8, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	9	1	0	0
26	Based on the impact data results for ELA GRADE 8, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	7	3	0	0

Table 6. ELA Panel Grades 7 & 8 – Text Responses for Open-ended Questions

Question #	Question Text	Response
27	Please indicate any parts of the standard setting training and process that we should improve.	Great job! Thank you!
		I think that the training and process went smoothly, and everything was presented well and thought out.
		Provide clarity on the thinking behind creating the PLDs when considering passage complexity and genre.
		Standard setting for the second-grade level went more smoothly than for the first-grade level, because I had a better understanding of how to navigate the OIB and provide KSAs more efficiently. It would have been helpful to see a couple of examples of what it might look like to complete the KSA, notes, and ID match before beginning to know how much or how little to write.
		Break the work into smaller parts to prevent fatigue
		none
		Maybe let people know about the details sooner. It is a little easier to plan childcare and similar with more notice.
		Clearly articulating the expectations of the participants during breaks and downtime. There were lots of times that down time was ambiguous about how long or what participants were supposed to do/be.
		The process was straightforward, so I don't have any suggestions for this one.
		n/a
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	Loved our facilitator; loved the immediate data provided to inform each step of the process.
		I feel like the process was really organized and everything went really well.
		It went well when we are able to discuss our reasoning behind the items. However, some felt like we all had to have the same result.
		The debrief rounds with the breakdown of the participant results was super helpful in determining which questions we needed to discuss further.
		Cognia was great. Food was good. Isolation from home distractions allowed many teachers to focus and provide valued input.
		small groups
		The discussions held after the data was processed was valuable. It confirmed some of my ideas while challenging others.
		I really enjoyed the PLDs as well as the discussions. I did not love the independent work time, but it was helpful to have done that front loading, so our discussions were more productive. I also liked that we had a space to add comments or suggestions on things outside the work of Standards Setting even if we did constantly say them aloud anyway.
		Hearing the expertise in the room was helpful to inform my own judgments.
		I felt like it went well.
29	Please note any other feedback you would like us to consider.	Learned a lot this week! Going to buy a book on psychometrics this week!
		I think everything went really well and I enjoyed the experience of being on the panel.
		The process overall was well thought out, and the Cognia and SDE team did a great job keeping us on track.
		Thank you for the invitation.
		The hours, being in the summer, were a little long. I realize there is a lot to include, but it is a long day, especially when driving to the site.
		N/A
		I think we need norms for the discussion process. #11: I understood the progressions but encountered some PLD definitions that were vague in relationship to the item.
		I know there needs to include a good mixture of stakeholders on the panel, but it might be beneficial to have a couple more current classroom teachers who are in the trenches. Maybe like a 70/30 ratio. Just a suggestion. We did have a good group, though.

Table 7. Mathematics Panel Grades 3 & 4 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	4	7
2	I understood the procedures we followed to set standards.	0	0	0	3	8
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	1	10
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	4	7
5	I am confident about my understanding of this standard setting process.	0	0	1	5	5
6	The workshop facilitator explained things clearly to us.	0	0	0	5	6
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	3	8
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	7	4
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	5	6
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	3	8
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	0	6	5
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	0	0	0	4	7
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item, and matching those item-response demands to PLDs.	0	0	0	5	6
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	2	9
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	2	9
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	2	9
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	0	5	6
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	0	5	6
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for Mathematics GRADE 3, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	2	9	0	0	0
20	Based on the impact data results for Mathematics GRADE 3, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	4	5	2	0	0
21	Based on the impact data results for Mathematics GRADE 3, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	5	6	0	0
22	Based on the impact data results for Mathematics GRADE 3, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	2	7	1	1	0
23	Based on the impact data results for Mathematics GRADE 4, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	7	4	0	0	0
24	Based on the impact data results for Mathematics GRADE 4, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	4	5	2	0	0
25	Based on the impact data results for Mathematics GRADE 4, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	5	6	0	0
26	Based on the impact data results for Mathematics GRADE 4, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	1	5	4	1	0

Table 8. Mathematics Panel Grades 3 & 4 – Text Responses for Open-ended Questions

Question #	Question Text	Responses
27	Please indicate any parts of the standard setting training and process that we should improve.	I enjoyed learning about this whole process. I think a good job was done by everyone to make us understand what was required of us.
		I feel like it would have been more beneficial to diversify the people in this group. The majority of people in this group were from small rural schools and I feel like it should have been a better mixture. (Title 1, larger school)
		The informant on the PLD we had on the 3rd round of 3rd grade was very informative and helpful, since she was in on the PLD conversations. I wish we had her present earlier in the rounds, that would have clarified some more things.
		I'll be honest, some of it was confusing, but as we dug deeper, I did understand it better. Day 4 I was a lot more confident than I was on Day 1.
		While I understand that it is important to have different people in each portion to help keep the results from skewing one way or another, I think that having the same person participate in 2/3 procedures would help with explaining. We had someone in our group who was on the item review and she was able to give helpful feedback (not specific, but helpful) during the process. Having several people in the room who had participated in multiple portions of the standards/item/PLD portion would have been even more beneficial. Quite a few of us were very frustrated with item quality and/or the PLD layout. I was concerned that I could not effectively evaluate and place some of the items due to this frustration.
		Once we began to use the materials the entire training became super clear.
		maybe a better explanation on how tests are rated after the rounds-
		The first day of training was long and repetitive.
		If a panel is divided on items after multiple discussions, the question should be thrown out.
		On Thursday after viewing final results, I would have liked a condensed recap of the Monday morning training and description of the process, next steps, etc. The bug in the standard setting toolkit needs to be fixed.
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	I liked being able to review the material as a group and listen to other people talk about their idea of what the answer is and the reason for it.
		I appreciated the sharing and "debate" in each round. I felt that the overall process worked well
		I believe you were very informative and gave all the information between the standards, PLD and OIB
		Everyone from Cognia to OSDE were very helpful when we did have questions. Our facilitator, Karen Whisler, was amazing, too! It really did go pretty well. It was a great experience for me!
		I thought the people from Cognia and the SDE were very knowledgeable and helpful with understanding the process and allowing us to really talk through the process. Karen was especially helpful to bring us back to the process at hand when we got sidetracked. The food and snacks were really varied and a welcome addition to the day!
		We had plenty of time and really good discussions about the PLDs/how the items aligned. I really appreciated the insights into the whole process.
		I am grateful to know the PLD will be made available for teachers for the next school year. It will help in thinking about lesson to determine if they are meeting the needs of the skill set.
		I understood our rating process well and it was easy to work with
		I liked being part of the process and learning about the PLD and how the assessment is scored.
		The discussion part was super helpful for clarity. It was great to have mix of different grade levels to appreciate different perspectives.
29	Please note any other feedback you would like us to consider.	The ID matching process and use of the standard setting toolkit was a good concept.
		It was very helpful to have SDE and Breanne here to explain and answer questions that we needed. Our facilitator, Karen, did a wonderful job of politely and patiently getting everyone back on task and recapping the discussion. She was really good at taking our questions and finding the correct person to ask to answer that question.
		Thank you for this informative
		trying to hear how our rating impact the students final score was foggy
		I did not feel like there was equal "air time" given to each person on the committee to speak. There was a lot of interruption and being talked over.

Table 9. Mathematics Panel Grades 5 & 6 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	0	12
2	I understood the procedures we followed to set standards.	0	0	0	0	12
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	0	12
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	2	10
5	I am confident about my understanding of this standard setting process.	0	0	0	3	9
6	The workshop facilitator explained things clearly to us.	0	0	0	0	12
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	0	12
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	0	12
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	0	12
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	0	12
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	0	4	8
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	0	0	0	2	10
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item and matching those item-response demands to PLDs.	0	0	0	2	10
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	1	11
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	0	12
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	1	11
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	1	0	11
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	1	0	11
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for Mathematics GRADE 5, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	8	3	0	1	0
20	Based on the impact data results for Mathematics GRADE 5, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	6	5	0	1	0
21	Based on the impact data results for Mathematics GRADE 5, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	3	8	1	0
22	Based on the impact data results for Mathematics GRADE 5, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	9	2	1	0
23	Based on the impact data results for Mathematics GRADE 6, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	10	1	0	1	0
24	Based on the impact data results for Mathematics GRADE 6, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	6	3	2	1	0
25	Based on the impact data results for Mathematics GRADE 6, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	4	7	1	0
26	Based on the impact data results for Mathematics GRADE 6, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	3	8	1	0

Table 10. Mathematics Panel Grades 5 & 6 – Text Responses for Open-ended Questions

Question #	Question Text	Responses
27	Please indicate any parts of the standard setting training and process that we should improve.	Sometimes, there is a lot of down time during the day. I am not sure if that can be fixed or modified, but it can be frustrating to feel like there is nothing to do.
		Maybe take a moment at the beginning to look at some of the work that has been done already with item development and PLD development so that people who may question some of these artifacts would have a better understanding of them and how they are formed.
		PLD need aligned to specific standard a little tighter or split to a standalone instead of 2 or more standards on one line.
		I left feeling like I really didn't have enough information to talk sensibly about the cut score that was set. I really enjoyed the process and know that what I have learned will help instruct my teaching, but I would like to be able to help my district more. I am not looking for a magic wand just some guided help.
		Maybe timing, but it wasn't bad, having extra time as a group was nice
		Nothing to improve at this time.
		The only confusion I noticed was a result of not addressing how DOK of questions relates to this process.
		The training was well done. The information was introduced the first day and then our facilitator built on that. She answered any questions. She did a fantastic job.
		None
		I would like to see more items presented to the students so that that the Below Basic is not so easy to attain, and I would like the Advanced items to be more available.
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	Overall, I think it was a successful meeting from my POV as a participant.
		Discussions about application of PLD
		Training on using the PLDs to make content-based decisions.
		I felt like the timing allowed worked really well. The presenter was well versed in what we were doing. I enjoyed the experience
		Being allowed to have a voice and have the panel listen. To have a better understanding of testing
		Our facilitator was amazing!
		I thought the open discussions at the end of each round worked really well.
		Our facilitator kept us going. Kept our room positive and on task. It's hard to keep a room of teachers on task and not talking. ha-ha She was well prepared for that. Loved her.
		The discussions when we were we able to state our viewpoint and hear others' viewpoints were very helpful.
		None
29	Please note any other feedback you would like us to consider.	This was a very interesting and informational experience. I think that the facilitator, Katie, was perfectly chosen because of her bubbly personality. She made everyone feel comfortable to express any concerns, questions, or thoughts. I feel Mathematics grades 5-6 were very fortunate to have her be our facilitator because she made the environment so welcoming. I also feel confident in the fact that I know my knowledge of HOW to do everything was correct; I was properly trained.
		less spicy food
		I feel like there is still a disconnect in communication of the students' performance converted to the score. I would personally like to see, not just a summary of the data, but the actual data being summarized. I also would like to think about how we are communicating this information to others, there seems to be a general idea that we do not need to understand the inner "magical" workings of the psychometrics when that is exactly what we need to understand. Questions about the process were often partially answered or dismissed by the psychometrics people as though we may not be able to understand.
		Please consider a crash course in how to decipher the cut scores so that we can better help those in our district.
		Katie is the best!! She set the tone for the week. Her friendliness and passion was infectious. Everyone involved seemed to have the passion.
		I would love to attend a workshop or continuing education to help me understand the statistics that are used to move forward.
continued		

Question #	Question Text	Responses
29	Please note any other feedback you would like us to consider.	Thank you so much for including the classroom teachers. It helps to know we are heard. Thanks again.
		Questions 19 to 26 are difficult to answer. Changing the impact level may increase or decrease students from a category, but I feel that lowers the level of where our students truly should be.
		I really hope to be able to come back to do more Standard Settings, IRW, PLD reviews, etc. I am very thankful for everything that Cognia/OSDE has done for me here. Thank you for letting me be a part of this very important process.
		I would love to be part of the standard setting panel. I wish we emphasized more on number operations and less on algebraic reasoning in elementary. We seem to reteach the same thing year after year, (fractions for instance). Students need more time for mastery of number operations and number sense before being introduced to algebraic reasoning. I also wish more emphasis would be placed on using correct mathematics terms. I saw places in our PLD's this week where mathematics terminology needs to be looked at (numerical expression vs algebraic expression). Correct terms should be in the PLD's if we expect teachers to know exactly what the standard is. The PLD's are for the teachers, not the students.

Table 11. Mathematics Panel Grades 7 & 8 - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the standard setting workshop.	0	0	0	0	12
2	I understood the procedures we followed to set standards.	0	0	0	0	12
3	I understood that my role was to make content-based judgments about the alignment between the items and the performance level descriptors.	0	0	0	0	12
4	The workshop procedures made sense to me, and I learned how to apply them efficiently.	0	0	0	1	11
5	I am confident about my understanding of this standard setting process.	0	0	0	2	10
6	The workshop facilitator explained things clearly to us.	0	0	0	0	12
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	0	12
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	0	0	12
9	The workshop facilitator took steps to help the standard setting process run smoothly.	0	0	0	0	12
10	Sufficient time was allotted for training and practice on the standard setting concepts, tasks, and procedures.	0	0	0	0	12
11	I understood the progressions in expectations across the Basic, Proficient, and Advanced performance levels as defined by the Performance Level Descriptors.	0	0	1	3	8
12	I became sufficiently familiar with the assessment to make item-PLD judgments, based on responding to items on the test and considering the knowledge, skills, and abilities required by the items.	0	0	0	2	10
13	I understood the ID Matching task, including considering the knowledge, skills, and abilities required by each item and matching those item-response demands to PLDs.	0	0	0	1	11
14	I understood how to use the standard setting tool to record my responses regarding skills and notes as instructed.	0	0	0	0	12
15	I understood how to use the standard setting tool to record my item-PLD alignment judgments.	0	0	0	0	12
16	I understood how to use the feedback after round 1, in preparation for round 2.	0	0	0	0	12
17	I understood what the content-based benchmarks, introduced in round 2, represented.	0	0	0	1	11
continued						

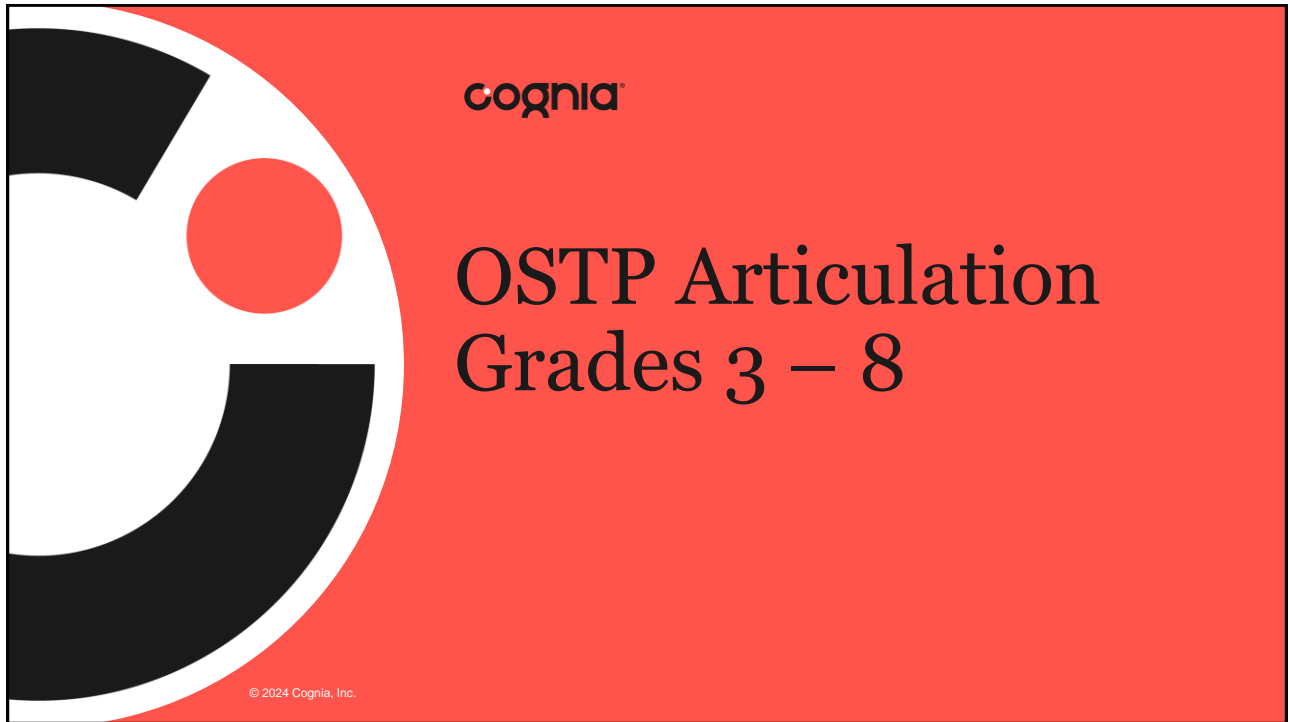
Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
18	I understood how to consider the content-based benchmarks in rounds 2 and 3, as I made my item-PLD alignment judgments.	0	0	0	1	11
Q#	Question Text	Less	About the same	More	Unsure	Not Applicable
19	Based on the impact data results for Mathematics GRADE 7, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	8	2	0	1	1
20	Based on the impact data results for Mathematics GRADE 7, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	3	4	4	0	1
21	Based on the impact data results for Mathematics GRADE 7, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	3	8	0	1
22	Based on the impact data results for Mathematics GRADE 7, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	10	1	0	1
23	Based on the impact data results for Mathematics GRADE 8, do you feel the percentage of students in the BELOW BASIC category should be less, about the same, or more?	8	2	0	2	0
24	Based on the impact data results for Mathematics GRADE 8, do you feel the percentage of students in the BASIC category should be less, about the same, or more?	3	4	4	1	0
25	Based on the impact data results for Mathematics GRADE 8, do you feel the percentage of students in the PROFICIENT category should be less, about the same, or more?	0	4	8	0	0
26	Based on the impact data results for Mathematics GRADE 8, do you feel the percentage of students in the ADVANCED category should be less, about the same, or more?	0	10	2	0	0

Table 12. Mathematics Panel Grades 7 & 8 – Text Responses for Open-ended Questions

Question #	Question Text	Responses
27	Please indicate any parts of the standard setting training and process that we should improve.	<p>Add the instructions / information about how to consider the benchmark data to the slide that is displayed during judgement for round 2 and round 3.</p> <p>Display the panelists round results bar graph in colors that are considerate to ADA/color blind participants; examples could be adding a pattern or displaying in shades of gray.</p> <p>Provide a printed copy of the panelists round results bar graphs for review during discussion; they could be handed back in during judgement if deemed too influential</p> <p>The original time sent to participants was 9:00 - 4:00; the week before the training an updated schedule was an additional hour and a half, 8:30 - 5:00. For participants traveling daily, a week before may not be enough time to adjust their schedule with kids and other family member s.</p> <p>Recognize the Juneteenth federal holiday and not have work on that day.</p> <p>I thought taking section 1 of the test before matching PLDs was extremely helpful. I wish we would have also done this for grade 7 as well instead of limiting it to just grade 8.</p> <p>In future please make all graphs color blind friendly both on screen and on projectors where color washes. The graphs at the end of each round were difficult for me to visually follow due to the yellow/green merging visually.</p> <p>The PLDs could be copied not front and back so you don't have to flip back and forth. The graph after round 1 was not easy to read for color blind individuals.</p> <p>The panelist round results bar graph is not able to be read by those who have a visual impairment (color blind, poor sight), It would make it easier if it was printed out or show on each individual computer. Having non-carbonated drink options for breakfast and lunch are important for those who do not drink soda. Water is great but juice, tea, flavored stuff is great too. Afternoon snacks should have non sugar options each day.</p> <p>On the PLD tool, I would like the Strand Descriptions at the top of each page, and I would like each category on a single page. (Less flipping)</p>
		continued

Question #	Question Text	Responses
27	Please indicate any parts of the standard setting training and process that we should improve.	<p>All in all, this was the BEST standard setting I have ever participated in as a very old teacher I have seen several different testing companies and numerous different SDE staffs. If I could have one suggestion, it would be on the way the data is presented on the Panelist Round Results Bar Graph. The yellow and green are too similar for some eyes.</p> <p>I would like to see the panelist round results while I am going through round 2 and round 3. It would help me make choices as I re-read the items and revisit my judgements. The colors on the round results bar graph could be different colors from yellow and green. They were very hard to distinguish on the screen.</p> <ol style="list-style-type: none"> 1. Consider panelists who may have needs such as color-blindness or hard of hearing. 2. Consider flexible seating options within the panel room. Sitting for long periods of time can make it difficult to focus. I would have loved to have the option to sit by myself to focus more during independent work time. 3. In Grade 7 mathematics, items 30-31 would be great TEI items! 4. When doing PLD work, be mindful of wording in the sentences. We had several conversations about what the intention of the sentence was. Be clear and concise. Fewer sentences is not necessarily better. 5. As the scores were explained to us, it would be nice if SDE could give guidance to parents, stakeholders and administrators about the scores. I think a big misunderstanding is that students who score below basic or basic only got "x" amount of questions correct. <p>n/a I was very impressed with the whole process</p>
28	Please indicate any parts of the standard setting training and process that you felt worked really well.	<p>I enjoyed the variety of food and snack options daily. All the tech set-up worked well for participants.</p> <p>The psychometrists, workshop facilitators, content specialists, SDE & other observers, were all knowledgeable and helpful when asked for clarifications or information.</p> <p>I thought Round 1 and the discussion process after Round 1 was the best part. It was the most insightful and impactful portion to decision making.</p> <p>I felt the information given was succinct and easily followed. As we progressed, we were better able to connect instructions to our actions.</p> <p>The process as a whole was very straight forward and made sense. the directions were also clear</p> <p>Jill was amazing about being a facilitator. She was pleasant and made sure that we stayed on task as well as everyone's voice was heard. Bri and Sandra also were amazing.</p> <p>I felt very good about all of it. Jill did an excellent job training each of us. Our panel had great discussions each time we discussed.</p> <p>I feel that the Cognia and SDE staff did an excellent job in preparing us for the task before we began. I also felt they did a phenomenal job of answering our questions as we went through the process. They did this while carefully assuring that they were not influencing anyone. Jill was a fabulous facilitator. She kept everyone moving forward and reminded not to try to influence others. Bri is exceptionally knowledgeable and was a terrific asset when we had questions about PLD language.</p> <p>The training was beneficial on day one in the opening session and in our 7/8 room. Jill did a great job keeping us on task and helping us focus on discussion on the task at hand. The mathematics specialist the joined our rooms were very helpful and answered all of our questions to the best of their ability.</p> <p>The process was very well organized and efficient. Jill did a great job of keeping us fair and ensuring that we all felt heard.</p> <p>Jill was a superior moderator. She kindly kept us on task and was extremely professional and personal at the same time</p> <p>Staff was very helpful and responsive to all our questions.</p>
29	Please note any other feedback you would like us to consider.	<p>For Mathematics standard 7.D.1.1, there was no proficient category; should this be a standard for 7th grade if students must be advanced in their understanding? Will other opportunities to continue in this type of work be sent to participants as they occur?</p> <p>Jill was an AMAZING asset to have as a facilitator. The process would not have gone as well without her.</p> <p>Several of the questions would have made some actually awesome technology assisted answers. GR7 Item 31 for example could use a drag/drop to put parenthesis. Jill was awesome. Bri helped many times, psychometricians were all super helpful.</p> <p>The temperature versus humidity made it hard to focus at times in our meeting room.</p> <p>Sandra and Qui did great at explaining what all our work was going to be used for. EVERYONE from Cognia, SDE, outside observers, and hotel staff were courteous and helpful. I felt very supported and appreciated!!!!!!</p> <p>Bre was very knowledgeable as well. This week was a great learning experience for me.</p>

APPENDIX—L
ARTICULATION POWERPOINT PRESENTATION



1



Articulation agenda

- Introductions, meeting norms, and overview
- The “why” and “how” of the articulation process
- The Consensus Process for Articulation
- Modeling our standard setting panel decisions
- Familiarization with standards, blueprints and PLDs
 - Across unfamiliar grades
- Expectations for between-grade transitions
- Presentation of Impact Data and discussion
- Recommendations (if any) for adjustments



2

Welcome & Introductions - Panelists

- Introduce yourself:
 - Name
 - District
 - Which grade-band you were with during standard setting
 - Grades and content areas you've taught
 - Fun fact about yourself?



