This document only applies to results from 2017-2025.
Do Not Use for 2026.

Summary of the ACCESS 2.0 2016 Standard Setting Study

Overview

In July and August of 2016, the WIDA Consortium conducted standard setting studies to reexamine the ACCESS for ELLs 2.0 assessment's proficiency level scores. The following factors motivated WIDA to conduct these studies: migrating from a paper-and-pencil to an online assessment, employing a new centrally scored, revised speaking assessment, and adapting to the influence of college and career ready standards. The standard setting study occurred in two phases. The first phase, which was led by WIDA, identified scores that panelists felt represented the marginally English proficient English language learner (ELL) student on the ACCESS 2.0 listening, reading, speaking and writing domain tests. Phase one occurred in July of 2016. Phase one information was used to set up phase two, which was led by the Center for Applied Linguistics (CAL). The goal in phase two was to determine where to set cut scores between WIDA's six proficiency levels, as described by the WIDA English Language Development (ELD) standards for grades 1-12 in each of the four domains. After each phase, recommended cut scores were smoothed[1] to assure appropriate vertical articulation[2] of cut scores across grades. Information from individual domain scores was used to create composite cut scores.

ACCESS 2.0 Score Tables

The new proficiency level cuts scores are shown in Table 1 through Table 8.

Table 1. Final Cut Scores: Listening

Grade	2.0	3.0	4.0	5.0	6.0
K	229	251	278	286	308
1	236	259	291	303	327
2	245	283	314	330	354
3	262	300	331	349	374
4	275	313	343	363	388
5	285	323	354	375	401
6	294	332	363	385	411
7	302	340	370	394	420
8	308	347	377	402	427
9	314	353	383	409	434
10	325	358	389	415	441
11	335	364	394	420	447
12	342	368	398	426	452

1 Smoothing is an empirical technique used to address uneven or irregular cut scores across grades, e.g., adjusting for a lower cut score in a higher grade or an inordinately large jump in a cut score from one grade to the next.

2 An English proficient performance standard is expected to increase across grades in a systematic fashion because the language needed for English proficiency in higher grades is different from that in lower grades. Accordingly, cut scores should vertically align (articulate) to support this.

This document only applies to results from 2017-2025. Do Not Use for 2026.

Table 2. Final Cut Scores: Reading

Grade	2	3	4	5	6
K	241	259	279	289	310
1	264	286	304	315	334
2	283	307	326	337	355
3	297	323	342	352	370
4	307	335	354	364	382
5	316	345	364	373	391
6	323	353	373	382	399
7	329	360	380	389	406
8	335	366	386	395	412
9	340	372	392	401	418
10	344	377	397	406	423
11	348	382	402	410	427
12	352	386	407	414	432

Table 3. Final Cut Scores: Speaking

Grade	2	3	4	5	6
K	191	250	301	349	392
1	205	261	311	361	403
2	220	273	322	374	415
3	234	283	332	386	425
4	246	293	342	397	435
5	258	302	350	407	443
6	268	310	360	417	451
7	277	317	369	425	457
8	284	323	377	433	463
9	290	328	385	440	468
10	295	333	393	446	471
11	299	337	400	451	474
12	302	340	406	455	476

This document only applies to results from 2017-2025. Do Not Use for 2026.

Table 4. Final Cut Scores: Writing

Grade	2	3	4	5	6
K	234	271	311	367	389
1	238	275	337	382	405
2	242	279	341	388	411
3	247	283	346	394	418
4	266	288	351	401	425
5	267	293	356	407	433
6	268	298	361	413	441
7	273	305	367	419	450
8	281	311	372	424	459
9	289	319	378	430	469
10	298	326	385	436	479
11	308	335	391	441	490
12	318	344	398	447	501

Table 5. Final Cut Scores: Oral Composite

Grade	2	3	4	5	6
K	210	251	290	318	350
1	221	260	301	332	365
2	233	278	318	352	385
3	248	292	332	368	400
4	261	303	343	380	412
5	272	313	352	391	422
6	281	321	362	401	431
7	290	329	370	410	439
8	296	335	377	418	445
9	302	341	384	425	451
10	310	346	391	431	456
11	317	351	397	436	461
12	322	354	402	441	464

This document only applies to results from 2017-2025. Do Not Use for 2026.