3

Meeting Norms

- All conversations are **confidential**.
- What happens here, stays here.
- Outside of this meeting, please **DO** talk about the process we undertake, but **DO NOT** disclose the specifics.
- Please **DO NOT**:
 - use any personal devices in the room; you may step out at any time if needed.
 - use the Chromebooks for anything other than standard setting or articulation activities.



4



Overview

Our shared goals:

- collect your recommendations on performance standards for OSTP ELA or Math assessments that provide meaningful and actionable information.

Your goals as panelists:

- adapt to forming consensus recommendations.
- listen carefully to your fellow panelists.
- make content and student-based judgments about the rigor of grade-to-grade transitions.
- rely on your expertise about the content standards, blueprints, PLDs and student learning throughout the process.



5



Purpose

Capture panelist expectations for differences in rigor between grades

- Does student performance on the test, calculated with the new cut scores, align with those expectations?
- If they don't align, how are they different?
- Use educator expectations to assess the reasonableness of the cut scores
- Recommend adjustments to smooth differences between grade panels
- Inform policy decisions regarding the rigor of the OSTP assessment



6

Articulation process: The “why”

- Why do we want to **COMPARE** the challenge of demonstrating proficiency for students in different grades?
 - Each of our panelists and facilitators are different (thank goodness)
 - On a different day, with different people and different facilitators (reviewing different items) there would likely be different judgments. That's okay and expected!
 - We know each grade has greater expectations in general (that's learning!), but...
 - We want to compare the challenge for a 5th grader (for example) who has had a full year of 5th grade instruction and development to a 6th grader!



7

Articulation process: Comparing Rigor



Is 5th grade more challenging
for a 5th grader than 6th grade is
for a 6th grader?



8

Articulation process: The “why” (cont.)

- Once we capture those expectations, the panel will look at impact data.
 - The percentage of students in each performance level using the cuts we developed this week.
- You'll compare your expectations to those empirical percentages
- You'll arrive at consensus advice to inform policymakers where the panel thinks those percentages don't fully agree with your expectations for rigor.
- BECAUSE we want to smooth the variation of different panel results to align with your expectations.



9

Articulation process: The “how”

- Review previous PLD alignments for select items.
- Review unfamiliar PLDs, standards and blueprints.
- Determine expectations for transition between grades based on content demands as reflected in PLDs, standards, and blueprints.
- Review impact data based on standard setting cut scores and compare these results to the expectations identified in the previous step
- Recommend adjustments



10

Articulation is an Advisory Process

Item-student
Judgments

Apply Your
Expertise

Consensus
Judgments



11

Content-based Judgment - Overview



Useful

- Standards and PLDs
- Blueprints
- Compare rigor between grades
- How students progress through each grade



Not Useful

- Compare rigor between grades for the same student
- Your aspirations or concerns regarding student test scores



12

Reviewing Previous PLD Alignments

- We will present some items and judgments from the standard setting panels
- Panelists who worked on an item during standard setting will present their reasoning for the item-PLD alignment
- We will look at one item from each grade-band (3-4,5-6,7-8)
- Our goal is to become familiar with the judgment tasks from unfamiliar grade-bands



13

Items for Review

- We will review one item each from grades 4, 5, and 7, respectively.
- Starting with the 4th grade item, we will look at the item in the Toolkit
 - Panelists from the 3-4 panel will summarize their PLD alignment & reasoning for the item
 - Panelists from other panels comment and ask questions
 - Repeat for the other two items



14

Blueprints: Review Across Grades

Grade	Reading & Writing Process	Critical Reading & Writing	Vocabulary	Language	Research
3	38 – 42 %	12 – 18 %	22 – 26 %	12 – 18 %	12 – 18 %
4	30 – 34 %	18 – 22 %	22 – 26 %	12 – 18 %	12 – 18 %
5	30 – 34 %	22 – 26 %*	18 – 22 %	12 – 18 %	12 – 18 %
6	34 – 38 %	18 – 22 %	18 – 22 %	12 – 18 %	12 – 18 %
7	34 – 38 %	18 – 22 %	14 – 20 %	12 – 18 %	14 – 20 %
8	24 – 30 %	24 – 30 %*	14 – 20 %	12 – 18 %	12 – 18 %



15

Standards: Independent Review & Discussion

- Review the **standards and PLDs** across grades 3 – 8.
- Consider differences and progressions across the grades
- Discuss findings with the group.



16

You're familiar with the standard setting process

1. Review the item and identify the KSAs

- Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.

What does a student need to know or be able to do to correctly respond to this item?

2. Make an item-PLD alignment judgment

- Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced performance level descriptor (PLD).

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?



17

Now consider what it means to demonstrate KSAs from one grade to the next

1. Review unfamiliar PLDs, standards and blueprints

- Consider how rigorous the demands are for a student in this grade

How challenging are these PLDs, blueprints, and standards for a student in one grade?

2. Consider how rigorous the content demands of the next grade are for a student in the next grade.

- Example: Is it more, less, or about the same difficulty for a 4th grader to demonstrate proficiency on 4th grade standards than it is for a 4th grader to demonstrate proficiency on 3rd grade standards?

Compared to the PLDs, blueprints and standards for a student in the next grade



18

What are we looking for?

- How do the standards and expectations for students at performance levels change from grade to grade?
 - How do the verbs change?
 - How do the students change from grade to grade?
 - Does your expectation for the pace of learning align with the change in standards and performance level expectations?
- We will review and discuss five transitions
 - Transition from grade 3 to 4, grade 4 to 5, grade 5 to 6, grade 6 to 7, and grade 7 to 8.



19

For each of five grade transitions

- Review the blueprints, standards, and PLDs, blueprints for the proximal grades
- Answer guided questions by considering
 - Differences in standards
 - Blueprints: % of items in domains
 - PLDs: Verbs, etc.
- We will make a consensus judgment
- Facilitator will take notes on the discussions



20

Transition between Grades 7 and 8

- How much more or less challenging is it for 8th graders to demonstrate proficiency in an 8th grade test (blueprint), assessing 8th grade standards, as described by 8th grade PLDs

THAN IT IS

- For 7th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



21

Transition between Grades 7 and 8

- Do we expect a similar difference for other performance levels?
 - Basic
 - Advanced
- If not, what are the expected differences?
- Provide our reasoning for our expectations to help inform policy makers

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



22

Transition between Grades 6 and 7

- How much more or less challenging is it for 7th graders to demonstrate proficiency in a 7th grade test (blueprint), assessing 7th grade standards, as described by 7th grade PLDs

THAN IT IS

- For 6th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



23

Transition between Grades 6 and 7

- Do we expect a similar difference for other performance levels?
 - Basic
 - Advanced
- If not, what are the expected differences?
- Provide our reasoning for our expectations to help inform policy makers

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



24

Transition between Grades 5 and 6

- How much more or less challenging is it for 6th graders to demonstrate proficiency in a 6th grade test (blueprint), assessing 6th grade standards, as described by 6th grade PLDs

THAN IT IS

- For 5th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



25

Transition between Grades 5 and 6

- Do we expect a similar difference for other performance levels?
 - Basic
 - Advanced
- If not, what are the expected differences?
- Provide our reasoning for our expectations to help inform policy makers

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



26

Transition between Grades 4 and 5

- How much more or less challenging is it for 5th graders to demonstrate proficiency in a 5th grade test (blueprint), assessing 5th grade standards, as described by 5th grade PLDs

THAN IT IS

- For 4th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



27

Transition between Grades 4 and 5

- Do we expect a similar difference for other performance levels?
 - Basic
 - Advanced
- If not, what are the expected differences?
- Provide our reasoning for our expectations to help inform policy makers

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



28

Transition between Grades 3 and 4

- How much more or less challenging is it for 4th graders to demonstrate proficiency in a 4th grade test (blueprint), assessing 4th grade standards, as described by 4th grade PLDs

THAN IT IS

- For 3rd graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



29

Transition between Grades 3 and 4

- Do we expect a similar difference for other performance levels?
 - Basic
 - Advanced
- If not, what are the expected differences?
- Provide our reasoning for our expectations to help inform policy makers

1. Much less challenging
2. Less challenging
3. About the same
4. More challenging
5. Much more challenging



30

Expectations compared to Standard Setting results

- We have captured our consensus expectations on a white board here in the room
- We will look at impact data based on Standard Setting cut scores
 - This data shows us what percentage of students we would expect in each performance level for each grade
- Compare the impact data to our consensus expectations. Do they match expectations?
 - If not, discuss and make recommendations for adjustments
 - Our facilitators will capture notes on the discussion and recommendations



31

For each grade

- Review impact data
- Consider the expectations we identified
- Answer the following question:

Do we think the percentage of students in the proficient and above category should be...

1. Much less
2. Less
3. About the same
4. More
5. Much more



32

APPENDIX—M
ARTICULATION EVALUATION SURVEY & RESULTS

Table 1. ELA Articulation - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the articulation workshop.	0	0	0	5	5
2	I understood the procedures we followed to advise policymakers on cut recommendations.	0	2	0	3	5
3	I understood that my role was to communicate educator expectations regarding the progression of rigor in student transitions from lower to higher grade-levels in my content area of expertise.	0	0	0	5	5
4	The workshop procedures made sense to me.	0	1	1	4	4
5	I am confident about my understanding of our consensus recommendations	0	1	1	5	3
6	The workshop facilitators explained things clearly to us.	0	2	0	5	3
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	3	7
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	1	0	0	5	4
9	The workshop facilitator took steps to help the process run smoothly.	1	0	0	4	5
10	Sufficient time was allotted for training and discussion.	1	1	1	3	4
11	I understood the progressions in expectations across grade-levels for Oklahoma students.	0	0	2	3	5
12	I became sufficiently familiar with blueprints, standards and PLDs for each content area to help inform our consensus recommendations to Oklahoma policymakers.	0	1	0	7	2
13	Our facilitators captured notes for our discussion that represented our process to arrive at consensus recommendations.	0	0	2	5	3
14	My expertise and input helped our group arrive at our consensus recommendations.	0	1	1	4	4

Table 2. ELA Articulation Panel – Text Responses for Open-ended Questions

Question #	Question Text	Response
15	Please indicate any parts of the articulation training and process that we should improve.	I thought that the articulation training process was well done. The only note I have is that I thought that we were rushed to look through the standards and PLDs. I wished that we had a some more time to look through them especially for the grade levels that we are not familiar with.
		There should be more time to familiarize ourselves with the standards AND blueprints (these were not available to us). More time for discussion as well so that we could really dig in and analyze differences in the standards across grade levels.
		I believe that the articulation training and process would be more beneficial if educators were given more opportunities to view assessments, or assessment questions, across grade levels. Only working with 2 grades does not allow me to fully capture what the other grade levels are attempting to accomplish.
		Maybe more data ahead of time would alleviate the outrage
		Please make sure that people listen and stay on task.
		The meeting today was brief, so I think we needed more time to flush out ideas.
		This felt like it should have been an important process, but the allotted time was not enough to actually get valuable data. I am not confident at all in the consensus and many of the other panelists were very confused and therefore the graph that the facilitator made did not match was on the board. I cannot perceive how this information could be valuable. With such a small group and such little time, the data gathered during vertical articulation seems like it will be damaging to the process. I really enjoyed Sandra's explanations and felt that she explained things very well and helped to correct several confused panelists.
		The workshop procedures and expectations could have been explained better, clearer. Time should have been allotted to give teachers opportunities to ask questions about the articulation workshop process.
16	Please indicate any parts of the articulation training and process that you felt worked really well.	I felt like the overall organization and flow worked really well. I also like the process used. The only breakdown I felt there was, is that teachers were hesitant to put in graph form the idea that 6th-8th grade proficiency should be less than in already was on the graph. Our facilitators were great and patient in helping us to dig through and overcome challenges we experienced.
		I am thankful that educators across grade levels and state were provided the opportunity to bring their expertise to the articulation training and process. It was also beneficial to deep dive into the state standards and the PLDs to determine the differences.
		I liked seeing the data and seeing that scores are adjusted so that we have a better idea of how the students are taking to the standards
		I thought that the open discussion parts were well done and that everyone respected each other's thoughts and opinions.
		The expertise of the facilitators was most impactful for me.
		Sandra is great, easy to understand and communicate with.
		Monday through Thursday worked very well when we were in our 5-6 group with Lisa who kept us on task and focused!
		Including teachers' perspectives and opinions, and relevant teaching experiences was valuable and appreciated.
continued		

Question #	Question Text	Response
17	Please note any other feedback you would like us to consider.	<p>I would love to participate in more!</p> <p>Thank you for incorporating teachers into the process. I believe that if a group of educators were given the opportunity to work across all of the grade level standards and assessments then the articulation process would have run smoother.</p> <p>I have no other notes at this time.</p> <p>If this is going to be a part of the process, it should be over several days with a larger group of teachers. We should be given more direction, have more time with the standards, and be comfortable with the items. If student experience is going to be considered, there should be social science data provided as well as past test scores.</p> <p>Thank you for including actual teachers in the process.</p> <p>I would recommend the articulation workshop being longer maybe, one full day to 2 days in length.</p> <p>Setting norms and expectations prior to meeting.</p> <p>I really enjoyed this experience. I know I have learned and grown a lot through this experience.</p>

Table 3. Math Articulation Panel - Frequency of Responses for Likert-type Questions

Q#	Question Text	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree
1	I understood the goals of the articulation workshop.	0	0	1	6	4
2	I understood the procedures we followed to advise policymakers on cut recommendations.	0	0	1	6	4
3	I understood that my role was to communicate educator expectations regarding the progression of rigor in student transitions from lower to higher grade-levels in my content area of expertise.	0	0	0	7	4
4	The workshop procedures made sense to me.	0	0	5	3	3
5	I am confident about my understanding of our consensus recommendations	0	0	2	6	3
6	The workshop facilitators explained things clearly to us.	0	1	1	5	4
7	The workshop facilitator encouraged us to raise questions and put our understandings into our own words.	0	0	0	6	5
8	The workshop facilitator provided clear and helpful responses to my questions and other requests for clarification.	0	0	2	6	3
9	The workshop facilitator took steps to help the process run smoothly.	0	0	0	8	3
10	Sufficient time was allotted for training and discussion.	0	5	0	4	2
11	I understood the progressions in expectations across grade-levels for Oklahoma students.	0	0	1	6	4
12	I became sufficiently familiar with blueprints, standards and PLDs for each content area to help inform our consensus recommendations to Oklahoma policymakers.	0	0	0	6	5
13	Our facilitators captured notes for our discussion that represented our process to arrive at consensus recommendations.	0	0	0	5	6
14	My expertise and input helped our group arrive at our consensus recommendations.	0	0	1	6	4

Table 4. Math Articulation Panel – Text Responses for Open-ended Questions

Question #	Question Text	Response
15	Please indicate any parts of the articulation training and process that we should improve.	n/a
		None
		It was hard for some participants to stay on task and wanted to solve much larger issues rather than answer the questions we had to answer right now. It may have been easier to agree on difficulty changing grade to grade if we got to experience the test rather than just looking at the PLDs.
		A more formal way of having discussions, people were talking over other people and having side conversations.
		I understand we were ahead of schedule, but moving Friday to Thursday afternoon did make it feel rushed. I would also have liked more time to process through the other grade level PLDs before this meeting, if possible.
		I think more time to compare the grade level PLDs and standards before being asked to compare them.
		I did not feel totally clear on what some of the procedures were or maybe more so where they were going.
		I am honestly not sure
		I would like to see an improvement in how the recommendations are made. Unfortunately, by the time we got to 8th grade we were out of wiggle room for it to make sense.
16	Please indicate any parts of the articulation training and process that you felt worked really well.	None
		I enjoyed the process...I learned a lot about testing and scores.
		vertical alignment was beneficial
		I appreciate it is a smaller group.
		We were able to eventually come to a consensus on most points.
		grouping the teachers by having a mixture of the groups
17	Please note any other feedback you would like us to consider.	Everyone is able to share.
		I enjoyed the process
		Honestly, I do feel that overall, the cut scores, though better than say last iteration, I do feel it still does a disservice to Oklahoma students.
		I enjoyed the process but feel totally overwhelmed with the responsibility we were given. I don't feel like I was totally comfortable covering standards in the articulation process that I don't teach. It takes me a while to process things, and I don't feel like I had enough time to do that.
		I have enjoyed this entire process and the dialog that has transpired. I truly feel that I have grown from all of the collaboration that has occurred

APPENDIX—N
STANDARD-SETTING MEMO

Oklahoma Standard Setting Memo

OSTP ELA and Mathematics Grades 3 - 8

June 17-21, 2024

Overview

Cognia and the Oklahoma State Department of Education (OSDE) convened six panels of ELA and Mathematics educators during June 17–21, 2024, to establish Basic, Proficient, and Advanced cut scores to enable reporting of student performance on the OSTP ELA and Mathematics Grades 3 – 8 assessments. Each panel included 10–12 educators from around the state and completed the standard setting activities for two grades, starting with the upper grade in their respective panels. The standard setting panelists reviewed test content and performance level descriptors and followed the modified Item-Descriptor (ID) Matching standard setting method. The standard setting portion of the meeting was conducted over the first four days of the meeting from Monday, June 17 to Thursday, June 20. At the conclusion of the standard setting portion, two articulation panels (one each for ELA and Mathematics) were convened to complete a half day of articulation activities across all grades within their respective content areas. The articulation panelists included three–four panelists from each of the original standard setting panels.

The purpose of this memo is to present the results from the standard setting and articulation meeting, including cut scores and associated impact data.

Methods

Standard Setting Procedure

During the standard setting meeting, the panelists were trained on and followed the modified ID-Matching method. Each panelist reviewed each item in a content and grade-specific ordered item booklet (OIB) and considered the knowledge, skills, and abilities required by the item. Panelists then matched those item-response demands to the knowledge and skill expectations in the performance level descriptors for the Basic, Proficient, and Advanced levels. Working independently, the standard setting panelists conducted the ID matching process over three rounds and made item-PLD alignment judgements for each item. Before each round, panelists completed a round readiness survey. After rounds 1 and 2, the Cognia workshop facilitator led panelists through a discussion of agreements and disagreements among the panelists and rationales for their various item-PLD alignment judgements. The ensuing discussion enabled panelists to consider their colleagues' insights about item response demands and rationales for matching items to descriptors, and to consider adjusting their judgements in rounds 2 and 3.

At the beginning of round 2, content-based benchmarks were introduced to panelists, which served as additional information for panelists to consider as they made their item-PLD alignment judgements in rounds 2 and 3. Panelists completed the activities for two grades, beginning with the upper grade in their respective panels. At the completion of both grades, standard setting cut scores were calculated and the associated impact data for both grades were presented to panelists within their respective panels. Impact data are the percentages of students who would be sorted into the Below Basic, Basic, Proficient, and Advanced performance levels, using their scores from the 2024 administration of the OSTP ELA and Mathematics grades 3-8 assessments. Panelists then completed a final evaluation survey about their overall experience with the standard setting workshop, as well as their opinions on the results (impact data) presented.

Analyses Procedure

During the standard setting meeting, a subject matter expert (SME) reviewed the qualitative data for panelists as the data became available. Specifically, the SME reviewed panelists' notes on the knowledge, skills, abilities required by the items, as well as their reasoning notes to determine if the panelists were on task.

Additionally, Cognia psychometricians conducted statistical analyses of panelists' item-PLD alignment data by calculating the percent exact, adjacent, and discrepant for each panelist on each performance level.

At the conclusion of Round 3 for each grade, Cognia psychometricians conducted initial logistical regression analyses. Since the logistical regression method is sensitive to statistical outliers and the presence of such outliers violates the assumptions of the model, outlier analyses were performed in the form of visual inspection of the initial logistic regression curves. Statistical outliers were identified, and the associated data points were removed and then the final logistic regression analyses were conducted to calculate the proficient and advanced cut scores. After calculating the proficient and advanced cut scores, the TCC method was used to calculate the Basic cut score.

Finally, the resulting cut scores were applied to student data from the spring 2024 administration of the OSTP ELA and Mathematics grades 3-8 assessments to calculate the impact data (i.e., the percentage of students that would be classified into each performance level based on the standard setting cut scores).

Articulation Procedure

At the conclusion of the standard setting meeting, an articulation panel was convened for each content area. Three to four panelists from each of the original standard setting panels participated in the articulation meeting. During the articulation meeting, panelists engaged in a cross-grade qualitative review of test blueprints, standards, and PLDs. In a consensus-based process and based on their review, panelists then identified performance expectations for transitions between grades (i.e., whether it is more or less challenging for a student in grade 4 to reach proficiency on the 4th grade assessment, than it is for a student in grade 3 to reach proficiency on the 3rd grade assessment). After identifying the performance expectations across grades, panelists review impact data based on the standard setting cut scores in comparison to the expectations identified in the previous step. Finally, panelists made consensus-based recommendations for adjustments. The meeting concluded with an articulation workshop survey.

Results

This section details the results from the standard setting and articulation meetings and is organized by content area, starting with the ELA grades 3–8 results.

ELA Grades 3–8: Standard Setting Results

Table 1 shows the three cut scores (basic, proficient, and advanced) for each ELA grade that resulted from the standard setting meeting and analyses. The table includes the OIB page range, theta, and associated standard error for each cut. In addition, the same information is presented based on the benchmark cut scores. Finally, the prior (pre-standard setting) theta cut scores are also listed for reference.

Table 1. OSTP ELA Grades 3-8 Cut Score Details based on Standard Setting, Benchmarks, and Prior

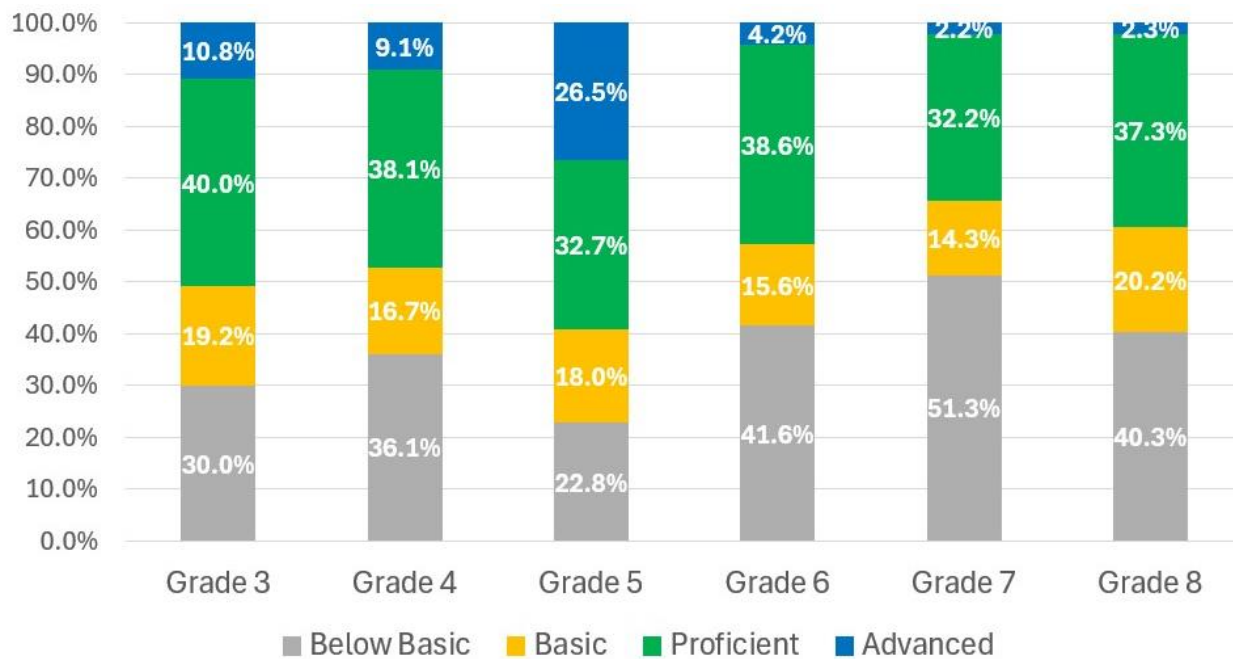
Subject Grade	Performance Cut Placement Level	Standard Setting			Benchmarks			Prior Theta
		OIB #	Theta	Standard Error	OIB #	Theta	Standard Error	
ELA 03	Basic	3 - 4	-0.890	--	6 - 7	-0.600	--	-0.531
	Proficient	11 - 12	-0.288	0.035	16 - 17	-0.102	0.114	0.341
	Advanced	41 - 42	0.949	0.042	45 - 46	1.667	0.609	1.396
ELA 04	Basic	4 - 5	-0.700	--	4 - 5	-0.670	--	-0.527
	Proficient	17 - 18	-0.225	0.042	14 - 15	-0.432	0.186	0.386
	Advanced	35 - 36	0.941	0.043	34 - 35	0.903	0.166	1.499
ELA 05	Basic	5 - 6	-1.120	--	5 - 6	-0.830	--	-0.783
	Proficient	11 - 12	-0.531	0.042	32 - 33	0.000	0.102	0.325
	Advanced	42 - 43	0.315	0.038	50 - 51	0.948	0.311	1.172
ELA 06	Basic	2 - 3	-0.670	--	8 - 9	-0.280	--	-0.909
	Proficient	9 - 10	-0.232	0.044	19 - 20	0.051	0.267	0.285
	Advanced	45 - 46	1.222	0.059	48 - 49	1.552	0.347	1.392
ELA 07	Basic	8 - 9	-0.380	--	8 - 9	-0.470	--	-0.498
	Proficient	15 - 16	0.015	0.070	17 - 18	0.139	0.152	0.467
	Advanced	47 - 48	1.551	0.124	47 - 48	1.599	0.436	1.259
ELA 08	Basic	8 - 9	-0.740	--	8 - 9	-0.570	--	-0.695
	Proficient	10 - 11	-0.207	0.068	16 - 17	0.061	0.244	0.451
	Advanced	50 - 51	1.351	0.172	50 - 51	1.606	0.524	1.208

Table 2 shows the impact data (percentage of students classified in each performance level) for each ELA grade based on the cut scores from the Standard Setting meeting and benchmarks. In addition, impact data based on the prior (pre-standard setting) cut scores are listed for reference. Note that percentages related to the standard setting, benchmark, and prior cut scores were calculated by applying the cut scores to student data from the Spring 2024 OSTP ELA test administration. Finally, where relevant, percentages based on NAEP data for Oklahoma are also shown. The NAEP data are based on the 2022 test administration for “Reading” and represent the most recent data available ([NAEP OK State Profile Website](#)).

Figure 1 gives a visual representation of the impact data based on the ELA standard setting cut scores across grades 3–8.

Table 2. OSTP ELA Grades 3-8 Impact Data based on Standard Setting, Benchmarks, Prior, and NAEP

Subject Grade	Impact based on	Below Basic	Basic	Proficient	Advanced	Basic & above	Proficient & above
ELA 03	Standard Setting	30.0	19.2	40.0	10.8	70.0	50.8
	Benchmarks	38.4	17.7	42.3	1.6	61.6	43.9
	Prior	40.7	31.6	23.9	3.9	59.3	27.8
ELA 04	Standard Setting	36.1	16.7	38.1	9.1	63.9	47.2
	Benchmarks	37.0	8.2	44.9	9.9	63.0	54.7
	Prior	41.8	33.9	22.5	1.8	58.2	24.3
	OK NAEP (2022)	--	--	--	4.0	55.0	24.0
ELA 05	Standard Setting	22.8	18.0	32.7	26.5	77.2	59.2
	Benchmarks	30.6	30.8	29.8	8.8	69.4	38.6
	Prior	32.1	41.8	21.0	5.2	67.9	26.1
ELA 06	Standard Setting	41.6	15.6	38.6	4.2	58.4	42.8
	Benchmarks	55.5	11.7	31.0	1.7	44.5	32.7
	Prior	34.2	41.0	22.3	2.6	65.8	24.9
ELA 07	Standard Setting	51.3	14.3	32.2	2.2	48.7	34.5
	Benchmarks	48.0	21.7	28.4	1.9	52.0	30.3
	Prior	47.0	32.9	15.2	4.8	53.0	20.0
ELA 08	Standard Setting	40.3	20.1	37.3	2.3	59.7	39.6
	Benchmarks	46.6	23.3	29.1	1.0	53.4	30.1
	Prior	42.0	40.6	14.0	3.4	58.0	17.5
	OK NAEP (2022)	--	--	--	1.0	62.0	21.0

Figure 1. OSTP ELA Grade 3-8 Impact Data based on Standard Setting Cut Scores

ELA Articulation Results

During the articulation portion of the meeting, panelists reviewed test blueprints, standards, and PLDs across grades and discussed their expectation for student performance relative to between grade transition. The discussion was facilitated with guided questions to consider for each grade transition. Table 3 shows the articulation guided questions alongside the panel’s consensus or majority response and panel discussion notes associated with each grade transition. Response options for the transition questions were on a Likert-type scale: (1) Much less challenging, (2) less challenging, (3) about the same, (4) more challenging, or (5) much more challenging.

Based on the panel’s consensus response for each grade transition, Cognia psychometricians adjusted the standard-setting cut scores to achieve articulation as recommended by the articulation panelists. Table 4 shows the articulation adjustments and associated articulated impact data percentages. The OIB page numbers and theta based on the standard setting results are provided in the first two columns. In addition, the change (unit additions or subtractions) in OIB page numbers and theta values based on articulation adjustments are listed for reference.

Figure 2 gives a visual representation of the impact data based on the ELA articulated cut scores across grades 3–8.

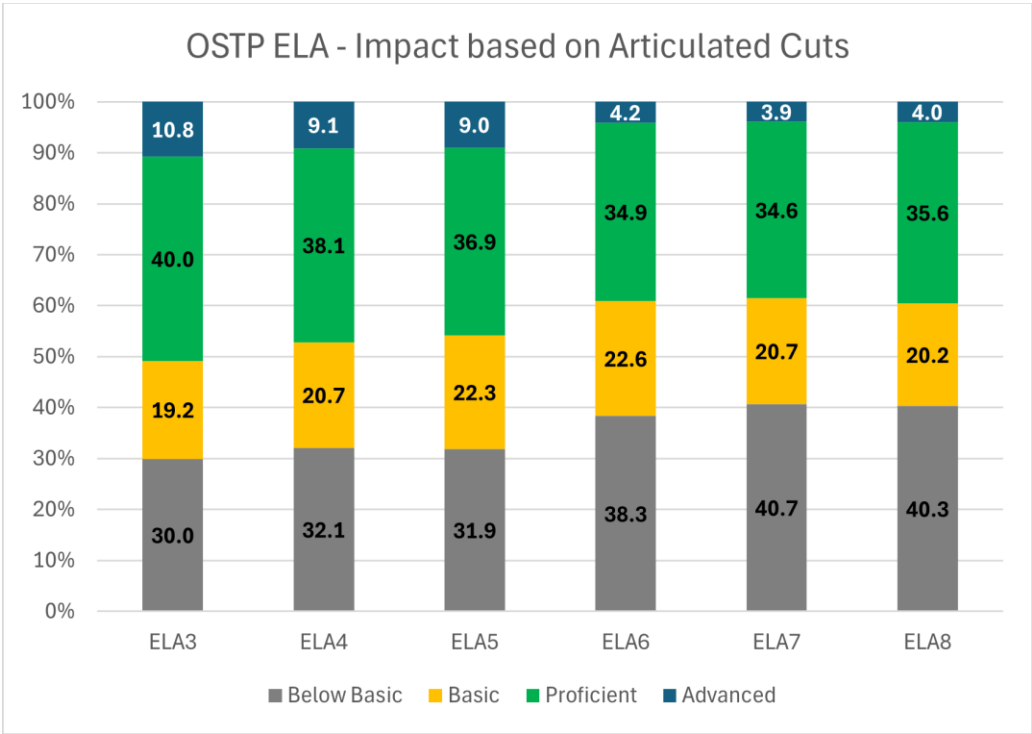
Table 3. OSTP ELA Articulation – Performance Expectations for Grade Transitions

Question	Panel Response	Panel Discussion Notes
Transition 1: How much more/less challenging is it for 4th graders to demonstrate proficiency in a 4th grade test (blueprint), assessing 4th grade standards, as described by 4th grade PLDs THAN IT IS for 3rd graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	About the same	a. Transition from 3-4 i. Third grade is the last year for learning to read. Fourth grade they should make the transition to reading to learn. ii. Historically, fourth grade is an extension of the standards. For example, the very first standard indicates that third grade is harder, and fourth grade is easier. (Main idea/supporting details) iii. For the writing standards, fourth grade IS harder. iv. 3.W.1 represents a cognitive leap from 3rd to 4th grade, BUT that is the only one. All the other standards represent an extension of writing. v. Reading is less, but writing is more. Some of their examples of “reading to learn” are shown by their writing. vi. About the same = 3 votes. vii. More difficult = 2 votes viii. One panelist thinks it is more difficult because reading to learn is hard. But a 3rd grade teacher felt like that shift happens in 3rd grade, NOT from 3rd to 4th. ix. 4th grade is more application of what they’ve learned in 3rd grade. x. Based on the standards, 4th grade is an extension of grade 3, not a huge leap. xi. About the same – 6
Transition 2: How much more/less challenging is it for 5th graders to demonstrate proficiency in a 5th grade test (blueprint), assessing 5th grade standards, as described by 5th grade PLDs THAN IT IS for 4th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	More challenging	a. Transition from 4-5 i. Especially in standard 3, this seemed to be a big leap; there are harder concepts in the standards. For example, 4.R.1 describing the purpose, vs. 5th grade more evaluation of achieving the purpose. ii. Writing is essentially the same, but reading is more challenging. iii. More inference required in grade 5. iv. Votes for more challenging: consensus
Transition 3: How much more/less challenging is it for 6th graders to demonstrate proficiency in a 6th grade test (blueprint), assessing 6th grade standards, as described by 6th grade PLDs THAN IT IS for 5th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	Much more challenging	a. Transition from 5-6 i. 6th grade begins puberty for many students, which makes learning more difficult. 6.W.2 – the jump is huge. They must develop a thesis statement, which is a huge leap beyond the 5th grade standard. Research paper is another big jump. ii. Maybe there are not so many huge leaps in the other standards, but the writing demands are much larger. iii. There are other changes in 6th grade, like changing classes, etc. It is hard for them to show proficiency because the structure of the classes is difficult. iv. Much more challenging: almost unanimous; one vote for more challenging.
Transition 4: How much more/less challenging is it for 7th graders to demonstrate proficiency in a 7th grade test (blueprint), assessing 7th grade standards, as described by 7th grade PLDs THAN IT IS for 6th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	About the same	a. Transition from 6-7 i. About the same – The jump from 5-6 was much more significant than the jump from 6-7. Seventh graders are going through some things (physically, emotionally) but it’s not as much as the shifts for 6th grade. The demands of the standards and the PLDs are about the same. ii. A little more challenging, because they must look at short articles instead of paragraphs. Parts of speech has made a big jump; iii. Consensus – about the same. There were two who were on the fence with less challenging.
Transition 5: How much more/less challenging is it for 8th graders to demonstrate proficiency in an 8th grade test (blueprint), assessing 8th grade standards, as described by 8th grade PLDs THAN IT IS for 7th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	About the same	Transition from 7-8 i. About the same – although another layer is added to the standard/PLDs, it is just a continuation of growth. Although we are adding onto their learning, it is not beyond what you would expect from grade to grade. ii. 3.R.5 – 7th grade theme and mood; 8th grade, just adding tone; this is just the next level and isn’t a huge leap. iii. Seeing very few standards that are different. iv. Less challenging – 3.R.4 – in 8th grade, just supporting interpretations; not a huge leap. v. Students are not going through huge transitions in the 8th grade. vi. One panelist would never say less challenging, because the standards are so challenging for the majority of the students. This allows all their learning/physical/emotional changes to “gel” so that they are ready for high school. vii. Less challenging – because the standards and PLDs are about the same, and the other challenges (physical, emotional, etc. viii. About the same – almost all; one vote for less challenging)

Table 4. OSTP ELA Standard Setting Cut Score Articulation Adjustments

Grade	Performance Level	Standard Setting OIB page	Standard Setting Theta	Change in OIB page	Change in Theta	Articulated Theta Value	Articulated Impact %
ELA 03	Below Basic	--	--	--	--	--	29.96
	Basic	3 - 4	-0.890	--	--	-0.890	19.22
	Proficient	11 - 12	-0.288	--	--	-0.288	40.03
	Advanced	41 - 42	0.949	--	--	0.949	10.79
	Prof + Adv	--	--	--	--	--	50.82
ELA 04	Below Basic	--	--	--	--	--	32.11
	Basic	4 - 5	-0.700	0	- 0.130	-0.830	20.69
	Proficient	17 - 18	-0.225	--	--	-0.225	38.11
	Advanced	35 - 36	0.941	--	--	0.941	9.09
	Prof + Adv	--	--	--	--	--	47.20
ELA 05	Below Basic	--	--	--	--	--	31.88
	Basic	5 - 6	-1.120	0	+ 0.330	-0.790	22.25
	Proficient	11 - 12	-0.531	+14	+ 0.350	-0.181	36.89
	Advanced	42 - 43	0.315	+8	+ 0.620	0.935	8.99
	Prof + Adv	--	--	--	--	--	45.88
ELA 06	Below Basic	--	--	--	--	--	38.34
	Basic	2 - 3	-0.670	0	- 0.100	-0.770	22.56
	Proficient	9 - 10	-0.232	+1	+ 0.100	-0.132	34.94
	Advanced	45 - 46	1.222	--	--	1.222	4.16
	Prof + Adv	--	--	--	--	--	39.10
ELA 07	Below Basic	--	--	--	--	--	40.70
	Basic	8 - 9	-0.380	-4	- 0.300	-0.680	20.73
	Proficient	15 - 16	0.015	-3	- 0.120	-0.105	34.63
	Advanced	47 - 48	1.551	0	- 0.210	1.341	3.93
	Prof + Adv	--	--	--	--	--	38.57
ELA 08	Below Basic	--	--	--	--	--	40.28
	Basic	8 - 9	-0.740	--	--	-0.740	20.15
	Proficient	10 - 11	-0.207	--	--	-0.207	35.60
	Advanced	50 - 51	1.351	-2	- 0.200	1.151	3.96
	Prof + Adv	--	--	--	--	--	39.57

Figure 2. OSTP ELA Impact Data based on Articulated Cut Scores



Mathematics Grades 3-8: Standard Setting Results

Table 5 shows the three cut scores (basic, proficient, and advanced) for each Mathematics grade that resulted from the standard setting meeting and analyses. The table includes the OIB page range, theta, and associated standard error for each cut. In addition, the same information is presented based on the benchmark cut scores. Finally, the prior (pre-standard setting) theta cut scores are also listed for reference.

Table 6 shows the impact data (percentage of students classified in each performance level) for each Mathematics grade based on the cut scores from the Standard Setting meeting and benchmarks. In addition, impact data based on the prior (pre-standard setting) cut scores are listed for reference. Note that percentages related to the standard setting, benchmark, and prior cut scores were calculated by applying the cut scores to student data from the Spring 2024 OSTP Mathematics test administration. Finally, where relevant, percentages based on NAEP data for Oklahoma are also shown. The NAEP data are based on the 2022 test administration and represent the most recent data available.

Figure 2 gives a visual representation of the impact data based on the mathematics standard setting cut scores across grades 3–8.

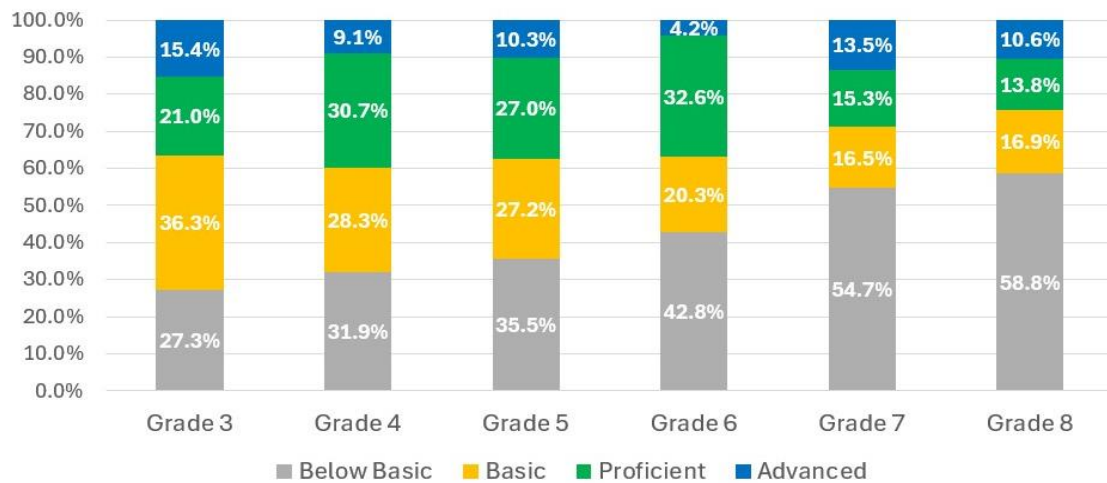
Table 5. OSTP Mathematics Grades 3-8 Cut Score Details based on Standard Setting, Benchmarks, and Prior

Subject Grade	Performance Cut Placement Level	Standard Setting			Benchmarks			Prior Theta
		OIB #	Theta	Standard Error	OIB #	Theta	Standard Error	
Mathematics 03	Basic	11 - 12	-1.000	--	11 - 12	-0.910	--	-0.840
	Proficient	21 - 22	0.106	0.041	19 - 20	0.071	0.140	0.187
	Advanced	42 - 43	0.739	0.058	47 - 48	1.156	0.359	0.988
Mathematics 04	Basic	5 - 6	-0.770	--	5 - 6	-0.730	--	-0.771
	Proficient	12 - 13	0.092	0.023	12 - 13	0.121	0.071	0.270
	Advanced	47 - 48	1.180	0.076	47 - 48	1.301	0.270	1.062
Mathematics 05	Basic	7 - 8	-0.660	--	7 - 8	-0.680	--	-0.829
	Proficient	18 - 19	0.141	0.025	18 - 19	0.153	0.081	0.427
	Advanced	45 - 46	1.109	0.017	46 - 47	1.190	0.157	1.170
Mathematics 06	Basic	9 - 10	-0.480	--	6 - 7	-0.520	--	-0.759
	Proficient	19 - 20	0.078	0.027	21 - 22	0.204	0.068	0.440
	Advanced	48 - 49	1.503	0.120	49 - 50	1.627	0.515	1.511
Mathematics 07	Basic	6 - 7	-0.180	--	6 - 7	-0.190	--	-0.336
	Proficient	14 - 15	0.314	0.026	14 - 15	0.297	0.112	0.447
	Advanced	32 - 33	0.881	0.024	39 - 40	1.160	0.113	1.471
Mathematics 08	Basic	6 - 7	-0.090	--	6 - 7	0.030	--	-0.027
	Proficient	10 - 11	0.416	0.021	11 - 12	0.443	0.073	0.756
	Advanced	32 - 33	0.971	0.028	36 - 37	1.033	0.096	1.267

Table 6. OSTP Mathematics Grades 3-8 Impact Data based on Standard Setting, Benchmarks, Prior, & NAEP

Subject Grade	Impact based on	Below Basic	Basic	Proficient	Advanced	Basic & above	Proficient & above
Mathematics 03	Standard Setting	27.3	36.3	21.0	15.4	72.7	36.4
	Benchmarks	29.6	32.7	30.7	6.9	70.4	37.6
	Prior	31.6	35.0	23.7	9.7	68.4	33.4
Mathematics 04	Standard Setting	31.9	28.3	30.7	9.1	68.1	39.8
	Benchmarks	33.0	28.2	31.7	7.1	67.0	38.8
	Prior	31.9	34.4	22.3	11.4	68.1	33.7
	OK NAEP (2022)	--	--	--	3.0	71.0	27.0
Mathematics 05	Standard Setting	35.5	27.2	27.0	10.3	64.5	37.3
	Benchmarks	34.9	28.2	28.0	8.9	65.1	36.9
	Prior	30.4	41.9	18.5	9.2	69.6	27.8
Mathematics 06	Standard Setting	42.8	20.3	32.6	4.2	57.2	36.9
	Benchmarks	41.4	26.3	29.1	3.2	58.6	32.3
	Prior	33.4	41.9	20.5	4.1	66.6	24.7
Mathematics 07	Standard Setting	54.7	16.5	15.3	13.5	45.3	28.8
	Benchmarks	54.3	16.4	21.1	8.2	45.7	29.3
	Prior	49.1	26.3	20.3	4.3	50.9	24.6
Mathematics 08	Standard Setting	58.8	16.9	13.8	10.6	41.2	24.4
	Benchmarks	62.8	13.7	14.1	9.4	37.2	23.5
	Prior	60.8	24.2	9.2	5.8	39.2	15.0
	OK NAEP (2022)	--	--	--	3.0	52.0	16.0

Figure 2. OSTP Mathematics Grade 3-8 Impact Data based on Standard Setting Cut Scores



Mathematics Articulation Results

Table 7 shows the questions alongside the panel’s consensus or majority response and panel discussion notes associated with each grade transition. Response options for the transition question were on a Likert-type scale: (1) Much less challenging, (2) less challenging, (3) about the same, (4) more challenging, or (5) much more challenging.

Based on the panel’s consensus response for each grade transition, Cognia psychometricians adjusted the mathematics standard setting cut scores to achieve articulated impact data as recommended by the articulation panelists. Table 8 shows the articulation adjustments and associated articulated impact data percentages. The OIB page numbers and theta based on the standard setting results are provided in the first two columns. In addition, the change (unit additions or subtractions) in OIB page numbers and theta values based on articulation adjustments are listed for reference.

Figure 3 gives a visual representation of the impact data based on the mathematics articulated cut scores across grades 3–8.