Table 6. Final Cut Scores: Comprehension Composite

Grade	2	3	4	5	6
K	237	257	279	288	309
1	256	278	300	311	332
2	272	300	322	335	355
3	287	316	339	351	371
4	297	328	351	364	384
5	307	338	361	374	394
6	314	347	370	383	403
7	321	354	377	391	410
8	327	360	383	397	417
9	332	366	389	403	423
10	338	371	395	409	428
11	344	377	400	413	433
12	349	381	404	418	438

Table 7. Final Cut Scores: Literacy Composite

Grade	2	3	4	5	6
K	238	265	295	328	350
1	251	281	321	349	370
2	263	293	334	363	383
3	272	303	344	373	394
4	287	312	353	383	404
5	292	319	360	390	412
6	296	326	367	398	420
7	301	333	374	404	428
8	308	339	379	410	436
9	315	346	385	416	444
10	321	352	391	421	451
11	328	359	397	426	459
12	335	365	403	431	467

This document only applies to results from 2017-2025.
Do Not Use for 2026.

Table 8. Final Cut Scores: Overall Composite

Grade	2	3	4	5	6
K	229	261	293	325	350
1	242	274	315	344	368
2	254	289	329	359	383
3	265	300	340	371	396
4	279	309	350	382	406
5	286	317	358	390	415
6	291	324	365	399	423
7	298	331	372	406	431
8	304	337	378	412	438
9	311	344	385	418	446
10	318	350	391	424	453
11	325	356	397	429	459
12	331	362	402	434	466

Standard Setting Panelists

In the fall of 2015, states were asked to recommend panelists to be part of phase one and phase two of the standard setting. Member states were asked to nominate panelist candidates who had experience teaching or directly supporting the instruction of ELs in K-12 public schools in the United States or its associated territories. States were asked to provide panelist nominations for educators having experience or qualifications in English language arts, mathematics, science, social studies, and special education.

In total, members from 35 WIDA states or U.S. territories participated in at least one of the standard setting sessions. A list of participating states is shown in the Table 9. Table 10 summarizes panelists' qualifications in each phase.

Table 9. Standard Setting Panelists by State

State	Participants		State	Participants	
	Phase 1	Phase 2		Phase 1	Phase 2
AK	1	1	MO		3
AL	1	1	MT	2	1
CNMI*	1	1	NC	1	5
CO	1	2	ND	2	
DC		1	NH	3	4
DE	2		NJ	2	
FL	2	2	NM		3
GA	3		NV	3	3
HI	2	2	OK	2	1
ID		2	RI	1	1

State	Participants		State	Participants	
	Phase 1	Phase 2		Phase 1	Phase 2
IL	6	3	**SD**	3	
IN	1	1	**TN**	2	1
KY	2		**USVI***		1
MA	3	2	**UT**	1	1
MD	2	4	**VT**	2	1
ME	2	2	**WI**	2	2
MI	1	1	**WY**	2	1
MN	1	1			

* CNMI=Commonwealth of Northern Marianas Islands, USVI=U.S. Virgin Islands

Table 10. Standard Setting Panelist Characteristics

Phase One Panelist Characteristics	Phase Two Panelist Characteristics
• A total of 59 Panelists • 30 States (79% of WIDA) represented • Teacher and Policymaker mix • 88% Female • 88% with a Master's degree or more • 81% with 10 or more years of experience • 80% familiar with WIDA standards, ELP levels and ACCESS	• A total of 54 panelists • 29 States (76% of WIDA) represented • Mostly teachers • 93% Female • 81% with a Master's degree or more • 83% with 10 or more years of experience • 75% familiar with WIDA standards • 85% familiar with WIDA ELP levels and ACCESS

Standard Setting Methodology and Procedures

Table 11 shows group assignments for phase one and phase two of the standard setting study.

Table 11. Panelist Group Assignments by Grade for Phase One and Two

Group	Grade											
	1	2	3	4	5	6	7	8	9	10	11	12
1	Three tables with approximately 5 panelists at each table											
2				Three tables with approximately 5 panelists at each table								
3						Three tables with approximately 5 panelists at each table						
4								Three tables with approximately 5 panelists at each table				

Grayed areas represent grades in which groups conducted standard setting activities.

Phase one involved two activities. The first was to describe and document the listening, speaking, reading, and writing characteristics that represented a minimally competent English proficient EL at each grade. Each group displayed the agreed-upon description of this minimally competent student so that all panelists could refer to it.

The second activity identified domain scale scores that represented the minimally competent student performance at each grade and for each domain test. For the listening and reading tests, a modified Bookmark method was adopted using an ordered item book online program. For the speaking and writing tests, an adapted Body-of-Work method was used, again with online materials. In this Body-of-Work method, panelists reviewed portfolios of students' work (consisting of all of the student's responses to the speaking or writing prompts) which were ordered from the highest to the lowest scores. Judgments were then made about where a minimally competent student's performance would lie. For each domain test, there were three rounds. In round one, panelists assigned ratings based on the group's description of the minimally competent student. In round two, panelists could revisit their initial ratings after they discussed round one ratings and saw the group's median scores from round one. Prior to the last round, impact information (i.e., the proportion of students above and below the median cut score) based on the group's round two median scores were shared. Panelists then completed round three. The round three ratings[3] were used as the group's final recommendation. For all rounds, it was made clear that panelists did not need to change their ratings, only that their final ratings be informed by conversations, group median ratings, and impact information.

Phase two focused on setting cut scores for the six proficiency levels described by the WIDA English Language Development standards in each of four domains (listening, reading, writing and speaking) for grades 1-12. The methods and procedures used for this phase followed closely those used in phase one, with some differences that arose from the different goals of the two phases. Phase two only had two rounds. To start round one of phase two, initial recommended cut scores which were calculated prior to convening phase two were provided for all domains except speaking. To calculate the proposed cut scores, CAL examined smoothed recommended cut scores from phase one, reviewed the previous ACCESS 1.0 cut score information, and created initial recommended cut score values. During phase two, it was made clear to panelists that these initial recommendations were only a starting point. Panelists were free to make changes to the initial recommended values, and in many cases, they did. Prior to round two, impact data was shared. For phase two, round two ratings were used as panelists' final recommended cut scores. Following this, CAL used a smoothing procedure on the phase two final recommendations to create the recommended domain, grade, and proficiency level cut scores.

Notice that in both phase one and phase two, standard setting activities occurred in grades 1 through 12. Kindergarten was not part of the standard setting studies. To obtain Kindergarten cut scores, researchers at WIDA fit a regression model based on grades 1, 2, and 3 cut scores. They then interpolated results to obtain preliminary cut scores. Once this was completed, CAL examined the preliminary cuts to see if actual Kindergarten student performances were consistent with identified cut scores. After deliberations between CAL and WIDA, final recommended Kindergarten cut scores were created.

[3] Two groups (group 2 and 3) were unable to finish the third round of speaking test because of time constraints. For these groups, impact data was shared after round one.

Communication and Approval of Final Cut Scores

On September 23, 2016, a memo was sent to WIDA states presenting the recommended cut scores for grades 1 to 12. Recommended cut scores for Kindergarten were not provided since analysis was underway at that time. On September 27, 2016, a WebEx session was held to give WIDA member states the opportunity to ask questions and/or seek clarification about the standard setting recommendations. Most comments at this WebEx related to the impact of the new cut scores on accountability and ELL reclassification criteria. A memo was sent to the WIDA's Executive Committee outlining the procedures used to establish Kindergarten cuts, and on October 11, 2016, a WebEx session was held for the Executive Committee's approval of the recommended cut scores, including Kindergarten. The cut scores were approved at that meeting. Subsequently, approved cut score values were provided to the Data Recognition Corporation (DRC) to be used for the 2016–2017 ACCESS 2.0 score reports. States were also given the opportunity to receive an updated 2015-16 state student response (SSR) file with post-standard setting scale scores and proficiency level values.