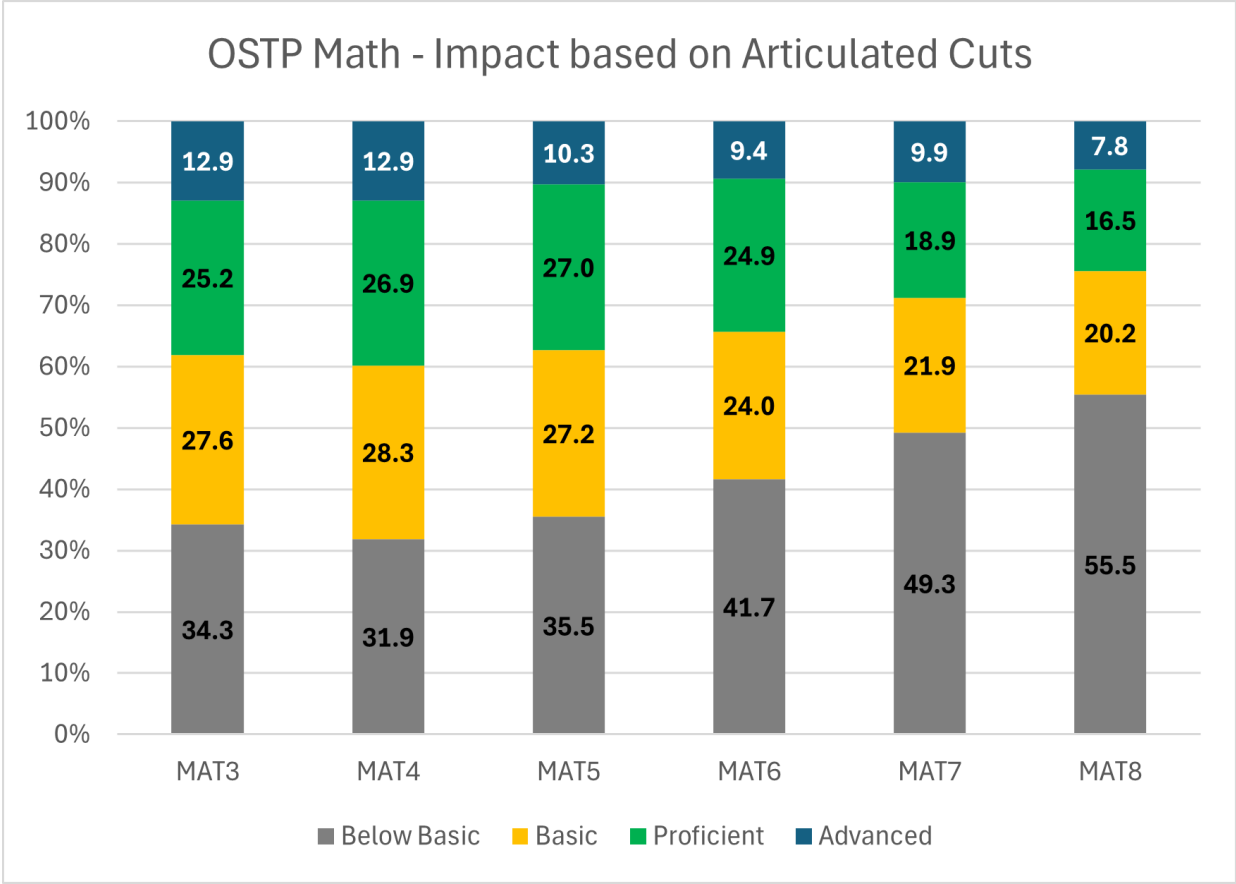
Table 7. OSTP Mathematics Articulation – Performance Expectations for Grade Transitions

Question	Panel Response	Panel Discussion Notes
Transition 1: How much more/less challenging is it for 4th graders to demonstrate proficiency in a 4th grade test (blueprint), assessing 4th grade standards, as described by 4th grade PLDs THAN IT IS for 3rd graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	Less challenging	Same concepts, but just extended. Lots of practice, not as many new concepts as other grades.
Transition 2: How much more/less challenging is it for 5th graders to demonstrate proficiency in a 5th grade test (blueprint), assessing 5th grade standards, as described by 5th grade PLDs THAN IT IS for 4th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	More challenging	Many new and challenging concepts in 5th grade. First real application tasks, students have multi-operational task with meaning – getting to the WHY. Not a monumental shift, but an increase in challenge.
Transition 3: How much more/less challenging is it for 6th graders to demonstrate proficiency in a 6th grade test (blueprint), assessing 6th grade standards, as described by 6th grade PLDs THAN IT IS for 5th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	More challenging	From grade 5 to grade 6, the concepts are moving from concrete to abstract. Now students must illustrate tougher concepts, and some new concepts. The material is more challenging. Basic and abstract are not different.
Transition 4: How much more/less challenging is it for 7th graders to demonstrate proficiency in a 7th grade test (blueprint), assessing 7th grade standards, as described by 7th grade PLDs THAN IT IS for 6th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	More challenging	Notes: several panelists (3-4) felt that the transition was MUCH MORE challenging. 7th grade skills go heavy into percents, other big blueprint changes include less Number and Operations, more Algebraic Reasoning and Algebra, and WAY more Geometry and Measurement (22-26% in 6th grade to 30-36% in 7th grade). If students don't have strong Number and Operations skills, it affects all other areas. 7th grade starts to use operations with rational numbers. 7th grade flows around proportional reasoning.
Transition 5: How much more/less challenging is it for 8th graders to demonstrate proficiency in an 8th grade test (blueprint), assessing 8th grade standards, as described by 8th grade PLDs THAN IT IS for 7th graders to demonstrate proficiency on the blueprint, standards and PLDs of their grade	Much more challenging	"Geometry for 8th grade is a very small percentage of the blueprint. The majority is algebraic reasoning and algebra. New concepts galore, solving multi-step problems, variables on both sides. There is scientific notation, other abstract concepts too. Foundation started early and progressed. (for Algebraic Reasoning & Algebra). More dramatic flip from concrete topics in 7th grade to abstract concepts in 8th grade. Students feel the stress of the new content. 7th graders seem to feel more comfortable, still in elementary school."

Table 8. OSTP Mathematics Standard Setting Cut Score Articulation Adjustments

Grade	Performance Level	Standard Setting OIB page	Standard Setting Theta	Change in OIB page	Change in Theta	Articulated Theta Value	Articulated Impact %
Mathematics 03	Below Basic	--	--	--	--	--	34.26
	Basic	11 - 12	-1.000	+1	+0.250	-0.750	27.61
	Proficient	21 - 22	0.106	-2	-0.050	0.056	25.23
	Advanced	42 - 43	0.739	+3	+0.100	0.839	12.89
	Prof + Adv	--	--	--	--	--	38.13
Mathematics 04	Below Basic	--	--	--	--	--	31.88
	Basic	5 - 6	-0.770	--	--	-0.770	28.34
	Proficient	12 - 13	0.092	--	--	0.092	26.92
	Advanced	47 - 48	1.180	-1	-0.190	0.989	12.86
	Prof + Adv	--	--	--	--	--	39.78
Mathematics 05	Below Basic	--	--	--	--	--	35.50
	Basic	7 - 8	-0.660	--	--	-0.660	27.20
	Proficient	18 - 19	0.141	--	--	0.141	27.03
	Advanced	45 - 46	1.109	--	--	1.109	10.27
	Prof + Adv	--	--	--	--	--	37.30
Mathematics 06	Below Basic	--	--	--	--	--	41.70
	Basic	9 - 10	-0.480	-1	-0.030	-0.510	24.00
	Proficient	19 - 20	0.078	0	+0.070	0.148	24.93
	Advanced	48 - 49	1.503	-2	-0.410	1.093	9.37
	Prof + Adv	--	--	--	--	--	34.30
Mathematics 07	Below Basic	--	--	--	--	--	49.28
	Basic	6 - 7	-0.180	0	-0.150	-0.330	21.90
	Proficient	14 - 15	0.314	--	--	0.314	18.88
	Advanced	32 - 33	0.881	+3	+0.180	1.061	9.94
	Prof + Adv	--	--	--	--	--	28.82
Mathematics 08	Below Basic	--	--	--	--	--	55.45
	Basic	6 - 7	-0.090	0	-0.100	-0.190	20.16
	Proficient	10 - 11	0.416	--	--	0.416	16.54
	Advanced	32 - 33	0.971	+3	+0.150	1.121	7.84
	Prof + Adv	--	--	--	--	--	24.39

Figure 3. OSTP Mathematics Impact Data based on Articulated Cut Scores



APPENDIX—O
FINAL CUT POINTS

Table 1. OSTP ELA Final Cut Scores and Impact Percentages by Grade

Grade	Performance Level	OIB Page Range	Theta Value	Impact %
3	Below Basic	--		29.96
	Basic	3 - 4	-0.890	19.22
	Proficient	11 - 12	-0.288	40.03
	Advanced	41 - 42	0.949	10.79
	Prof + Adv	--	--	50.82
4	Below Basic	--	--	32.11
	Basic	4 - 5	-0.830	20.69
	Proficient	17 - 18	-0.225	38.11
	Advanced	35 - 36	0.941	9.09
	Prof + Adv	--	--	47.20
5	Below Basic	--	--	31.88
	Basic	5 - 6	-0.790	22.25
	Proficient	25 - 26	-0.181	36.89
	Advanced	50 - 51	0.935	8.99
	Prof + Adv	--	--	45.88
6	Below Basic	--	--	38.34
	Basic	2 - 3	-0.770	22.56
	Proficient	10 - 11	-0.132	34.94
	Advanced	45 - 46	1.222	4.16
	Prof + Adv	--	--	39.10
7	Below Basic	--	--	40.70
	Basic	4 - 5	-0.680	20.73
	Proficient	12 - 13	-0.105	34.63
	Advanced	47 - 48	1.341	3.93
	Prof + Adv	--	--	38.57
8	Below Basic	--	--	40.28
	Basic	8 - 9	-0.740	20.15
	Proficient	10 - 11	-0.207	35.60
	Advanced	48 - 49	1.151	3.96
	Prof + Adv	--	--	39.57

Table 2. OSTP Mathematics Final Cut Scores and Impact Percentages by Grade

Grade	Performance Level	OIB Page Range	Theta Value	Impact %
3	Below Basic	--		34.26
	Basic	12 - 13	-0.750	27.61
	Proficient	19 - 20	0.056	25.23
	Advanced	45 - 46	0.839	12.89
	Prof + Adv	--	--	38.13
4	Below Basic	--	--	31.88
	Basic	5 - 6	-0.770	28.34
	Proficient	12 - 13	0.092	26.92
	Advanced	46 - 47	0.989	12.86
	Prof + Adv	--	--	39.78
5	Below Basic	--	--	35.50
	Basic	7 - 8	-0.660	27.20
	Proficient	18 - 19	0.141	27.03
	Advanced	45 - 46	1.109	10.27
	Prof + Adv	--	--	37.30
6	Below Basic	--	--	41.70
	Basic	8 - 9	-0.510	24.00
	Proficient	19 - 20	0.148	24.93
	Advanced	46 - 47	1.093	9.37
	Prof + Adv	--	--	34.30
7	Below Basic	--	--	49.28
	Basic	6 - 7	-0.330	21.90
	Proficient	14 - 15	0.314	18.88
	Advanced	35 - 36	1.061	9.94
	Prof + Adv	--	--	28.82
8	Below Basic	--	--	55.45
	Basic	6 - 7	-0.190	20.16
	Proficient	10 - 11	0.416	16.54
	Advanced	35 - 36	1.121	7.84
	Prof + Adv	--	--	24.39

APPENDIX—P
COMMISSION FOR EDUCATIONAL QUALITY AND
ACCOUNTABILITY (CEQA) PRESENTATION

Performance Levels and Cut Scores for the OSTP ELA and Math Assessments

Presentation to the Commission for
Educational Quality and Accountability

July 10, 2024



OKLAHOMA
Education

cognia™

1

Members of the Team

- **Catherine Boomer**, Program Director, State Assessments, OSDE
- **Alyssa Tyra**, Project Manager, ELA Assessments, OSDE
- **Corinne Beasler**, Project Manager, Math Assessments, OSDE
- **Dr. Frank Padellaro**, Vice President Psychometrics and Reporting Services, Cognia
- **Julie DiBona**, Vice President, Program Management, Cognia



2

Background on Grades 3-8 ELA & Math Assessments



3

Oklahoma Statute on Performance Levels

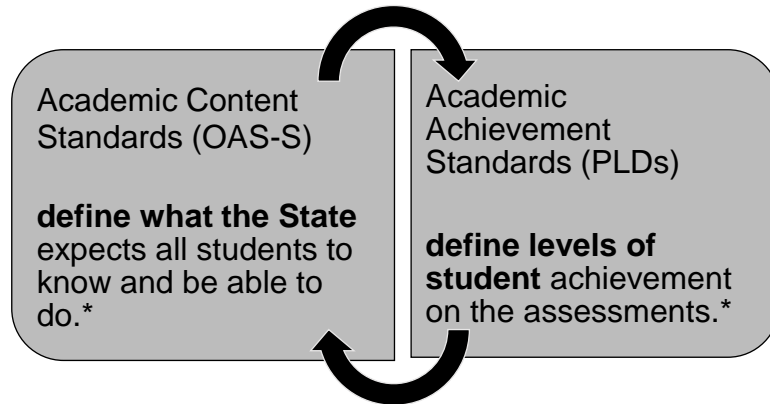
- OSTP Performance is divided into performance levels.
- The Performance levels shall be set by a method that indicates students are ready for the next grade, course, or level of education, as applicable.
- The Commission for Educational Quality and Accountability (CEQA) shall determine and adopt a series of student performance levels and the corresponding cut scores pursuant to the Oklahoma School Testing Program Act.
- §70-1210.541



4

4

Content Standards and PLDs



**U.S. Department of Education Peer Review of State Assessment Systems Non-Regulatory Guidance for States, September 25, 2015*

5 |



5

Logistics of the Standard Setting Meeting

- Standard Setting: June 17-20, 2024
- Location: Stoney Creek Hotel, Tulsa-Broken Arrow, OK

Grade Span	Content	Number of Panelists
Grades 3-4	Math	11
Grades 5-6	Math	12
Grades 7-8	Math	12
Grades 3-4	ELA	11
Grades 5-6	ELA	10
Grades 7-8	ELA	10



6

6

Logistics of the Standard-Setting Meeting

- Articulation Meeting:
 - Math: Afternoon of June 20, 2024
 - ELA: Morning of June 21, 2024
- Location: Stoney Creek Hotel, Tulsa-Broken Arrow, OK

Grade Span	Content	Number of Panelists
Grades 3-8	Math	12
Grades 3-8	ELA	11



7

7

Logistics Continued

- How long have you been teaching?

Years Teaching	Number of Panelists
1-5 Years	29
6-10 Years	11
11-20 Years	16
21+ Years	10

- Location Demographics

- *Based on National Center for Education Statistics
- https://nces.ed.gov/ccd/districtsearch/district_list.asp?Search=1&State=40

Location*	Number of Panelists
City: Large	14
City: Small	1
Rural: Distant	10
Rural: Fringe	8
Rural: Remote	3
Suburb: Large	11
Town: Distant	12
Town: Fringe	1
Town: Remote	5



8

8

Standards and Assessments



What are standards?

The content students are expected to know by the end of a grade level and subject.

Guideposts for teachers to build their lesson plans and develop “can-do” statements.

They answer: What can students do as a result of learning these standards?



What are large-scale assessments?

They are designed to cover the depth (complexity) and breadth (scope) of the standards across a year.

They provide large grain-size information on how student performance compares to end-of-grade level expectations.



9

Assessments and Performance Expectations



There is a lot of content to cover in an assessment based on the breadth and depth of the state’s standards.



How much content is enough to say students are on track to meet the challenges of the *next grade, course, or level of education, as applicable?*



Setting achievement **standards** (i.e., standard setting) requires expert judgment from teachers of the content to determine what content represents Below Basic, Basic, Proficient, or Advanced knowledge.



10

What are Performance Level Descriptors?

- PLDs provide a narrative account of the knowledge, skills, and abilities demonstrated by students in each level of achievement:
 - Below Basic, Basic, Proficient and Advanced
- Describe what students know and can do based on the Oklahoma Academic Standards.
- Inform stakeholders of how to interpret student test scores in relation to the Oklahoma Academic Standards.
- Are typically used for standard setting and score reporting.



11

Background on PLD development

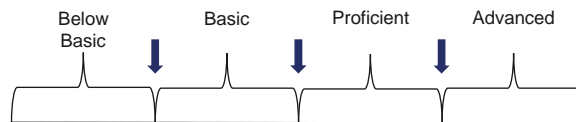
- New standards were adopted by OSDE. As a result, the PLDs needed to be updated so that they accurately reflect what students know and can do at each performance level.
- After adopting new standards, OSDE and Cognia staff collaborated on the development of new PLDs using the updated standards as a foundation.
- Teacher committees reviewed and discussed draft PLDs. After this discussion, OSDE finalized the PLDs.



12

Standard Setting for OSTP ELA and Math Grades 3 – 8 Content Assessments

- Standard setting is a deliberative process used to establish the test scores that separate achievement levels (e.g., basic/proficient) on a test.
- A total of 66 Oklahoma educators from various districts were selected to participate in this process.
- These Oklahoma experts matched test performance to descriptions of the knowledge, skills, and abilities defining each of the four performance levels on the OSTP assessments.
- Note: Oklahoma educators were organized into grade-band panels where each panel completed the standard setting activities for two grades



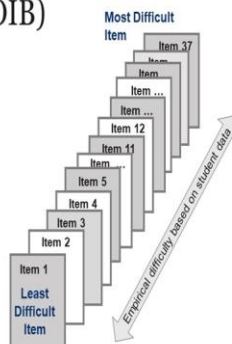
13

Standard Setting for OSTP ELA and Math Grades 3 - 8 Content Assessments

The Expert Judgment Task

Ordered item booklet (OIB)

- The OIB contains test items ordered by difficulty.
- Each OIB page represents an item.
- The difference in difficulty is not exactly the same between each pair of neighboring items.
- Difficulty is based on data from the students who took the test during prior administrations.



ID Matching process

For each item in the OIB:

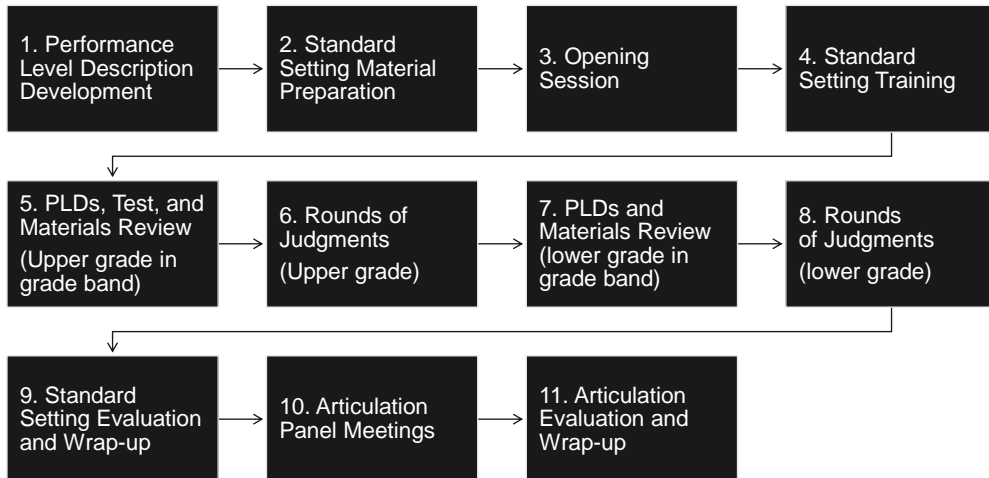
- 1. Review the item and identify the KSAs**
 - Identify the knowledge, skills, and abilities (KSAs) required to respond to the item correctly.
- 2. Make an item-PLD alignment judgment**
 - Match the KSAs required by the item with the expectations described in either the Basic, Proficient, or Advanced performance level descriptor (PLD).

What does a student need to know or be able to do to correctly respond to this item?

Which PLD most closely matches the knowledge, skills, and abilities (KSAs) required by the item?

14

Standard Setting for OSTP ELA and Math Grades 3 - 8 Content Assessments



15

Independent Observer Feedback

"Cognia implemented the ID matching approach with fidelity. Panelist exit surveys clearly indicates that panelists felt that they: understood the task, tools and feedback at each step in the process; had sufficient time for training and practice as well as opportunities to pose questions; and felt like the facilitator provided clear responses to questions and requests for clarification. Our observations confirm these results – the training, facilitation, tools, and participation were all the highest quality" – Dr. Erika Landl

16

16

Articulation process: The “why”

- Why do we **COMPARE** the challenge of demonstrating proficiency for students in different grades?
 - Each of our panelists and facilitators are different (thank goodness)
 - On a different day, with different people and different facilitators (reviewing different items) there would likely be different judgments. That’s okay and expected!
 - We know each grade has greater expectations in general (that’s learning!), but...
 - We had Oklahoma educators examine the challenge for a 5th grader (for example) who has had a full year of 5th grade instruction and development compared to that for a 6th grader.



17

Why is it reasonable to articulate (adjust) cuts?

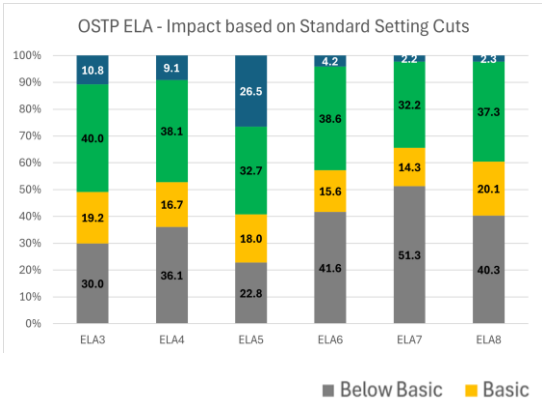
- Because there is no perfect cut judgment from a single standard setting activity, it is reasonable to make adjustments
 - Large jumps in impact data (performance level percentages) that can’t be explained by differences in the grade level challenges for students may be the result of random differences in panel results
 - This difference creates a lack of program coherence that is hard to explain to stakeholders
 - Minor changes to the cuts were reviewed by SDE and TAC members who noted the changes (for the most part) were trivial compared to panelist variance
 - The recommended articulation cuts reflected the feedback of OSDE, TAC and OK educators
 - This process is a normal part of most standard settings involving multiple grades in the same content area



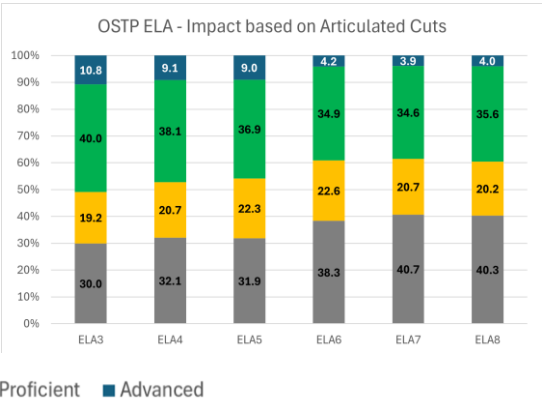
18

Comparison of an unarticulated to smoothed content area (ELA)

Unarticulated



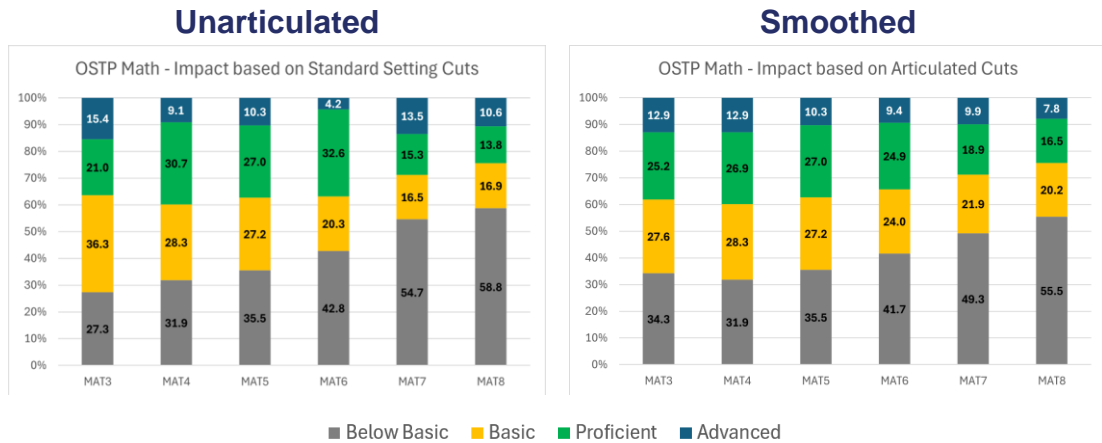
Smoothed



ELA Recommended Cut Scores

Grade	Performance Level	Standard Setting Theta	Articulation	Recommended Theta Cuts	Impact %	Grade	Performance Level	Standard Setting Theta	Articulation	Recommended Theta Cuts	Impact %
ELA 03	Below Basic	--	--	--	29.96	ELA 06	Below Basic	--	--	--	38.34
	Basic	-0.89	--	-0.89	19.22		Basic	-0.67	-0.1	-0.77	22.56
	Proficient	-0.288	--	-0.288	40.03		Proficient	-0.232	0.1	-0.132	34.94
	Advanced	0.949	--	0.949	10.79		Advanced	1.222	--	1.222	4.16
	Prof + Adv	--	--	--	50.82		Prof + Adv	--	--	--	39.1
ELA 04	Below Basic	--	--	--	32.11	ELA 07	Below Basic	--	--	--	40.7
	Basic	-0.7	-0.13	-0.83	20.69		Basic	-0.38	-0.3	-0.68	20.73
	Proficient	-0.225	--	-0.225	38.11		Proficient	0.015	-0.12	-0.105	34.63
	Advanced	0.941	--	0.941	9.09		Advanced	1.551	-0.21	1.341	3.93
	Prof + Adv	--	--	--	47.2		Prof + Adv	--	--	--	38.57
ELA 05	Below Basic	--	--	--	31.88	ELA 08	Below Basic	--	--	--	40.28
	Basic	-1.12	0.33	-0.79	22.25		Basic	-0.74	--	-0.74	20.15
	Proficient	-0.531	0.35	-0.181	36.89		Proficient	-0.207	--	-0.207	35.6
	Advanced	0.315	0.62	0.935	8.99		Advanced	1.351	-0.2	1.151	3.96
	Prof + Adv	--	--	--	45.88		Prof + Adv	--	--	--	39.57

Comparison of an unarticulated to smoothed content area (Math)



21

Math Recommended Cut Scores

Grade	Performance Level	Standard Setting Theta	Articulation	Recommended Theta Cuts	Impact %
Math 03	Below Basic	--	--	--	34.26
	Basic	-1	0.25	-0.75	27.61
	Proficient	0.106	-0.05	0.056	25.23
	Advanced	0.739	0.1	0.839	12.89
	Prof + Adv	--	--	--	38.13
Math 04	Below Basic	--	--	--	31.88
	Basic	-0.77	--	-0.77	28.34
	Proficient	0.092	--	0.092	26.92
	Advanced	1.18	-0.19	0.989	12.86
	Prof + Adv	--	--	--	39.78
Math 05	Below Basic	--	--	--	35.5
	Basic	-0.66	--	-0.66	27.2
	Proficient	0.141	--	0.141	27.03
	Advanced	1.109	--	1.109	10.27
	Prof + Adv	--	--	--	37.3
Math 06	Below Basic	--	--	--	41.7
	Basic	-0.48	-0.03	-0.51	24
	Proficient	0.078	0.07	0.148	24.93
	Advanced	1.503	-0.41	1.093	9.37
	Prof + Adv	--	--	--	34.3
Math 07	Below Basic	--	--	--	49.28
	Basic	-0.18	-0.15	-0.33	21.9
	Proficient	0.314	--	0.314	18.88
	Advanced	0.881	0.18	1.061	9.94
	Prof + Adv	--	--	--	28.82
Math 08	Below Basic	--	--	--	55.45
	Basic	-0.09	-0.1	-0.19	20.16
	Proficient	0.416	--	0.416	16.54
	Advanced	0.971	0.15	1.121	7.84
	Prof + Adv	--	--	--	24.39

22

APPENDIX R

PERFORMANCE LEVEL DISTRIBUTIONS

Table R-1. Performance Level Distributions by Grade and Year*—ELA

Grade	Performance Level	% in Level 2025	% in Level 2024	% in Level 2023	% in Level 2022	% in Level 2021	% in Level 2019	% in Level 2018
3	1	42	30	41	40	44	31	34
	2	30	19	31	32	32	30	33
	3	24	40	25	23	21	29	27
	4	3	11	4	6	4	10	6
4	1	43	32	41	43	45	36	30
	2	34	21	35	33	33	33	34
	3	22	38	22	21	20	24	28
	4	2	9	2	2	2	6	7
5	1	22	32	23	26	31	25	23
	2	46	22	46	43	41	40	42
	3	25	37	25	23	21	27	22
	4	6	9	6	8	6	8	13
6	1	34	38	32	31	31	22	22
	2	41	23	42	43	44	42	40
	3	23	35	23	22	21	28	29
	4	3	4	3	4	4	8	9
7	1	45	41	44	44	46	35	32
	2	32	21	34	34	34	36	41
	3	17	35	17	16	15	21	20
	4	5	4	4	5	4	8	8
8	1	34	34	33	30	33	25	24
	2	42	21	46	42	43	43	43
	3	18	40	17	22	18	24	24
	4	6	4	4	6	6	9	9

*Tests were not administered in 2019-20 due to COVID-19.

Table R-2. Performance Level Distributions by Grade and Year*—Mathematics

Grade	Performance Level	% in Level 2025	% in Level 2024	% in Level 2023	% in Level 2022	% in Level 2021	% in Level 2019	% in Level 2018
3	1	24	26	25	33	35	24	24
	2	37	29	37	33	35	33	35
	3	27	29	27	22	20	26	26
	4	12	15	12	11	9	17	15
4	1	23	23	23	35	37	26	27
	2	37	30	36	32	35	36	37
	3	26	31	26	20	18	26	25
	4	14	16	14	13	10	12	11
5	1	21	27	23	32	37	24	25
	2	47	29	45	41	41	45	46
	3	21	32	21	18	15	19	20
	4	11	12	11	8	8	11	10
6	1	25	34	28	38	37	27	29
	2	46	26	45	40	42	43	43
	3	23	29	22	18	16	25	23
	4	5	11	5	5	5	6	5
7	1	41	43	43	48	55	38	34
	2	29	24	30	28	25	29	32
	3	24	22	23	20	17	26	26
	4	6	12	5	4	3	7	8
8	1	54	50	57	61	65	50	52
	2	27	22	27	23	21	30	28
	3	11	19	10	10	9	11	10
	4	8	9	6	6	5	10	10

*Tests were not administered in 2019-20 due to COVID-19.

Table R-3. Performance Level Distributions by Grade and Year*—Science

Grade	Performance Level	% in Level 2025	% in Level 2024	% in Level 2023	% in Level 2022	% in Level 2021	% in Level 2019	% in Level 2018
5	1	21	20	20	28	28	22	20
	2	41	41	39	34	40	40	39
	3	32	33	34	31	27	30	32
	4	5	6	7	7	5	8	9
8	1	30	29	27	48	45	39	40
	2	31	32	32	21	22	21	21
	3	32	32	33	24	26	31	29
	4	8	7	7	6	6	9	10
11	1	58	53	57	54	52	57	--
	2	21	23	21	21	24	20	--
	3	16	17	16	18	17	17	--
	4	6	7	6	8	6	7	--

*Tests were not administered in 2019-20 due to COVID-19.

Table R-4. Performance Level Distributions by Grade and Year*—U.S. History

Grade	Performance Level	% in Level 2025	% in Level 2024	% in Level 2023	% in Level 2022	% in Level 2021	% in Level 2019	% in Level 2018
11	1	43	40	42	40	--	--	--
	2	14	15	14	14	--	--	--
	3	34	36	35	36	--	--	--
	4	9	9	9	10	--	--	--

*Tests were not administered in 2019-20 due to COVID-19.

APPENDIX S
CLASSICAL RELIABILITY

Table S-1. Subgroup Reliabilities Grade 3—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	24,367	52	26.86	10.38	0.91	3.13
Male	25,266	51	25.87	10.46	0.91	3.13
American Indian/Alaskan Native	10,626	52	23.61	10.12	0.90	3.17
Hispanic or Latino	5,067	50	25.85	9.73	0.89	3.17
Asian	1,206	50	29.39	10.67	0.92	3.08
Black/African American	3,806	50	21.95	9.87	0.90	3.17
Pacific Islander	257	46	21.06	8.73	0.87	3.20
White/Caucasian	21,320	51	28.40	10.30	0.91	3.10
Two or More Races	7,351	51	26.72	10.29	0.91	3.14
Economically Disadvantaged	32,401	51	24.25	10.01	0.90	3.17
Individual Education Program	9,928	50	19.42	9.27	0.88	3.15
Plan 504	1,510	50	27.17	9.61	0.89	3.14
English Language Learners	6,893	50	21.28	9.27	0.88	3.19

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-2. Subgroup Reliabilities Grade 4—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	24,034	52	29.45	10.09	0.90	3.12
Male	24,942	52	28.26	10.33	0.91	3.12
American Indian/Alaskan Native	10,454	51	25.86	10.07	0.90	3.17
Hispanic or Latino	5,015	51	28.67	9.75	0.90	3.13
Asian	1,188	51	31.53	10.83	0.92	3.04
Black/African American	3,745	51	24.23	9.55	0.89	3.19
Pacific Islander	237	48	24.22	8.71	0.87	3.17
White/Caucasian	20,977	52	30.92	9.95	0.90	3.08
Two or More Races	7,360	51	29.33	9.97	0.90	3.12
Economically Disadvantaged	31,555	52	26.68	9.88	0.90	3.16
Individual Education Program	9,835	51	21.22	9.28	0.88	3.18
Plan 504	1,777	51	29.64	9.38	0.89	3.12
English Language Learners	6,383	49	22.51	8.84	0.87	3.21

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-3. Subgroup Reliabilities Grade 5—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,416	55	38.38	9.61	0.90	2.99
Male	20,540	55	37.51	9.89	0.91	3.00
American Indian/Alaskan Native	8,017	55	35.35	10.34	0.91	3.11
Hispanic or Latino	4,229	55	37.05	9.20	0.89	3.05
Asian	1,020	54	40.16	9.84	0.91	2.89
Black/African American	3,105	54	32.63	10.75	0.91	3.17
Pacific Islander	169	51	29.88	9.91	0.89	3.27
White/Caucasian	18,407	55	40.04	8.84	0.89	2.90
Two or More Races	6,010	54	38.24	9.40	0.90	2.99
Economically Disadvantaged	24,574	55	35.82	9.94	0.90	3.09
Individual Education Program	3,473	54	30.62	11.59	0.93	3.17
Plan 504	1,869	54	37.95	9.03	0.89	3.00
English Language Learners	3,053	52	28.42	9.50	0.88	3.31

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-4. Subgroup Reliabilities Grade 6—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	23,932	52	30.39	10.59	0.91	3.10
Male	25,363	52	28.61	10.67	0.92	3.10
American Indian/Alaskan Native	10,494	52	26.46	10.59	0.91	3.17
Hispanic or Latino	5,279	52	29.10	10.07	0.90	3.12
Asian	1,259	52	33.13	11.34	0.93	3.02
Black/African American	3,724	52	25.08	10.36	0.91	3.18
Pacific Islander	228	50	23.75	10.16	0.90	3.21
White/Caucasian	21,404	52	31.48	10.34	0.91	3.06
Two or More Races	6,907	51	30.01	10.38	0.91	3.10
Economically Disadvantaged	30,541	52	27.17	10.35	0.91	3.16
Individual Education Program	9,175	52	20.25	9.04	0.88	3.19
Plan 504	2,112	52	30.75	9.41	0.89	3.11
English Language Learners	4,848	48	19.55	7.92	0.84	3.21

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-5. Subgroup Reliabilities Grade 7—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	23,940	52	28.89	10.59	0.91	3.18
Male	25,028	52	27.18	11.03	0.92	3.18
American Indian/Alaskan Native	10,183	51	24.71	10.54	0.91	3.24
Hispanic or Latino	5,324	52	27.52	10.32	0.90	3.22
Asian	1,290	52	32.03	11.30	0.93	3.07
Black/African American	3,822	52	23.22	10.13	0.90	3.25
Pacific Islander	277	50	21.66	10.46	0.91	3.21
White/Caucasian	21,450	52	30.27	10.53	0.91	3.15
Two or More Races	6,622	51	28.44	10.72	0.91	3.19
Economically Disadvantaged	30,051	52	25.58	10.46	0.90	3.24
Individual Education Program	8,536	51	18.84	8.77	0.87	3.21
Plan 504	2,199	52	29.39	10.12	0.90	3.18
English Language Learners	5,118	49	18.55	7.73	0.82	3.24

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-6. Subgroup Reliabilities Grade 8—ELA*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	21,416	57	37.39	10.30	0.91	3.14
Male	21,524	57	34.67	10.96	0.91	3.22
American Indian/Alaskan Native	8,668	57	33.20	11.06	0.91	3.28
Hispanic or Latino	4,749	56	35.49	10.14	0.90	3.21
Asian	1,037	57	39.39	10.65	0.92	3.07
Black/African American	3,208	56	31.21	10.82	0.91	3.31
Pacific Islander	181	52	30.23	10.46	0.90	3.38
White/Caucasian	19,226	57	37.96	10.20	0.91	3.12
Two or More Races	5,872	57	36.50	10.40	0.91	3.17
Economically Disadvantaged	25,059	57	33.51	10.66	0.91	3.27
Individual Education Program	3,300	56	27.44	10.63	0.90	3.34
Plan 504	2,033	56	36.26	10.42	0.91	3.18
English Language Learners	3,701	53	26.22	9.18	0.86	3.43

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-7. Subgroup Reliabilities Grade 3—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,485	50	32.51	9.89	0.92	2.83
Male	20,211	50	34.91	9.86	0.92	2.71
American Indian/Alaskan Native	7,148	50	31.71	9.95	0.92	2.87
Hispanic or Latino	4,316	50	33.04	9.46	0.91	2.83
Asian	876	50	38.34	8.96	0.92	2.51
Black/African American	3,271	50	26.82	10.47	0.92	2.99
Pacific Islander	160	50	27.96	10.18	0.92	2.96
White/Caucasian	18,556	50	35.80	9.21	0.91	2.69
Two or More Races	6,369	50	33.36	9.88	0.92	2.79
Economically Disadvantaged	25,192	50	31.42	9.99	0.92	2.88
Individual Education Program	4,139	50	30.55	10.78	0.93	2.90
Plan 504	1,273	50	33.72	9.43	0.91	2.81
English Language Learners	3,390	50	30.32	9.93	0.91	2.92

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-8. Subgroup Reliabilities Grade 4—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	19,879	50	29.86	10.08	0.91	2.96
Male	19,636	50	32.84	10.23	0.92	2.86
American Indian/Alaskan Native	6,990	50	29.45	10.18	0.91	2.99
Hispanic or Latino	4,144	50	30.76	9.72	0.91	2.95
Asian	900	50	36.71	9.66	0.92	2.66
Black/African American	3,157	50	24.16	10.00	0.91	3.07
Pacific Islander	160	49	25.44	10.07	0.91	3.05
White/Caucasian	17,867	50	33.40	9.73	0.91	2.85
Two or More Races	6,297	50	30.97	10.13	0.92	2.93
Economically Disadvantaged	23,979	50	28.96	10.05	0.91	3.00
Individual Education Program	3,297	50	28.30	10.92	0.92	2.99
Plan 504	1,500	50	30.83	10.12	0.91	2.96
English Language Learners	3,067	50	26.97	9.97	0.91	3.05

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-9. Subgroup Reliabilities Grade 5—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	19,875	50	29.10	9.60	0.91	2.94
Male	19,790	50	31.54	9.86	0.91	2.89
American Indian/Alaskan Native	7,609	50	28.10	9.64	0.90	2.99
Hispanic or Latino	4,087	50	29.20	9.20	0.90	2.97
Asian	998	50	35.61	9.75	0.93	2.66
Black/African American	3,004	50	24.25	9.62	0.90	3.06
Pacific Islander	152	48	23.88	9.38	0.89	3.07
White/Caucasian	17,939	50	32.38	9.35	0.91	2.86
Two or More Races	5,877	50	30.03	9.62	0.91	2.94
Economically Disadvantaged	23,573	50	27.88	9.51	0.90	3.00
Individual Education Program	2,542	50	26.30	10.25	0.91	3.02
Plan 504	1,835	50	29.86	9.45	0.90	2.97
English Language Learners	2,722	49	23.54	8.64	0.87	3.10

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-10. Subgroup Reliabilities Grade 6—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,445	50	26.18	9.02	0.89	2.99
Male	20,619	50	28.38	9.46	0.90	2.93
American Indian/Alaskan Native	8,145	50	24.74	8.78	0.88	3.02
Hispanic or Latino	4,333	50	26.65	8.61	0.88	2.98
Asian	1,140	50	32.32	10.16	0.93	2.78
Black/African American	3,057	50	21.73	8.54	0.87	3.05
Pacific Islander	188	48	21.41	9.01	0.89	3.00
White/Caucasian	18,276	50	29.26	9.05	0.90	2.91
Two or More Races	5,925	50	27.23	9.20	0.90	2.97
Economically Disadvantaged	24,379	50	25.04	8.79	0.88	3.02
Individual Education Program	2,480	50	22.62	9.25	0.89	3.05
Plan 504	1,829	50	27.38	8.99	0.89	2.98
English Language Learners	2,937	48	19.46	7.02	0.81	3.08

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-11. Subgroup Reliabilities Grade 7—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,740	50	21.16	9.63	0.90	3.07
Male	20,326	50	23.25	10.49	0.92	3.06
American Indian/Alaskan Native	7,830	50	19.78	9.22	0.89	3.09
Hispanic or Latino	4,455	50	21.16	9.27	0.89	3.09
Asian	1,124	50	28.58	11.52	0.94	2.93
Black/African American	3,195	49	16.75	8.15	0.86	3.05
Pacific Islander	217	47	17.16	8.55	0.87	3.05
White/Caucasian	18,539	50	24.21	10.26	0.91	3.06
Two or More Races	5,706	50	21.74	9.89	0.90	3.07
Economically Disadvantaged	24,074	50	19.77	9.06	0.88	3.09
Individual Education Program	2,419	50	17.42	8.65	0.87	3.06
Plan 504	1,879	50	22.40	10.05	0.91	3.08
English Language Learners	2,989	48	15.74	6.95	0.81	3.07

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-12. Subgroup Reliabilities Grade 8—Mathematics*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,847	50	22.43	10.24	0.91	3.09
Male	20,734	50	22.77	10.94	0.92	3.07
American Indian/Alaskan Native	8,192	50	20.75	9.94	0.90	3.09
Hispanic or Latino	4,601	50	21.49	9.70	0.90	3.12
Asian	1,025	50	29.05	11.67	0.94	2.94
Black/African American	3,117	50	17.80	9.06	0.89	3.07
Pacific Islander	180	44	18.38	7.88	0.84	3.16
White/Caucasian	18,726	50	24.28	10.77	0.92	3.07
Two or More Races	5,741	50	22.24	10.55	0.91	3.08
Economically Disadvantaged	24,004	50	20.14	9.62	0.90	3.10
Individual Education Program	2,340	50	16.96	8.81	0.88	3.06
Plan 504	1,987	50	22.19	10.64	0.92	3.08
English Language Learners	3,294	50	15.99	7.02	0.81	3.09

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-13. Subgroup Reliabilities Science (OSTP)—Grade 5*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	18,667	45	24.88	8.04	0.86	2.96
Male	18,589	45	26.16	8.37	0.88	2.90
American Indian/Alaskan Native	7,088	45	23.35	7.97	0.86	3.00
Hispanic or Latino	3,976	45	24.83	7.71	0.85	2.97
Asian	880	45	27.71	8.18	0.88	2.86
Black/African American	2,708	43	20.56	7.78	0.85	3.04
Pacific Islander	147	42	19.76	7.34	0.83	3.01
White/Caucasian	16,929	45	27.31	7.98	0.87	2.88
Two or More Races	5,529	45	25.53	8.11	0.87	2.94
Economically Disadvantaged	22,155	45	23.77	8.01	0.86	2.99
Individual Education Program	2,634	45	22.32	8.77	0.88	2.99
Plan 504	1,678	45	25.38	8.12	0.87	2.94
English Language Learners	2,567	42	18.74	6.49	0.78	3.07

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-14. Subgroup Reliabilities Science (OSTP)—Grade 8*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	20,484	48	25.34	8.99	0.88	3.16
Male	20,512	48	26.23	9.79	0.90	3.11
American Indian/Alaskan Native	8,037	48	23.78	8.98	0.87	3.19
Hispanic or Latino	4,586	48	25.07	8.93	0.87	3.17
Asian	1,023	48	29.97	9.72	0.90	3.00
Black/African American	2,973	47	21.23	8.48	0.86	3.21
Pacific Islander	177	43	19.58	7.28	0.80	3.24
White/Caucasian	18,570	48	27.44	9.38	0.89	3.10
Two or More Races	5,630	48	25.60	9.34	0.89	3.14
Economically Disadvantaged	23,638	48	23.77	8.87	0.87	3.19
Individual Education Program	2,679	47	20.83	8.68	0.86	3.20
Plan 504	1,963	48	26.03	9.77	0.90	3.11
English Language Learners	3,301	45	18.98	6.66	0.76	3.25

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-15. Subgroup Reliabilities Science (CCRA)—Grade 11*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	24,196	60	25.78	10.17	0.88	3.53
Male	24,152	60	26.85	11.75	0.91	3.48
American Indian/Alaskan Native	9,908	60	23.69	9.72	0.87	3.52
Hispanic or Latino	4,990	58	25.35	10.14	0.88	3.52
Asian	1,149	59	31.83	12.62	0.93	3.41
Black/African American	3,677	56	21.54	8.69	0.84	3.51
Pacific Islander	205	55	21.47	8.82	0.84	3.51
White/Caucasian	21,957	60	28.26	11.43	0.91	3.49
Two or More Races	6,462	58	26.38	10.97	0.9	3.51
Economically Disadvantaged	24,019	59	24.12	9.97	0.88	3.52
Individual Education Program	5,340	58	20.36	8.27	0.82	3.49
Plan 504	2,617	58	27.84	11.49	0.91	3.49
English Language Learners	3,867	54	18.92	6.18	0.68	3.48

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-16. Subgroup Reliabilities U.S. History (CCRA)—Grade 11*

Description	Number of Students	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
Female	24,164	50	25.44	10.32	0.91	3.13
Male	24,141	50	27.08	11.56	0.93	3.06
American Indian/Alaskan Native	9,906	50	24.08	10.37	0.91	3.14
Hispanic or Latino	4,996	50	25.6	10.4	0.91	3.13
Asian	1,144	50	30.65	11.33	0.93	2.97
Black/African American	3,672	49	22	9.85	0.9	3.15
Pacific Islander	203	45	19.74	9.66	0.89	3.15
White/Caucasian	21,941	50	27.92	11.18	0.93	3.06
Two or More Races	6,443	50	26.32	10.87	0.92	3.1
Economically Disadvantaged	24,021	50	24.11	10.35	0.91	3.14
Individual Education Program	5,339	50	19.97	9.54	0.89	3.15
Plan 504	2,627	50	28	11.23	0.93	3.06
English Language Learners	3,851	48	18.41	7.61	0.82	3.19

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-17. Reliabilities by Reporting Category—ELA Grade 3*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	19	20	10.38	4.38	0.81	1.90
2	6	7	3.43	1.89	0.68	1.07
3	11	11	5.90	2.36	0.61	1.47
4	6	6	2.47	1.57	0.50	1.10
5	8	8	4.16	2.04	0.62	1.26

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-18. Reliabilities by Reporting Category—ELA Grade 4*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	15	16	8.96	3.61	0.77	1.73
2	9	10	4.64	2.30	0.66	1.33
3	12	12	7.62	2.72	0.71	1.46
4	7	7	3.79	1.59	0.46	1.17
5	7	7	3.83	1.87	0.60	1.18

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-19. Reliabilities by Reporting Category—ELA Grade 5*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	15	15	9.97	3.22	0.74	1.63
2	11	11	7.34	2.47	0.69	1.37
3	11	11	8.77	2.24	0.73	1.17
4	7	7	6.21	1.25	0.65	0.74
5	6	6	3.54	1.62	0.55	1.09

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-20. Reliabilities by Reporting Category—ELA Grade 6*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	17	18	9.57	3.93	0.78	1.85
2	9	10	5.50	2.42	0.67	1.38
3	11	11	7.69	2.57	0.73	1.34
4	6	6	2.80	1.61	0.56	1.07
5	7	7	3.91	1.90	0.64	1.15

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-21. Reliabilities by Reporting Category—ELA Grade 7*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	16	16	8.58	3.44	0.74	1.77
2	11	12	6.35	2.96	0.74	1.49
3	9	9	4.95	2.36	0.70	1.29
4	7	7	4.00	1.75	0.54	1.19
5	7	8	4.13	2.10	0.61	1.31

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-22. Reliabilities by Reporting Category—ELA Grade 8*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	14	14	9.12	3.23	0.76	1.57
2	14	14	9.54	2.95	0.74	1.49
3	10	10	6.64	2.30	0.69	1.27
4	6	6	3.69	1.52	0.48	1.09
5	6	6	3.53	1.67	0.59	1.06

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-23. Reliabilities by Reporting Category—Mathematics Grade 3*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	23	23	17.17	4.81	0.86	1.78
2	7	7	5.57	1.59	0.67	0.91
3	13	13	6.89	2.92	0.72	1.54
4	7	7	4.07	1.92	0.67	1.11

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-24. Reliabilities by Reporting Category—Mathematics Grade 4*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	23	23	14.02	5.14	0.85	1.99
2	8	8	5.72	1.90	0.69	1.05
3	13	13	7.34	2.96	0.72	1.56
4	6	6	4.26	1.51	0.58	0.98

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-25. Reliabilities by Reporting Category—Mathematics Grade 5*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	23	23	14.32	4.97	0.85	1.94
2	9	9	5.84	2.10	0.68	1.19
3	12	12	6.55	2.53	0.66	1.48
4	6	6	3.61	1.59	0.58	1.03

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-26. Reliabilities by Reporting Category—Mathematics Grade 6*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	20	20	12.60	4.04	0.80	1.81
2	11	11	6.64	2.69	0.76	1.33
3	13	13	5.56	2.69	0.66	1.57
4	6	6	2.49	1.46	0.46	1.07

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-27. Reliabilities by Reporting Category—Mathematics Grade 7*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	8	8	4.10	2.10	0.65	1.24
2	13	13	5.92	3.12	0.75	1.55
3	18	18	7.32	3.69	0.75	1.86
4	11	11	4.85	2.58	0.70	1.40

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-28. Reliabilities by Reporting Category—Mathematics Grade 8*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	9	9	4.32	2.53	0.75	1.27
2	24	24	10.83	5.29	0.84	2.12
3	9	9	3.91	2.37	0.70	1.30
4	8	8	3.54	1.97	0.61	1.23

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-29. Reliabilities by Reporting Category—Science (OSTP) Grade 5*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	15	15	8.00	3.00	0.66	1.75
2	12	12	6.98	2.67	0.69	1.48
3	18	18	10.54	3.61	0.74	1.83

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-30. Reliabilities by Reporting Category—Science (OSTP) Grade 8*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	15	16	8.54	3.65	0.76	1.81
2	21	22	12.29	4.50	0.79	2.08
3	9	10	4.95	2.32	0.58	1.49

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-31. Reliabilities by Reporting Category—Science (CCRA) Grade 11*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	30	31	13.33	6.09	0.83	2.52
2	29	29	12.99	5.49	0.80	2.44

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

Table S-32. Reliabilities by Reporting Category—U.S. History (CCRA) Grade 11*

Reporting Category	Number of Items	Raw Score Maximum	Raw Score Mean	Raw Score Standard Deviation	Alpha	Standard Error
1	25	25	12.97	5.87	0.86	2.17
2	25	25	13.29	5.56	0.84	2.21

*Reliability calculations were based on the same subset of examinees that were included in the equating procedures.

APPENDIX T

DECISION ACCURACY AND CONSISTENCY RESULTS

Table T-1. Summary of Decision Accuracy and Consistency Results by Content Area and Grade—Conditional on Cutpoint

Content Area	Grade	Accuracy (consistency)	Below Basic / Basic		Accuracy (consistency)	Basic / Proficient		Accuracy (consistency)	Proficient / Advanced	
			Positive	False		Positive	False		Positive	False
				Negative			Negative			Negative
ELA	3	0.93 (0.90)	0.04	0.04	0.93 (0.90)	0.04	0.03	0.98 (0.96)	0.02	0.01
	4	0.93 (0.90)	0.04	0.04	0.93 (0.90)	0.04	0.03	0.98 (0.98)	0.01	0.01
	5	0.94 (0.92)	0.03	0.03	0.92 (0.89)	0.04	0.04	0.96 (0.95)	0.02	0.01
	6	0.94 (0.91)	0.03	0.03	0.93 (0.90)	0.04	0.03	0.98 (0.97)	0.01	0.01
	7	0.93 (0.90)	0.04	0.04	0.93 (0.90)	0.04	0.03	0.97 (0.95)	0.02	0.01
	8	0.93 (0.91)	0.03	0.03	0.93 (0.91)	0.03	0.03	0.97 (0.95)	0.02	0.01
Mathematics	3	0.94 (0.92)	0.03	0.03	0.93 (0.90)	0.04	0.04	0.95 (0.94)	0.03	0.02
	4	0.94 (0.91)	0.03	0.03	0.94 (0.91)	0.03	0.03	0.96 (0.94)	0.02	0.02
	5	0.93 (0.90)	0.04	0.03	0.94 (0.91)	0.03	0.03	0.96 (0.95)	0.02	0.02
	6	0.92 (0.89)	0.04	0.04	0.94 (0.91)	0.03	0.03	0.98 (0.97)	0.01	0.01
	7	0.91 (0.88)	0.05	0.04	0.95 (0.93)	0.03	0.02	0.98 (0.97)	0.01	0.01
	8	0.93 (0.90)	0.04	0.03	0.96 (0.95)	0.02	0.02	0.98 (0.97)	0.01	0.01
Science	5	0.92 (0.89)	0.04	0.04	0.92 (0.88)	0.04	0.04	0.97 (0.96)	0.02	0.01
	8	0.90 (0.87)	0.05	0.04	0.93 (0.89)	0.04	0.03	0.97 (0.96)	0.02	0.01
	11	0.93 (0.90)	0.04	0.03	0.95 (0.93)	0.03	0.02	0.98 (0.97)	0.01	0.01
U.S. History	11	0.92 (0.89)	0.04	0.03	0.94 (0.91)	0.04	0.03	0.96 (0.95)	0.02	0.02

APPENDIX U

SAMPLE REPORTS



STUDENT/FAMILY REPORT

COLLEGE and CAREER READINESS ASSESSMENT

Dear Family,

This report showcases your student's performance on the spring 2025 College and Career Readiness Assessment (CCRA) in key academic areas. State test results, when combined with other information (i.e., homework, classwork, report card grades, and local assessments), can help you and the teacher work together to support your student's growth.

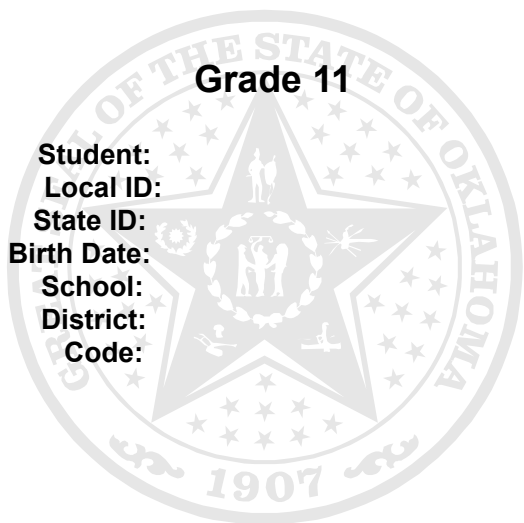
Your student's score report helps you know:

- how your student performed in each academic area
- where your student is doing well and where they may need additional support
- how your student performed compared to others
- how you can support your student at home and at school

If you have any questions, please contact your local school or the Office of Assessments at Assessments@sde.ok.gov.

Sincerely,

Ryan Walters
State Superintendent of Public Instruction

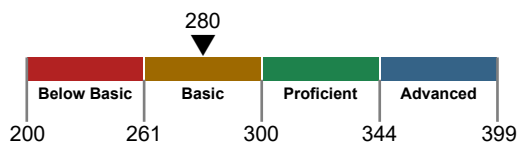


Grade 11

Student:
Local ID:
State ID:
Birth Date:
School:
District:
Code:

English Language Arts

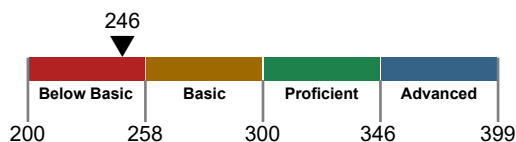
OPI: 280 Basic



demonstrates partial readiness in ELA for the next grade or course and may need additional support.

Mathematics

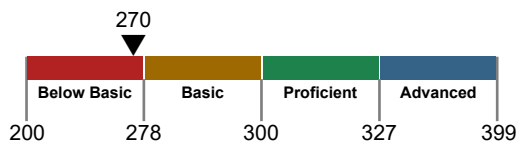
OPI: 246 Below Basic



demonstrates partial readiness in Mathematics for the next grade or course and may need targeted support to bring them to grade level.

Science

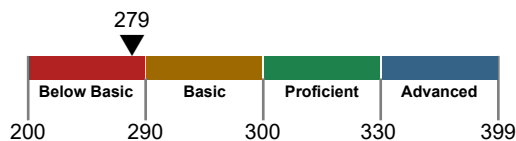
OPI: 270 Below Basic



demonstrates partial readiness in Science for the next grade or course and may need targeted support to bring them to grade level.

U.S. History

OPI: 279 Below Basic



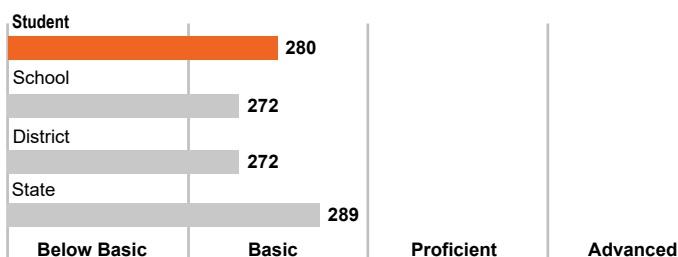
demonstrates partial readiness in U.S. History for the next grade or course and may need targeted support to bring them to grade level.

English Language Arts ► BASIC

Students scoring **Basic** demonstrate foundational skills and abilities but are still working toward being on-track. Students scoring basic typically:

- Comprehend, analyze, and synthesize information from literary and informational texts, applying limited close reading skills across a range of subject levels and complexity levels.
- Inconsistently locate explicitly stated details, make inferences about characters and actions, and identify central ideas when they are clearly stated
- Sometimes use knowledge about the author's craft and the text structure to determine the text's primary purpose and the function of key textual elements.
- Identify knowledge and ideas from across multiple related texts, comparing details that texts have in common.
- Attempt to blend multiple modes of writing to produce argumentative essays on substantive topics.
- Produce writing that responds to multiple perspectives, establishes a thesis claim that shows some clarity in thought, and provides limited analysis of the issue.
- Develop ideas and support claims with some relevant evidence that is often overly general, sometimes using basic reasoning and illustration that may be repetitious.
- Attempt to use a recursive writing process and create a simple organization with some transitions that establish relationships among ideas.
- Use language that is sometimes imprecise to convey meaning.
- Use sentence structures that are usually clear but show little variety.
- Interpret vocabulary, including basic figurative language, sometimes inferring the meaning of key words and phrases by using the context.
- Demonstrate understanding of familiar and some general academic vocabulary.
- Make inconsistent word choices and may use inappropriate tone in their writing.
- Inconsistently apply knowledge of the English language and rhetorical style to make meaning when analyzing, evaluating, producing, and revising texts.
- May recognize obvious disturbances in sentence structure.
- Demonstrate an inconsistent command of the conventions of English grammar, usage, and mechanics.

Performance Compared to School, District, and State

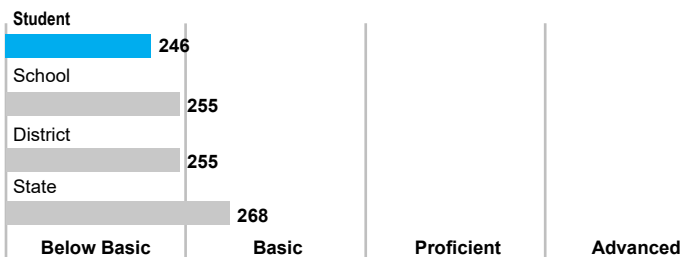


Mathematics ► BELOW BASIC

Students scoring **Below Basic** may need targeted support in developing skills and abilities to bring them to grade level. Students scoring below basic may:

- Add complex numbers and add matrices.
- Simplify square roots.
- Rewrite monomials with integer exponents to have positive exponents.
- Create linear expressions, equations or inequalities to model contexts.
- Create systems of two linear equations to model contexts.
- Solve systems of two linear equations with integer coefficients.
- Make connections between different representations of linear relationships between two variables.
- Create and use linear relationships to solve a problem.
- Multiply polynomials by monomials.
- Multiply binomials.
- Factor monomials from polynomial expressions.
- Factor trinomials.
- Add and subtract polynomials.
- Solve quadratic equations in the form $ax^2 = b$.
- Solve simple radical equations.
- Use function notation to represent functions.
- Evaluate absolute value functions.
- Evaluate simple algebraic expressions.
- Identify the shape of graphs from some of their points.
- Identify graphs of nonlinear relationships between two variables based on descriptions of characteristics.
- Read and interpret information presented in graphs, scatterplots, or tables.
- Find the median or mean of data sets.
- Find probabilities of simple events.
- Estimate expected population counts or proportions from sample counts or proportions.
- Find probabilities of simple compound events.
- Calculate simple conditional probabilities.
- Solve simple problems about geometric figures using the vertical angle theorem, the triangle angle sum theorem, or theorems about a transversal crossing parallel lines.
- Solve real-world problems using the Pythagorean Theorem.
- Solve simple problems involving perimeter, area and volume.
- Identify corresponding parts of congruent triangles.
- Translate points horizontally and vertically on a coordinate plane.

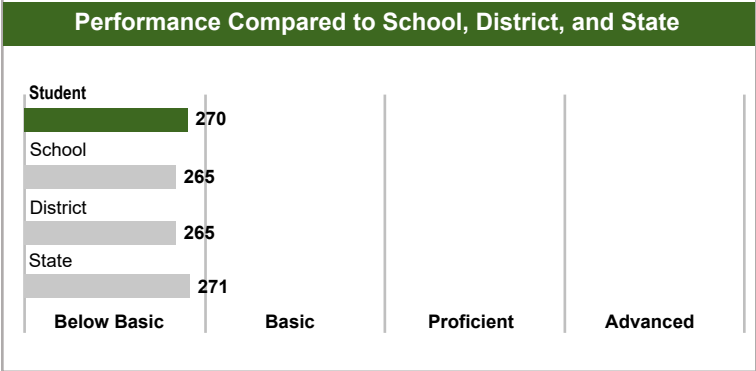
Performance Compared to School, District, and State



Science ► **BELOW BASIC**

Students scoring **Below Basic** may need targeted support in developing skills and abilities to bring them to grade level. Students scoring below basic may:

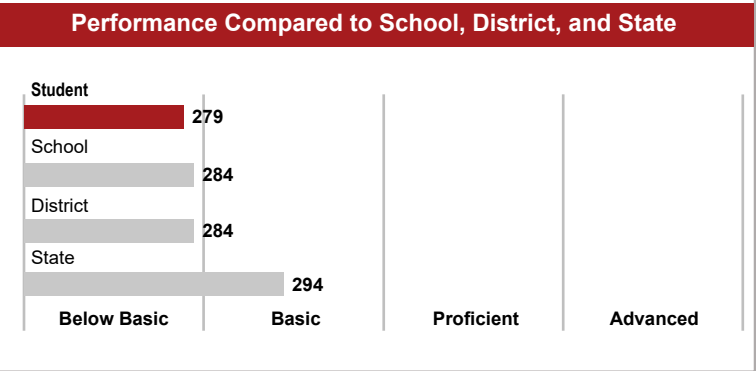
- Use basic patterns and models to identify and describe components between or within systems related to the energy of motion and the structure and properties of matter, and the relationships between energy and matter.
- Use simple mathematical models and conduct investigations to produce data or use provided data to support explanations or claims about the conservation of energy and matter during chemical reactions, the effects of different types of interactions, definitions of energy, conservation of energy and energy transfer within a system and/or system model, and how matter affects wave properties.
- Evaluate the validity and/or reliability of a simple claim about the effects of electromagnetic radiation on matter from a published source.
- Identify and describe basic relationships and construct explanations based on evidence from a variety of sources about patterns relating to the structure and properties of matter; identify how temperature or concentration affects the rate of chemical reactions; and define energy and matter in order to design solutions around defining and delimiting engineering problems and interdependence of science, engineering, and technology.
- Identify or describe basic components or relationships among components within systems and system models related to structure, function, growth and/or development of organisms, organization of matter and energy flow in organisms, cycles of matter and energy transfer in ecosystems, or energy in chemistry processes.
- Conduct investigations to produce data; use provided data to support explanations or claims about the stability related to structure and function of organisms, interdependent relationships in ecosystems at different scales, the cycling of matter and flow of energy among organisms in an ecosystem, the effect variation of traits has in a population, patterns that show evidence of natural selection or adaptation.
- Synthesize scientific information to communicate using a partial understanding of the patterns that show evidence of common ancestry, diversity, or adaptation.
- Ask questions to identify relationships about the effect of structure and function on inheritance of traits; or describe arguments based on evidence to communicate understanding of stability and change in ecosystem dynamics, function and resilience, the cause-and-effect relationships of social interactions, group behaviors, adaptation, and variation of traits.
- Identify and describe basic relationships based on evidence of the cause-and-effect relationships in natural selection, adaptation, and how the structure of DNA determines protein structure and impacts the function of the cell; or identify and describe explanations from evidence for how matter and energy is organized, cycled, and transferred within an organism or ecosystem.



U.S. History ► **BELOW BASIC**

Students scoring **Below Basic** may need targeted support in developing skills and abilities to bring them to grade level. Students scoring below basic may:

- Inconsistently apply social studies content knowledge in order to make connections between, and partially understand, how eras and events throughout United States history have influenced subsequent eras.
- Partially analyze how post-Reconstruction civil rights struggles, westward expansion, immigration, and American Indians were impacted by federal policy from 1865 to the 1920s.
- Partially evaluate how the American Industrial Revolution, the Progressive Movement, and the impact of key individuals transformed the United States from the 1870s to the 1920s.
- Partially describe the causes and effects of the United States developing into a world power through foreign and domestic policies from 1890 to 1920.
- Partially identify the factors that transformed the American government, economy, and society during the 1920s and 1930s.
- Summarize some of the major causes, events, and effects of the United States' involvement in World War II, from 1933 to 1946, transformed the nation, including the Nuremberg Trials.
- Partially describes the economic, political, and social effects of containment of Communism and Cold War from 1945 to 1975.
- Partially examine how the domestic events and policies, including various civil rights movements, transformed the United States from 1945 to 1975.
- Identify the impact of United States' foreign and domestic policy both at home and abroad from 1977 to 2001.
- Partially apply critical thinking skills, demonstrating an inconsistent ability to comprehend, interpret, evaluate, and utilize primary and secondary sources.





STUDENT/FAMILY REPORT

COLLEGE and CAREER READINESS ASSESSMENT

USING THIS REPORT TO MEET WITH YOUR STUDENT'S TEACHER OR SCHOOL

As your student's first teacher, you are a critical part of their education. It is important to remember that your student's strengths, abilities, and potential cannot be measured by a single test score. Each student grows at different rates both physically and academically. State tests help gauge how your student is growing in the knowledge and skills outlined in the Oklahoma Academic Standards. State test results, when combined with other information (i.e., report card grades, teacher feedback, classroom performance, and local tests), can help you and the teacher understand where your student is making progress and where they may need extra support. Ask your student's teachers and/or school:

- Where is my student excelling? How can I support this success?
- What do you think is giving my student the most trouble? How can I help my student improve in this area?
- What can I do to help my student with upcoming work?
- What curriculum and learning experiences do you provide to support my student?

OKLAHOMA STATE DEPARTMENT OF EDUCATION (OSDE) RESOURCES

The **OSTP Parent Portal** - is an interactive web-based tool you can use to access information about your student's OSTP results. (Note: You will need your student's state ID [STN] number and date of birth to set up an account. Your student's state ID [STN] number is located on the front of this report.) <https://okparentportal.emetric.net/login>

The **OSDE Graduation Resources** page provides links and tools you can use to help answer questions you may have about graduation requirements and career and college readiness.
<https://oklahoma.gov/content/dam/ok/en/osde/documents/services/assessments/assessment-guidance/academic-advisement/2024-25-Academic-Advisement-and-Policy-Guidebook.pdf>

The **OSDE Family Engagement** page is home to tools and resources that support partnerships between families and schools.
<https://oklahoma.gov/education/services/family-community-engagement.html>

The **OSDE Assessment Guidance for Families** page provides information and guidance on interpreting and using data from student assessments. <https://oklahoma.gov/education/services/assessments/assessment-guidance.html>

GLOSSARY OF TERMS

Performance Level: Reflect overall performance and are determined by where a student's OPI score falls within a defined range for each academic area. Oklahoma reports four performance levels: **Below Basic**, **Basic**, **Proficient**, or **Advanced**.

OPI Score: The Oklahoma Performance Index (OPI) score allows for a numerical comparison between students. For example, we can compare scale scores for students who took the 5th grade mathematics test this year with those who will take this test next year. Scale scores are not comparable across different subjects.

ADDITIONAL RESOURCES AND INFORMATION

Office of Assessments
Phone: (405) 521-3341

Office of Special Education
Phone: (405) 521-3351

Office of Standards and Learning
Phone: (405) 521-4287



APPENDIX V
REPORTING BUSINESS REQUIREMENTS

Reporting Business Requirements

156752– OSTP 2025 Oklahoma (Grade 3-8: Math, ELA, Science)

159552– CCRA 2025 Oklahoma (Grade 11: Science, US History)

Spring 2025 Version Number	Date	Updated Content Description	Updated By
1.0	10/10/24	Initial updates	Woreen Bogle
	1/9/25	Additional updates	Woreen Bogle
	1/30/25	Updates from 1/29 Planning Meeting	Woreen Bogle
	2/19/25	Media Redacted rules updated	T. LaPierre
	2/25/2025	Incorporated Program Management edits	W. Bogle

Table of Contents

I.	Overview.....	4
A.	Points of Contact.....	4
B.	Changes from 23-24.....	4
C.	Contract File Layouts and other documentation	4
D.	Risks.....	5
II.	General Information.....	5
A.	Assessments	5
B.	Reporting Phases.....	5
C.	Receivables	6
D.	Deliverables	6
III.	Pre-Assessment Processing.....	7
A.	Student Roster and Test Data Preparation	7
IV.	Post Test Assessment Administration.....	7
A.	Preliminary and Expedited Reporting.....	7
B.	Clean Up Window.....	8
C.	Student Data Processing	8
D.	Test Data	8
E.	Scan Paper Delivery and Data Denotation.....	8
F.	Data Validation.....	9
G.	Blank Books.....	9
H.	Login Discrepancy	9
I.	Spanish Tests	10
J.	Void Bubble.....	10
K.	Paper Booklet/Test Identity	10
L.	Overlay Data	10
M.	Test and Overlay Variable Reconciliation.....	11
N.	Demographic Cleanup	11
O.	Duplicate Processing.....	11
P.	Breach Processing.....	12
Q.	Merge Tests.....	12
R.	Processing ACT/SAT Score Data File.....	13
S.	Processing Scoring Data	13
V.	Student Participation and Exclusions.....	14

A.	Test Attempt Rules	14
B.	Test Design	14
C.	Not Tested Reasons.....	16
D.	Student Participation Status	16
VI.	Calculations	16
A.	Participation Status Summary.....	16
B.	Demographic Reporting.....	17
C.	Scoring Items	17
D.	Performance Levels	18
E.	Aggregate Calculations.....	18
VII.	Specific Reporting Rules	19
A.	Student Results Label(s)	19
B.	OSTP Student Report.....	19
C.	CCRA Student Report.....	20
D.	Datafile Deliverables	21
VIII.	Content Design and Development	23
A.	Assessment Content	23
IX.	Shipping Product Code Summary.....	25
A.	Reporting Products.....	25
X.	Appendix.....	26
XI.	Addenda.....	26

I. Overview

This document describes the Reporting requirements for the Spring 2025 assessments in Oklahoma. Assessments being administered by Cognia in Oklahoma are the Oklahoma School Testing Program (OSTP) and the College and Career Readiness Assessment (CCRA).

A. Points of Contact

Title	Name
Lead Program Manager	Elizabeth Garcia

B. Changes from 23-24

1. Grade 3 RSA is no longer a requirement for OSTP. There is no longer early reporting of RSA or reporting of any kind related to RSA.
2. Ways to Support text updated by SDE. Updated text used on the OSTP student report and online in English and Spanish.
3. The media redacted data file will be sorted so that state appears at the top of the file.
4. The media redacted data file masking rules updated.
5. Redaction rules updated for the media redacted data file.
6. 2 copies of the OSTP student report will be printed and shipped.
7. No build outs from the Overlay file beginning in 24-25. SDE accountability will add the records from the overlay file to their analyses.
8. No participation files are needed this year.
9. Students are allowed to test online remotely in 24-25. The students will be included in all analyses and reported in DI.

C. Contract File Layouts and other documentation

- 1) Data File Layouts and Schema Documentation Files
 - a) OKXXXX_StudentResultsLayout.xlsx
 - b) OKXXXXeMetricReportingTransfer.xlsx-used for both OSTP and CCRA results.
 - c) OKXXXXeMetricSummaryDataTransfer.xlsx-used for both OSTP and CCRA results.
 - d) OK_MediaRedacted_Layout_23-24.xlsx used for both OSTP and CCRA results. Column indicates which fields are valid for each program.
 - e) OKXXXXRosterOutboundSchemaDocumentation.xlsx
 - f) OSTPXXXXStudentLabelsSchemaDocumentation.xlsx
 - g) CCRAXXXXStudentLabelsSchemaDocumentation.xlsx
 - h) OSTPXXXXStudentReportSchemaDocumentation.xlsx
 - i) CCRAXXXXStudentReportSchemaDocumentation.xlsx
 - j) OKStudentDataDefinitions.xlsx
 - k) DemographicOverlayLayout.xlsx
 - l) eMetricPostAdminTransfer.xlsx
 - m) Pre-ID layout
 - n) File with the ACT/SAT equivalent scale scores and performance levels.

Where XXXX is the academic year

D. Risks

All stakeholders shall be notified of any risks associated to their responsible area's and be engaged as necessary.

II. General Information

A. Assessments

The CCRA testing window begins on April 1, 2025. The window for paper testers concludes on April 11, 2025. The online testing window concludes on April 25, 2025. The OSTP testing window begins April 15, 2025. The window for paper testers concludes on May 1, 2025. The online testing window concludes on May 14, 2025.

Overview of Operational Assessments			
OSTP	03	ELA, Math	Online Operational (in English)
OSTP	04	ELA, Math	Online Operational (in English)
OSTP	06	ELA, Math	Online Breach (in English)
OSTP	07	ELA, Math	Online Spanish (Math and Science only)
OSTP	05	ELA, Math, Science	Paper Operational (in English)
OSTP	08	ELA, Math, Science	Paper Breach (in English)
CCRA	11	Science	Online Operational (in English)
CCRA	11	Science	Online Breach (in English)
CCRA	11	Science	Online Spanish
CCRA	11	Science	Paper Operational (in English)
CCRA	11	Science	Paper Breach (in English)
CCRA	11	US History	Online (in English)
CCRA	11	US History	Paper (in English)
CCRA	11	US History	Online Spanish
CCRA	11	US History	Online Breach (in English)
CCRA	11	US History	Paper Breach in (in English)

B. Reporting Phases

See the Reporting Schedule for specific dates included in each phase.

- 1) Pre-ID – This is the period before the test administration window begins.
 - a) Using the Pre-ID file from SDE Cognia produces Pre-administration labels.
 - b) Pre-administration labels are provided only for those students indicating a Paper based test.
 - c) A pre-administration label is produced for each subject a student is expected to take depending on the student's grade.
 - d) Cognia provides eMetric with a data file of students that are identified as taking their tests online. The data are provided according to the OKStudentDataDefinitions file layout.
 - e) Pre-ID data is used to populate Outbound Rosters which accompany the pre-administration labels.
- 2) Preliminary Reporting
 - a) Cognia provides the SDE with preliminary student results data files.
 - b) Cognia provides eMetric with preliminary student results data.

- 3) Final Reporting - The period following state cleanup and receipt of the final Demographic Overlay file and final SSC files from SDE.
 - a) Cognia will provide SDE with the final state student results data files for OSTP and CCRA
 - b) Cognia will provide eMetric with the final data to populate Data Interaction and the Parent Portal
 - c) Cognia reporting team will provide Psychometrics data support for the Technical Report and Data Forensics deliverables.

C. Receivables

Receivable	Received from	Description	Method of Delivery
Pre-ID file (all grades)	SDE	WAVE and Non-WAVE student data	sftp
Post Admin Extract (by administration)	eMetric	Student Post-test data	Database backup
Demographic Overlay (all grades)	SDE	Student Demographic file to be used as the source of student demographic information for students with a verified student ID. The source is the WAVE file.	sftp
Student Status Code (all grades and subjects)	SDE	Lists tests to be invalidated (including writing as a subject)	sftp
CCRA ELA/Math scores	SDE	Student level data file containing the ACT/SAT equivalent scale scores for High School students	sftp

D. Deliverables

Contract	Deliverable	File Layout	Method of Delivery	Recipient
PRE-TEST ADMINISTRATION				
OSTP CCRA	Mock PreID Student File	OKStudentDataDefinitions	sftp	eMetric
OSTP CCRA	Outbound Roster	N/A	Printed/shipped	Schools
OSTP CCRA	Reporting Test Deck	eMetricReportingTransfer; eMetricSummaryTransfer	sftp	eMetric
PRELIMINARY REPORTING				
OSTP CCRA	State Student Results (OSTP: all grades and subjects; CCRA USH and Science) 1 file for OSTP and 1 file for CCRA	StudentResultsLayout	sftp	SDE
OSTP CCRA	Student Results Data File (OSTP:all grades and subjects; CCRA USH and Science)	eMetricReportingTransfer	sftp	eMetric
OSTP CCRA	Summary Data File (a file per grade) OSTP:all grades and subjects; CCRA USH and Science)	eMetricSummaryTransfer	sftp	eMetric
FINAL REPORTING				
OSTP CCRA	Student Results Data File (1 file for OSTP and 1 file for CCRA)	eMetricReportingTransfer	sftp	eMetric
OSTP CCRA	Summary Data File (one file per grade)	eMetricSummaryTransfer	sftp	eMetric
OSTP CCRA	Media Redacted file (one file for OSTP and 1 file for CCRA)	OK_Media_Redacted_Layout	sftp	SDE

Contract	Deliverable	File Layout	Method of Delivery	Recipient
OSTP	Individual Student Reports (2 copies)	N/A	printed pdf	Shipped to Districts
CCRA	Individual Student Reports (1 copy)	N/A	printed pdf	Shipped to Districts
OSTP CCRA	Student Results Labels	N/A	printed pdf	Shipped to Districts
OSTP CCRA	Student Results Data File (one file for OSTP and one file for CCRA)	StudentResultsLayout	sftp	SDE

III. Pre-Assessment Processing

Pre-Assessment activities are completed prior to the testing window. Using the pre-ID data provided by SDE, Cognia produces and ships test administration labels and the Outbound rosters to districts.

A. Student Roster and Test Data Preparation

Pre-ID data is received from the SDE to prepare for the test administration. The Pre-ID data contains student information, including demographics, and program information for students eligible to take the assessments. The source of the pre-ID from the SDE is the Accountability Reporting Application, the Student Information System in Oklahoma.

The SDE shares Student Information with the State of Texas, in the district of Texhoma, district code 70I061, located in Texas for grades three and four. The Oklahoma Student Information System includes the students located in Texas for Public School Funding purposes. The Students in the State of Texas are not included in any reporting or rostering activities.

- 1) Student information is provided by SDE in the WAVE file. Student data not available in the WAVE are provided by the districts in separate data files to Cognia directly. The data from the non-WAVE districts is provided in a different layout from the WAVE data. In 2023, Class/Teacher information was removed from the layouts for the WAVE and non-WAVE files.
- 2) Students in Grades 03 and 04, in the School District of Texhoma (70I061) Texhoma Elementary School (105) are removed from data to handed off to iCore and eMetric. These students do not take OK assessments.
- 3) Students in Texhoma district in grades 5-8 are expected to take OK assessments.
- 4) Cognia provides the final Pre-ID data to eMetric prior to the test administration window.
- 5) Cognia provides data for student test booklet labels to the iCore distribution group. These labels are printed and shipped to the district for all students taking paper tests.
- 6) Data from the Pre-ID files are used to produce the Outbound Rosters. These rosters are printed and shipped along with the pre-administration labels districts and schools.

IV. Post Test Assessment Administration

The Testing window is closed prior to processing and reporting. The commencement of the testing window initiates the following activities to report test results.

A. Preliminary and Expedited Reporting

- 1) eMetric provides the post testing data to Cognia in the post admin extract layout.

- 2) Cognia Reporting team provides Psychometrics with data to support Scaling and Equating.
- 3) Key verification will be done for all grades and subjects. Adjudication is also done for TEIs to ensure correct scoring of these items. Psychometrics and Content Development work together in this effort. This is done prior to reporting.
- 4) For Preliminary Reporting, machine scores and hand scores are available.

B. Clean Up Window

- 1) The SDE will perform post-test clean-up of Student Participation and Demographic data using the preliminary data from Cognia.
- 2) The following steps define the process to be followed:
 - a) SDE has 25 days to complete clean-up for each program.
 - b) Specific fields will be identified as editable.
 - c) SDE will provide FAY/NFAY status back to Cognia in the data files.
 - d) After cleanup by SDE, the updated data files are returned to Cognia reporting team.
 - e) The student results files will be used for data cleanup.
 - f) SDE will review student status codes and STNs.
- 3) The eMetric Reporting Portal will have a note to direct users to SDE's Accountability Systems for a more accurate student status.

C. Student Data Processing

- 1) Student IDs are provided by the SDE whenever possible. In the event the state does not provide a Student ID for a test, Cognia will assign a unique test ID for processing purposes.
 - a) If the Student ID is blank, Cognia creates a unique number using the eMetric ID. It is stored as the booklet number. The Student ID remains blank.
 - b) All created IDs will be a ten-digit number which may not begin with 0 (zero) or "100".
- 2) Student data from the Overlay datafile are used for reporting student demographic data if the student has a valid verified student ID that links to the Overlay file.

D. Test Data

- 1) Every imported test record must be associated with a student record.
- 2) Test Mode is captured in all test records as 1: Online or 2: Paper.
- 3) All Braille tests are new forms in 2025.
- 4) Duplicate test records are merged/resolved prior to reporting:
 - a) All attempted duplicate records are reviewed and updated accordingly based on SDE feedback as necessary.
 - b) In the event the student has a test record with no items attempted, Cognia does not suppress any records unless specifically directed to do so as part of the duplicate resolution or according to breach processing rules below or a completely blank book as defined in section G below.

E. Scan Paper Delivery and Data Denotation

Each Paper Booklet is scanned and delivered immediately to the Cognia Reporting team. At the time of receipt, the reporting team performs procedures to accurately identify discrepancies in the data.

The data is handed off in the agreed upon format specified in the Scan Delivery Layout and Scanning Specifications document.

- 1) Any and all discrepancies with the Scan File are resolved accordingly.
- 2) The reporting team provides a report of all discrepancies back to the Scanning department for research and/or re-scanning.

F. Data Validation

- 1) The Date of Birth field is set to blank if the value does not pass the 6 numeric value validation of (mmddyy).
- 2) All non-Alpha characters are set to blank for First Name, MI and DOB fields where there are non-Alpha characters in the fields.
- 3) SDE may provide information on any unresolved test data records that have no student association.
- 4) Ethnicity is reported as selected.
 - a) If Hispanic/Latino ethnicity is selected, the record is reported as Hispanic/Latino regardless of any additional ethnicity value selected.
 - b) If more than one ethnicity is selected and none of them Hispanic/Latino, the ethnicity is reported as Two or More Races
 - i) Valid Ethnicity Values include:
 - (1) Black/African American
 - (2) American Indian/Alaska Native
 - (3) Hispanic/Latino
 - (4) Asian
 - (5) Pacific Islander
 - (6) White/Caucasian
 - (7) Two or More Races

G. Blank Books

- 1) Records are suppressed from reporting if all the following fields are blank:
 - a) First Name
 - b) Last Name
 - c) Bubbled Student ID
 - d) Student Label
 - e) All item responses

H. Login Discrepancy

- 1) For paper tests, a comparison is made between the location where a label was sent and where the label is returned from. A login discrepancy occurs if these are different locations.
- 2) In the event of a login discrepancy, the Label location is used.
 - a) Schools/Districts can resolve during the clean-up period allowed.

I. Spanish Tests

- 1) There are no paper Spanish Tests. All Spanish tests are available online only. Spanish tests are available in grades 3-8 and CCRA.

J. Void Bubble

- 1) Preliminary reporting includes Void (that are not invalidated) records.
- 2) For final reporting, all remaining VOID booklets will be suppressed.

K. Paper Booklet/Test Identity

1. If a label exists, label always trumps bubbled information.
If a valid label exists and the barcode matched to label data, assign Student ID that was assigned to the barcode.

Apply demographic data from label data.
 - Name, DOB, District Student ID
2. If a label does not exist, and Bubbled State Student ID links to Overlay.
 - And the first 3 characters of bubbled Lname and Fname (or the inversion of Names) matches.
(Note: blank data indicates no conflict)
3. If a label does not exist, and Bubbled State Student ID links to Overlay.
 - And the first 3 characters of bubbled Lname and Fname (or the inversion of Names) matches.
(note: blank data indicates no conflict)
4. Bubbled LocalID link to District Student ID in Overlay
 - And school matches
 - And first name and last name matches.
5. Bubbled State Student ID link to District Student ID in Overlay
 - And school matches
 - And first name and last name matches.
6. Apply overlay demographic data when assigned **student ID** matched.
7. SDE will participate in resolution of any unidentified book or student.

L. Overlay Data

- 1) The Demographic Overlay file is provided by the SDE to Cognia for reporting purposes.
 - a) The demographic overlay file is the most up-to-date demographic information submitted by available in the student information systems.
 - b) If a Student ID is not unique within a school, the Program Manager will be notified for research and resolution.
 - c) A file will be delivered to the Program Manager with all requested resolutions.

- d) Demographics available in the Demographic Overlay file will be used in reporting a student if the Student ID exists in the Demographic Overlay file and has been verified. Otherwise, the demographics provided in the testing platform will be used.

M. Test and Overlay Variable Reconciliation

- 1) Set the IEP value in the test record to 0 if Student IEP from the Overlay file is not set to '1' for students that link to the Overlay file or 'Y' for students whose data is coming from the test record.
 - a. If Student IEP is '1' or 'Y' and test IEP is '1' then test IEP remains set to '1'.
 - b. If Student IEP is '0' or 'N' and test IEP is '1' then test IEP is set to '0'.
 - c. If Student IEP is '1' or 'Y' and test IEP is not '1' then test IEP is set to '2'.
- 2) Set test Plan504 as follows:
 - a. If Student Plan504 is '1' or 'Y' and test Plan504 is '1' then test Plan504 remains set to '1'.
 - b. If Student Plan504 is '0' or 'N' and test Plan504 is '1' then test Plan504 is set to '0'.
 - c. If Student Plan504 is '1' or 'Y' and test Plan504 is not '1' then test Plan504 is set to '2'.
- 3) Set Test EL as follows:
 - a. If Student EL is '1' or 'Y' and test EL is '1' then test EL remains set to '1'.
 - b. If Student EL is '0' or 'N' and test EL is '1' then test EL is set to '0'.
 - c. If Student EL is '1' or 'Y' and test EL is not '1' then test EL is set to '2'.

N. Demographic Cleanup

- 1) NFAY is not available in Preliminary Reporting. SDE will update during the cleanup period and provide back to Cognia.
- 2) Grade
 - a. Student Grade is provided by SDE in the Overlay file
 - b. Where Student Grade is not available, the Student Grade is set to the Tested Grade
- 3) If both the operational and breach forms are returned with a "tested" status, the breach form will be reported. Otherwise, if both tests are returned as the same not tested status, both tests will be reported as returned.

O. Duplicate Processing

Multiple Choice duplicate test items are identified when there is more than one record with the same verified Student ID. A record is attempted when five or more MC/PMC/TEI responses to any item(s) has been recorded. All duplicate records with less than five responses will be suppressed from reporting.

- 1) For all online test duplicate records in which there are five or more MC/PMC/TEI responses, the test record with the earliest Start Date will be used where there is a valid participation status
 - a) All records with a greater Start Date will be flagged as 'Do Not Report-Duplicate.'
- 2) For all paper test duplicate records in which there are five or more MC/PMC/TEI responses, all paper tests will be reported in Preliminary Reporting and SDE will resolve all paper duplicates for Final Reporting.

- 3) For all test duplicate records where there are five or more MC/PMC/TEI responses on the Online test and five or more MC/PMC/TEI responses on the Paper test, the Online test will be the record of source.
 - a) The duplicated Paper test will be flagged as ‘Do Not Report – Duplicate.’
- 4) For duplicate online tests with less than five MC/PMC/TEI responses, the test with the lower TestDate will be reported. The other test will be suppressed from reporting.
- 5) For all test duplicate records where there are two or more Paper records with less than five MC/PMC/TEI responses, the Paper record with the earliest Bubbled valid Test Date will be the record of source.
 - a) In the event of all duplicate paper tests that do not have a Bubbled Test Date, the Booklet number with the lowest sequence number will be the source of record.
- 6) For duplicate cases with online and paper records with less than five MC/PMC/TEI responses, the online record is reported. The paper record is suppressed from reporting.
- 7) Any duplicates not resolved are included in the data reported to eMetric. These duplicates are included in aggregations based on the participation status of the test and the schooltype.

P. Breach Processing

- 1) A valid breach test needs to be accompanied by an Invalidated operational test.
- 2) An Invalidated operational test needs to be indicated in the SSC file from SDE.
- 3) A breach test is Invalidated when an Invalidated operational test does not exist in SSC file, matching by STN and Subject.
- 4) Blank breach tests will be voided and suppressed from reporting if the student has another test in that subject.
- 5) If blank breach test is the only test for that subject, then it will be reported according to the participation status hierarchy below.
- 6) A student whose complete ELA test includes a valid Breach form (either for the writing prompt or the rest of the ELA test) will be treated as a valid participant. The form created will identify these students in data deliverables.
- 7) If either form (writing prompt or machine scored portion of the test) is an invalidated breach, it will not be merged with the rest of test. Only the operational portion or valid breach portion of the test is reported.

Q. Merge Tests

- 1) If we have more than one Writing booklet and one ELA booklet for the same student,
 - a) If the writing scores are the same, the writing score associated with the lower booklet number is merged with the ELA booklet.
 - b) If the scores are different between the Writing booklets, the ELA test will be reported with the Writing score with the lower test date (or Test ID). Cognia will send a report to SDE with the different scores. SDE will decide which Writing score to merge with the ELA test for Final Reporting.
 - c) If the Writing booklets are from different grades, merge the Writing booklet with the grade that matches the ELA grade.
- 2) If we have multiple ELA booklets and one Writing booklet for the same student, all ELA tests are reported with the same Writing score from the Writing booklet.
- 3) If we have an ELA booklet with no associated Writing booklet, the Writing score is reported as “B” for Blank.

- 4) If we have a Writing booklet with no matching ELA booklet, an ELA booklet is built out with blank ELA item responses.
- 5) If either the writing prompt or rest of the ELA test is on the paper and the other is online, the sections will not be merged.

R. Processing ACT/SAT Score Data File

The SDE will provide a data file containing the ACT/SAT equivalent scale scores and performance levels for High School students. The file will be used to populate the CCRA student report and student results data files.

- 1) Cognia will link to the data file using the Student's STN.
- 2) If performance level column is -1 or 0 this indicates the student did not test in the subject. The test will be reported as No Score Available on the student report.
- 3) If a student does not link to the ACT/SAT file, the ELA and Math will be reported as No Score Available on the student report.
- 4) The following cleanup/checks will be done on the ACT/SAT score data file from SDE:
 - a) Check for duplicate STN.
 - b) Check for valid STN that links to the overlay file.
 - c) Score range check
 - d) -1 values will be blanked.
- 5) If an STN does not link to the overlay file an updated STN will be provided by SDE or an updated overlay file with the added STN will be provided.
- 6) All demographic information is taken from the overlay file.

S. Processing Scoring Data

Scoring division will provide Reporting Services with the scores from all tests.

- 1) Each score record will be associated with a Booklet ID or a Test ID
 - a) If a score record is received without an associated Test or Booklet ID, resolution will be attempted with the Scoring Division
- 2) Every score record will contain a valid score value.
 - a) A validation of score values will be performed.
 - i) Multiple Choice responses must be a valid value to be considered attempted.
 - (1) Valid values will be A, B, C, D, blank and * for items with multiple response when only one should be given.
 - (2) Blank values will not be considered a response attempt.
 - ii) Technology Enhanced Items will be administered online only and scored based on the scoring rubric.
 - iii) Multiple Part Selected Response Items will be combined when each part has a valid response attempt.
 - (1) Valid values will be A, B, C, D or blank.
 - (2) The two parts will be combined for a final response.
 - iv) Writing Composite Score will be based on a single holistic rubric.
 - (1) Responses are 30% double scored, with a score range of 1-4. A 3rd score is required if scores are non-adjacent, or non-scorable codes do not match; the third score will be human scored. A final score is then calculated.

- v) Constructed response scores will be provided in ELA in grades 3,4,6, and 7.
- b) Score validation for each individual score is captured as follows:

Raw Data Value	Description	Reported Value	Point Value	Count as item attempted? (Y/N)
1-4 (per scorer)	Score	Final score	1-4	Y
NE	Non-English	L	0	Y
B	Blank response	N	0	N
OT	Off Topic	O	0	Y
NS	No Score	N	0	Y
RTP	Repeats the Prompt	P	0	N

- 3) All unresolved scoring records will be included in a report to the Scoring Division, as well as the Program Manager for research and resolution.
- 4) All scoring records will be resolved prior to reporting.

V. Student Participation and Exclusions

A. Test Attempt Rules

- 1) Test Attempted indicates that a student has answered a minimum of five (5) operational MC/PMC/TEIs test item(s) within a content area, regardless number of sessions.
 - a) Each of the five items must not be indicated as flawed or otherwise not scorable.
 - b) Items not able to be converted to Braille must be identified and excluded from attempted rules
- 2) In grades 5 and 8 English Language Arts (ELA) tests even if the writing composition is present, the student would still need to have attempted at least five operational multiple-choice test items to be considered meeting attemptedness
- 3) If the student doesn't meet test attemptedness then the test is reported with a Did Not Attempt status if no other not tested reason applies.
- 4) If there is no valid attempt, the record will use the Participation Status guidelines.

B. Test Design

Each test will be delivered Online or Paper. Operational items will be included in Raw Score. Raw score items will be a single common block across all forms.

Grade	Subject	Form(s)	Items included in Raw Score	Item Types
03-05	Mathematics	Paper Operational Paper Breach	If counttowardsstudentscore =Yes in NTS	Selected response items (Single part) only.(Grade 4 and 5 TEIs)
06-08	Mathematics	Online Operational: A1 Online Breach Paper Operational Paper Breach Online Spanish	If counttowardsstudentscore =Yes in NTS	Selected response items (Single part) and TEIs.
03-08	ELA	Paper/Online Operational Paper/Online Breach Online Spanish	If counttowardsstudentscore =Yes in NTS	A Writing Composition is present at grades 5 and 8. All other items are selected response items (Single or Multiple parts and TEIs).
05	Science	Paper Operational Paper Breach	If counttowardsstudentscore =Yes in NTS	Selected response items (Single part) and TEIs
08	Science	Online Operational Online Breach Online Spanish Paper Operational Paper Breach	If counttowardsstudentscore =Yes in NTS	Selected response items (Single or Multiple parts) and TEIs.
11	Science	Online Operational Online Breach Online Spanish Paper Operational Paper Breach	If counttowardsstudentscore =Yes in NTS	Selected response items (Single or Multiple parts) and TEIs.
11	US History	Online Paper Online English with Spanish TTS	If counttowardsstudentscore =Yes in NTS	Selected response items (Single or Multiple parts) and TEIs.

1) Item Reporting Categories

- a) Standards will be reported for all content areas.
- b) ELA Tests for grades 5 and 8 will have writing subtest information reported.
- c) Minimum item counts
 - i) Any content area attempt will be considered to have attempted all standards.
 - ii) If less than 6 points are included in a standard, the student score will not be reported within that category.
 - iii) All Item Reporting Categories are defined by Content Design and Development. The reporting categories are mapped and found in the CDD Test Delivery Blueprints
 - iv) The Primary Standard in NTS is the source of the Reporting Category.
 - v) Writing Prompt is its own category. The suppression rule is not applied since the number of points is less than 6.

2) Braille Item Content

- a) Paper Braille tests will be transcribed onto an answer booklet.
 - i) Paper Braille tests will be identified with the IEP Braille accommodation.
- b) Any test items that are not able to be transcribed Braille will be identified.

- i) School year 2024-2025 tests will not contain any items required to be excluded for Braille.

C. Not Tested Reasons

Not Tested Reasons are supplied by the SDE in the Student Status Code file or is flagged in iTester for online testers or the scannable for paper testers. Throughout the reporting cycles Cognia receives updated versions of the SSC.

- 1) If a student test record is assigned more than one Not Tested reason, the following hierarchy will be applied to assign only one status to a student test record:
 - a) Did Not Attempt
 - b) No Longer Enrolled
 - c) State Alternate Testing (OAAP)
- 2) If a student has participated and has a valid attempt, any Not Tested Reason indicated is ignored.

D. Student Participation Status

Student Participation Status reflects the participation of the test assessment performed by an individual student. Valid Participation Status values are provided by the SDE.

- 1) If a student has more than one of the below statuses, the Participation Status for each subject is set based on the following hierarchy:
 - a) Void, not invalidated (preliminary reporting only)
 - b) Emergency Exemption
 - c) Do Not Report
 - d) Do Not Report – Duplicate
(*May be set by Reporting*)
 - e) Invalidated Test
 - f) Invalidated Breach Tests
(*Breach tests without an Operational test that have been Invalidated will be set by reporting*)
 - g) Low Grade Invalidation
(*set by Reporting, not by the SDE*)
 - h) State Alternate Testing (OAAP)
 - i) No Longer Enrolled
- 2) If the student does not have any valid test attempt and none of the above apply, the test record is reported with a status of Did Not Attempt.

VI. Calculations

A. Participation Status Summary

- 1) Student Level Calculations will be summarized by Participation Status
- 2) Raw scores are only produced and available in datafiles and do not appear on any Printed reports (Points Possible will be provided for each subcategory)

Description	Part Status	Item Scores (Reports)	OPI Score (Reports)	Performance Level (Reports)	Data File Raw Scores	Data File Item Scores	Data File OPI Score	Data File Performance Level	Data File Student Status Code
Valid Participant	Z	ü	ü	ü	ü	ü	ü	ü	
Did Not Attempt	A								DNA
Emergency Exemption	D								EE
Do Not Report	E*								DNR
Invalidated (Breach)	F								INV
No Longer Enrolled	G*								NLE
State Alternate Testing (OAAP)	I*								OAAP
Do Not Report-Duplicate	L*								DNR-D
Invalidated Breach	M								INV-B
Low Grade Invalidation	V								INV-G
Voided Booklet	X ¹								VOID

* Student records only appear in the State results file. They do not appear in online or paper reports.

¹ Voided booklets will be provided in Preliminary and Participation state results datafile only and will not appear in Final Reporting.

B. Demographic Reporting

A student may have differing demographic information associated with each test record. This may occur when the STN is not provided and/or the STN does not link to the overlay file or SDE has not updated the record in the preliminary cleanup file. However, only one student report (OSTP) and one student label are generated for a unique student in a given school and tested grade. One of the tests will be selected, according to the below selection hierarchy, to be the associated demographics for all tests reported for that student in the eMetric data files, Student Report, and Student Labels.

Selection Hierarchy

1. Blank Student Status Code (Valid Participant).
2. Last Name is NOT null or blank.
3. First Name is NOT null or blank.
4. Class Name is provided.
5. Most recent Test
6. Largest Test ID value.

See Specific Reporting Rules section for demographics per report.

C. Scoring Items

- 1) The tests are pattern scored.
- 2) Open response scores are reported for only non-flawed items.
- 3) Writing Scores
 - a) Cut points are psychometrically determined making them consistent with other Reporting Categories and similarly interpretable.
 - b) Performance level is determined based on a psychometric scale method.
 - c) Final Score
 - i) If Scorer 1 score=Scorer 2 score, then the final score is set to either.

- ii) If Scorer 1 score is adjacent to Scorer 2 score, then the final score is the higher of the 2 scores.
- iii) If Scorer 1 and Scorer 2 both assign the same non-scorable codes, that value is the final trait score.
- iv) Otherwise, the final trait score is the third score or non-scorable value.
- d) Final Composite Score for grades 5 and 8 is calculated as follows:
 - i) Grade 5 is calculated as 5 times the final score, divided by 4
 - (1) The grade 5 score is rounded to the nearest whole value.
 - (2) Possible score values 0, 1, 3, 4 or 5. These values are used for Psychometrics.
 - (3) Reports contain the scores in the range 1-4.
 - ii) Grade 8 is calculated as 7 times the final score, divided by 4
 - (1) The grade 8 score is rounded to the nearest whole value.
 - (2) Possible score values 0, 2, 4, 5 or 7. These values are used for Psychometrics.
 - (3) Reports contain the scores in the range 1-4.
- e) Raw Score is calculated as a sum of the final calculated writing score and the operational multiple-choice raw score and used to get the final scale score and performance level.
- 4) Reporting Category Scores
 - a) Only calculated from Common, non-flawed items
 - b) Each Percent value is rounded to the nearest whole number.
 - c) The Reporting Category associated with the Writing is reported using the final score.

D. Performance Levels

- 1) Performance Levels are assigned based on the Scale Scores by grade and subject.
- 2) Four Performance Levels
 - a) Performance Level 1: Below Basic
 - b) Performance Level 2: Basic
 - c) Performance Level 3: Proficient
 - d) Performance Level 4: Advanced

E. Aggregate Calculations

- 1) Enrolled number of students (TotalN) includes students with the following participation status:
 - a) Valid Participant (Partstatus=Z)
 - b) Did Not Attempt (Partstatus=A)
 - c) Emergency Exemption (Partstatus=D)
 - d) Invalidated (Breach) (Partstatus=F)
 - e) State Alternate Assessment (OAAP) (Partstatus=I)
- 2) The number tested only includes Valid Participant status.
- 3) Aggregations include Valid Participants. However, school inclusion rules also apply:
 - a) The 3rd character of the district code is used to determine the school inclusion rules for aggregations.
 - b) Schools whose district codes contain B or P are not included in the State Summary.
 - c) Other Placement students are not included in Class, School or District aggregations. Other Placement students are identified in the Student Status Code file from SDE.
- 4) Standards Summary only include Valid Participant who meet school inclusion rules only and do not include Braille tests if there are items that cannot be brailled.
- 5) All Valid Participants are included for Performance Levels and scaled score aggregations at the Class, School, District and State levels based on school inclusion rules.

VII. Specific Reporting Rules


- 1) School information is taken from the iCore database.

A. Student Results Label(s)


- 1) Student Results Labels are printed, packed and shipped to each District for dissemination to each School separated by tested Grade.
- 2) Student Results Labels are created for each student and include all subjects tested at that particular school.
- 3) If a student has subjects at different schools, results for the subjects are reported back to the school where the test was taken.
- 4) When printed there will be 10 labels per page.
- 5) Student Results Labels are grouped by tested Grade, tested District and School and ordered alphabetically by Student Last Name, FirstName, MiddleName, Student ID (STN)
- 6) If the First and Last Name are both blank, No Name Provided is set as the student's name.
- 7) The sort is done so that No Name Provided are sorted to the top of each pack.
- 8) Labels are printed one per student per school with results from all tests taken at that school.
- 9) Demographics that are not consistent between reported subjects with not tested reasons are reported based on the selection hierarchy presented in the Demographic Reporting above. The following demographic fields are taken from the selected test record after application of the selection hierarchy, if necessary.
 - i) First Name, Last Name MI
 - ii) Gender
 - iii) Date of Birth
 - iv) Student ID
 - v) Student Grade

B. OSTP Student Report


- 1) 2 copies of each Student Reports are printed, packed and shipped to each District for dissemination to each School separated by tested Grade.
- 2) Reports are printed in color on 11 x 17 paper and folded in the middle.
- 3) Student Reports are created for each participating student.
- 4) Student name is formatted as FIRSTNAME MI LASTNAME
- 5) Packs are grouped by tested Grade, tested District and School and cpicode.
- 6) Within packs the reports are ordered alphabetically by Student Last Name, Student First Name, MI, Student ID. No Name Provided are sorted as to appear at the top of the pack.
- 7) Report templates for grades 3, 4, 6 and 7 include ELA and Mathematics results.
- 8) Report templates for grades 5 and 8 include ELA, Mathematics and Science results.
- 9) A Grade 5 template will be used for a Spanish template to be posted online. Results for Proficient in all subjects will be used.
- 10) Students that do not test in a subject related to their Grade tested receive text "No Score Available" instead of the subject results display on the front page.

- 11) Historical Scores:
- Science scores are displayed for current year results only. Due to Science only being tested in grades 5 and 8, prior year results are always unavailable.
 - 3 years' worth of scores are reported for ELA and Math where available. The current year and 2 previous years. In the reports for Spring 2025, the years will be 2025, 2024, and 2023.
 - Years with no available data are left blank on the graph and an * on the year indicates Score Not Available
- 12) In the absence of a Student First Name, the first name is replaced with "Your student" or "your student."
- 13) If a student has tested different subjects in different schools, a student report is sent to each testing school with the results for the subject taken at that school.
- 14) Reporting Category performance is reported for both subjects.
- The performance level for each reporting category is reported.
 - The Reporting Category Performance levels are: Approaching Expectations, Near/At Expectations and Achieving Expectations
 - Symbols represent the 3 performance levels as follows:
- 

Approaching Expectations



Near/At Expectations



Achieving Expectations
- 15) Lexile and Quantile was removed in 2024.

C. CCRA Student Report

- 1) 1 copy of each Student Report is printed and collated by testing school. A school pdf is created containing all Student Report PDFs for students being reported to that school.
- 2) Student Reports are created for students with any participation status.
- 3) Reports are printed in color on 8 ½ x 11 paper.
- 4) If a student has tested different subjects in different schools, a student report is sent to each testing school with the results for the subject taken at that school.
- 5) Within the school pdf the reports are ordered alphabetically by Student Last Name, Student First Name, MI, Student ID. No Name Provided are sorted as to appear at the beginning of the school pdf.
- 6) Student name is formatted as FIRSTNAME MI LASTNAME
- 7) In the absence of a Student First Name, the first name is replaced with "Your student" or "your student" wherever first name alone appears on the report.
- 8) Only current year results are reported on the student report.
- 9) The students' earned scaled score and performance level for Science and US History are reported on the front page.
- 10) CCRA students' ACT/SAT equivalent scale score for ELA and Mathematics are reported on the CCRA student report.
- 11) The earned performance level descriptor associated with the ACT/SAT equivalent scale score is printed on the report as well. See CCRA Math PLDs updated 7.28.23.pdf and CCRA ELA PLDs updated 7.28.23.pdf for the performance level descriptors.
- 12) On the back page the performance level descriptor associated with the earned performance level is printed for US History and Science.

- 13) Students that do not test in a subject receive text “No Score Available” “(Please contact your student’s school for more information.)” instead of the subject results display on the front page. There is no reported performance level. If Science or US History is not tested the corresponding comparison graphs are blank on the back page.

D. Datafile Deliverables

- 1) Student Results Datafiles are provided to SDE in a comma delimited format (csv) format.
 - a) The file contains students with their Student Status Code or results for each subject that they are eligible for based on tested grade.
 - b) Demographics reported for students are either from the Demographic Overlay file provided by SDE if the student has a validated student ID or from the test records as described above.
 - c) Rows in the data file represent students’ test records.
 - d) Naming convention for the data files to SDE: OSTPXXXXStudentResultsRelease[i].csv and CCRAXXXXStudentResultsRelease[i].csv
Where XXXX=academic year, i=1,2,3 etc.
 - e) WR_FinalScore is the final score. Final score is determined after all scores are available and arbitration is complete if necessary.
 - f) R1 score is the score given by scorer 1
 - g) R2 score is the score given by scorer 2
 - h) R3 score is the arbitrated score, if necessary.
 - i) An indicator is added to the results file indicating any student who tested remotely. If any session is marked as digitalproctoring, the remote flag is to 1 otherwise it is 0.
- 2) Student Results Datafile is provided to eMetric
 - a) eMetricReportingTransfer layout. The following tables contain the students’ results:
 - i) StudentData
 - ii) StudentScores
 - iii) Datafiles provided to eMetric only contain student records where status is Valid, Did Not Attempt, Emergency Exemption, Invalidated (Breach), Invalidated Breach and Low-Grade Invalidation. Remote tests will be included if they have one of these statuses.
 - b) eMetricSummary data file is provided to eMetric for both Preliminary and Final Reporting.
- 3) Student results data files and participation data files will follow the same layout. Participation data files to SDE will not have item and performance data populated. Demographics, form, accommodation information will be populated.
- 4) Summary Data is provided to eMetric to aid in their quality assurance process. The following files are posted to the ftp site for eMetric:
 - a) eMetricSummaryDataTransfer
 - i) Summary
 - ii) SummaryLookup
- 5) Media Redacted

Redaction is a general term describing the process of expunging sensitive data from the records prior to disclosure in a way that meets established disclosure requirements applicable to the specific data disclosure occurrence (e.g., removing, or obscuring PII from published reports to meet federal, state, and local privacy laws as well as organizational data disclosure policies). (See disclosure limitation method for more information about specific techniques that can be used for data redaction.)

- 1) Cognia provides two Media Redacted Datafiles to the SDE. One file for OSTP. One file for CCRA.

-
- a) The files are in comma separated file format (csv)
 - b) The naming convention for the file is OKOSTPXXXMediaRedacted.csv and OKCCRAXXXXMediaRedacted.csv where XXXX is the academic year.
- 2) Each file contains the Tested Grade, County Name, District or School Code (as the OrganizationID), District or School Name (as Group), and Administration Year. For each test subject the file contains total N, valid N, mean OPI, number and percent of students at each test performance level. Additionally, for each reporting category, the file includes valid count as well as number and percent of students at each reporting category performance level.
- 3) Each file is sorted by tested Grade, CountyName, District, School, Subject, Reporting Category and Reporting Subcategory
- 4) When a reporting category is not applicable to a subject/grade, or a subject is not applicable at a certain grade, N/A will be used in the cell.
- 5) To minimize the identification of any individual student, some data will be redacted. Redacted data will contain *** in the cell. Refer to redacted rules included below.
- 6) Redaction Rules
- A) A test level score set consists of Total N, Valid N, mean OPI score, performance level N and percent. A reporting category score set consists of test subject reporting category Valid N, performance level N and Percent.
- B) Apply the redaction rules to each score set in the order below.
- a) If Minimum Total N or Valid N is less than 10 then
 - a. Blank all scores in the score set, including Total N and Valid N
 - b) Otherwise, if one Performance Level Percent = 100
 - a. Blank the Performance Level N and Percent for the one with 100 percent and one additional Performance Level N and Percent (Pick the most adjacent lowest performance level to the one with 100 percent)
 - c) Otherwise, if the sum of two Performance Level Ns equals Valid N
 - a. Blank the Performance Level N and Percent for the smaller of the non-0 counts (if tie pick the lowest performance level) and blank one of the performance levels N and Percent that have a count of 0 (pick most adjacent lowest performance level to the non-0 count that was blanked)
 - d) Otherwise, if exactly one Performance Level N is less than four
 - a. Blank the Performance Level N and Percent for the one that is less than four and the one with the next lowest count (if there is a tie pick the most adjacent lowest performance level)
 - e) Otherwise, if more than one Performance Level N is less than four
 - a. Blank the performance level N and Percent for all performance levels with a count that is less than four

VIII. Content Design and Development

A. Assessment Content

- 1) CD provides Reporting with the Test Content delivery blueprint, both External/Public Blueprint and District Aggregate Reporting and Internal Target Blueprint which contains the following:
 - a) Reporting Category
 - b) Assessable Standards
 - c) Target Number of Clusters
 - d) Target Points
 - e) Percent of Total on Test
- 2) Reference Reporting Categories is mapped as follows:

Grade	Subject	Content Standard	Reporting Category	Student Report Display
3	OSTP Math	N	Number & Operations	Number & Operations
3	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
3	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
3	OSTP Math	D	Data & Probability	Data & Probability
3	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
3	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
3	OSTP ELA	4	Vocabulary	Vocabulary
3	OSTP ELA	5	Language	Language
3	OSTP ELA	6	Research	Research
4	OSTP Math	N	Number & Operations	Number & Operations
4	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
4	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
4	OSTP Math	D	Data & Probability	Data & Probability
4	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
4	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
4	OSTP ELA	4	Vocabulary	Vocabulary
4	OSTP ELA	5	Language	Language
4	OSTP ELA	6	Research	Research
5	OSTP Math	N	Number & Operations	Number & Operations
5	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
5	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
5	OSTP Math	D	Data & Probability	Data & Probability
5	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
5	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
5	OSTP ELA	4	Vocabulary	Vocabulary
5	OSTP ELA	5	Language	Language
5	OSTP ELA	6	Research	Research
5	OSTP ELA	Writing Prompt	Writing Composite Score	Writing Composite Score
5	OSTP Science	PS1.1, PS1.2, PS1.3, PS1.4	Physical Science	Physical Science
5	OSTP Science	LS1.1, LS2.1, LS2.2, PS3.1	Life Science	Life Science
5	OSTP Science	ESS1.1, ESS1.2, ESS2.1, ESS2.2, PS2.1	Earth & Space Science	Earth & Space Science
6	OSTP Math	N	Number & Operations	Number & Operations
6	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
6	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
6	OSTP Math	D	Data & Probability	Data & Probability
6	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
6	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
6	OSTP ELA	4	Vocabulary	Vocabulary

Grade	Subject	Content Standard	Reporting Category	Student Report Display
6	OSTP ELA	5	Language	Language
6	OSTP ELA	6	Research	Research
7	OSTP Math	N	Number & Operations	Number & Operations
7	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
7	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
7	OSTP Math	D	Data & Probability	Data & Probability
7	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
7	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
7	OSTP ELA	4	Vocabulary	Vocabulary
7	OSTP ELA	5	Language	Language
7	OSTP ELA	6	Research	Research
8	OSTP Math	N	Number & Operations	Number & Operations
8	OSTP Math	A	Algebraic Reasoning	Algebraic Reasoning & Algebra
8	OSTP Math	GM	Geometry & Measurement	Geometry & Measurement
8	OSTP Math	D	Data & Probability	Data & Probability
8	OSTP ELA	2	Reading/Writing Process	Reading & Writing Process
8	OSTP ELA	3	Critical Reading/Writing	Critical Reading & Writing
8	OSTP ELA	4	Vocabulary	Vocabulary
8	OSTP ELA	5	Language	Language
8	OSTP ELA	6	Research	Research
8	OSTP ELA	Writing Prompt	Writing Composite Score	Writing Composite Score
8	OSTP Science	8.PS1.5, 8.PS1.6, 8.PS2.1, 8.PS2.2, 8.PS4.1, 8.PS4.2	Physical Science	Physical Science
8	OSTP Science	8.LS1.7, 8.LS4.1, 8.LS4.2	Life Science	Life Science
8	OSTP Science	8.ESS1.4, 8.ESS2.1, 8.ESS2.2, 8.ESS2.3, 8.ESS3.1, 8.ESS3.2, 8.ESS3.4	Earth & Space Science	Earth & Space Science
11	CCRA Science	PS.PS1.1, PS.PS1.2, PS.PS1.5, PS.PS1.7, PS.PS2.5, PS.PS3.1, PS.PS3.2, PS.PS3.3, PS.PS3.4, PS.PS4.1, PS.PS4.4	Physical Science	Physical Science
11	CCRA Science	B.LS1.1, B.LS1.2, B.LS1.3, B.LS1.4, B.LS1.5, B.LS1.6, B.LS1.7, B.LS2.1, B.LS2.2, B.LS2.3, B.LS2.4, B.LS2.5, B.LS2.6, B.LS2.8, B.LS3.1, B.LS3.2, B.LS3.3, B.LS4.1, B.LS4.2, B.LS4.3, B.LS4.4, B.LS4.5	Life Science	Life Science

Grade	Subject	Content Standard	Reporting Category	Student Report Display
11	CCRA US History	1.2.A, 1.3.A, 1.3.D, 2.1.A, 2.1.B, 2.1.D, 2.1.E, 2.1.G, 2.2.B, 2.3.B, 3.1.A, 3.1.B, 3.1.C, 3.2.A, 3.2.B, 4.1.A, 4.1.D, 4.1.E, 4.2.A, 4.2.B, 4.2.D, 4.3.C, 5.1.B, 5.2, 5.3, 6.1.A, 6.1.B, 6.1.C, 6.1.D, 6.2.A, 6.2.B, 6.2.C, 6.4, 7.2.D, 8.1, 8.2, 8.3, 8.4, 8.5.A	US History	US History
11	CCRA US History	1.1, 1.2.B, 1.2.C, 1.3.B, 1.3.C, 2.1.C, 2.1.F, 2.2.A, 2.2.C, 2.3.A, 2.3.C, 3.1.D, 3.2.C, 3.2.D, 4.1.B, 4.1.C, 4.2.C, 4.3.A, 4.3.B, 5.1.A, 5.1.C, 6.3, 7.1.A, 7.1.B, 7.1.C, 7.2.A, 7.2.B, 7.2.C, 7.2.E, 7.2.F, 8.5.B, 8.6	Civics	Civics

IX. Shipping Product Code Summary

A. Reporting Products

Reporting Products is provided to iCore to identify the products that will be shipped to the client.

Contract Code	Description	Report For	Grade(s)	Report Subtype	Content Code	Qty
156752	OSTP Student Labels	1	03-08	03	00	1
159552	CCRA Student Labels	1	11	03	00	1
156752	OSTP Student Report-Parent Copy	1	03-08	02	00	1
156752	OSTP Student Report-School Copy	1	03-08	01	00	1
159552	CCRA Student Report-Parent Copy	1	11	02	00	1

X. Appendix

XI. Addenda

1. It was reported that the PBT Breach form was printed in a different item order than the CBT Breach form. Caution any merging across test modes.
2. SDE has asked that the cut scores for 3-8 Math and ELA be rolled back to the cut scores prior to standard setting in 2024. The following actions in reporting will be taken:
 - Reporting will handoff participation files for ELA and Math to SDE. These files will not have the scaled score, and achievement levels assigned.
 - Reporting will handoff a student results file for Grades 5 and 8 science. This file will contain the scaled score and achievement levels.
 - Both files will be used in the SDE cleanup window.
 - The following note will be printed on the front page of the student report “**Important Note:** This year’s results reflect recent cut score changes from the Oklahoma Commission for Educational Quality and Accountability (CEQA), which oversees the Office of Educational Quality and Accountability (OEQA). Any questions regarding interpretation of the cut scores and their validity should be directed toward the OEQA. To contact the OEQA please email info@oeqa.ok.gov or call (405) 522-5399.”
 - Only Results from Spring 2025 are reported.