

ACCESS for ELLs: Speaking

The ACCESS for ELLs Speaking test assesses students' abilities to produce spoken academic English, not conversational language skills. The test is designed to mimic an instructional scenario in which a student responds without practice or linguistic support.

Students taking the Speaking test join a conversation between a virtual test administrator and model student. The virtual test administrator asks questions and the model student provides an answer before the virtual test administrator prompts the test-taker to respond.

The test administrator will play a CD that contains the virtual test administrator and model student audio. When your student responds, you score the response as the CD continues to play, giving your student the next prompt.

TEST MATERIALS

Test Administrator Script, test booklets, pencils,
Listening & Speaking CD, CD player and speakers



Students listen to an audio CD and speak aloud. The test administrator scores the response during testing.

Tell Students

- Think before you speak!
- Match the model student's response. It demonstrates how much and what kind of language to use in an answer.
- Say as much as you can to show the words you know and your ability to connect ideas. Use words like "and," "because," "then," and "also." There is no penalty for being cut off in the middle of a thought.
- Speak loudly and clearly.
- Use the pictures to help you think of what to say.

Keep in Mind

- Students can score well for concise, clearly delivered responses that address the task with appropriate word choices.
- **Sufficient space and privacy is critical to student confidence and success on the Speaking test.**
- The Speaking test focuses on communication and comprehensibility. The questions are designed to elicit language at progressively higher proficiency levels. Students begin by responding with single words or short phrases and work up to tasks that require more connected ideas and extended speech.
- **Responses are scored on fluency, vocabulary, and discourse, relative to the model student's response.** Students can score well even with factual inaccuracies or grammatical errors, as long as they don't prevent clear communication. Responses are scored against a simplified version of the WIDA Speaking Rubric.



SCHEDULE 50 MINUTES Typical in-seat testing time: 15-35 minutes

Test times depend on the grade-level cluster and tier of the test as well as the proficiency of the student. Response times are unique to the task and always sufficient for high-scoring responses.

ACCESS for ELLs Speaking Scoring Scale

Score point	Response characteristics
Exemplary use of oral language to provide an elaborated response	<ul style="list-style-type: none"> • Language use comparable to or going beyond the model in sophistication • Clear, automatic, and fluent delivery • Precise and appropriate word choice
Strong use of oral language to provide a detailed response	<ul style="list-style-type: none"> • Language use approaching that of model in sophistication, though not as rich • Clear delivery • Appropriate word choice
Adequate use of oral language to provide a satisfactory response	<ul style="list-style-type: none"> • Language use not as sophisticated as that of model • Generally comprehensible use of oral language • Adequate word choice
Attempted use of oral language to provide a response in English	<ul style="list-style-type: none"> • Language use does not support an adequate response • Comprehensibility may be compromised • Word choice may not be fully adequate
No response (in English)	<ul style="list-style-type: none"> • Does not respond (in English)

Scoring processes

Select the score point that best describes the overall response relative to the qualities of the model

- Check to ensure each bullet point is met
- If not, check one level below

Scoring notes & rules

- For P1 tasks, assign a score of **Adequate and above** if the response includes more than one word in English. This includes an article plus noun (e.g., “a chair”), and words repeated verbatim from the model.
- For P3 and P5 tasks, students may take up and use language from the model and should not be penalized for this. This is particularly relevant for personal-preference tasks.
- At all task levels, simply repeating or reading all or part of the task question should be scored **Attempted**.
- At all task levels, responses of “I don’t know” should be scored **Attempted**.

Off-task response: The response shows no understanding of or interaction with the prompt. It may answer another, unrelated task. A response that is entirely off task receives a score of **Attempted**.

Off-topic response: The response shows a misinterpretation of the instructions. An off-topic response is *related* to the prompt, but does not address it. (Note that this does not refer to task completion—for example, if a student is asked for 3 reasons and gives 1, this should be scored based on language use and is not considered off topic.) **The maximum score for an off-topic response is Adequate.** If any part of the response is on topic, the entire response is scored as on topic.

For scoring use only

ACCESS for ELLs Paper Test Domain: Speaking

Speaking Test Notes:



The ACCESS for ELLs Paper Speaking test is an individually administered test that standardizes test administration across students. Speaking test items are media-delivered. Speaking test audio is provided on the same Listening and Speaking test CD. Your role as the Test Administrator is to monitor the student as they take the test and to score the student's responses during the test administration.

The Speaking test provides ELLs with the opportunity to demonstrate their academic English language proficiency in speaking across the WIDA ELD Standards through a set of constructed-response tasks. The Speaking test is tiered. Students will either take the Tier A form or a Tier B/C form, which are included in the same Speaking Test Booklet.

The Speaking test is multi-modal. The student hears audio input and also sees the input as text in their Speaking Test Booklet. This presentation format supports the student in understanding test input. Media-delivery of the Speaking test means that an audio recording will guide the student through the Speaking test.

The audio recording includes two voices: a Model Student and a Virtual Test Administrator.

- Each task on the Speaking test is preceded by a Model Student task and response. The questions posed to the Model Student are the same proficiency level as the tasks to which the student will respond, allowing the Model Student to demonstrate the expected language use at a given proficiency level. In most cases the model questions are designed to be parallel to but not exactly the same as the examinee questions. The model student also has an important function for scoring using a scoring scale that is designed to evaluate student responses relative to the Model Student's response.
- The Virtual Test Administrator guides the student through the test and asks the student questions designed to elicit language at targeted proficiency levels. While the Virtual Test Administrator will instruct and guide the student through the Speaking test, you may also need to assist the student in navigating test materials (e.g., turning the page when prompted).

The Speaking test includes standardized, built-in response time for every task. The amount of time varies according to the grade-level cluster, tier, and proficiency level of the task, and ranges from 15 to 50 seconds in Grades 1–3 and from 15 to 45 seconds in Grades 4–12. Five seconds prior to the end of the response time, a tone will sound. This alerts the student that the response time is coming to an end. Students may not require the entire time allotted. After the response time has ended, the test audio will automatically continue to the next Speaking task.

ACCESS for ELLs Paper Speaking Scoring

What is rater reliability?

Rater reliability is a technical term used in testing that refers to the consistency of scores awarded by multiple raters. Rater reliability is calculated by having the same group of raters evaluate the same students. We can then compare how consistent these raters are in evaluating the same spoken responses. If the raters are very consistent in how they evaluate the students' spoken language, their rate of reliability will be high. However, if they award many different scores to the same students, then the rate of reliability will be low, which is problematic as it indicates a risk of measurement error and students receiving scores that are not fair or appropriate. Reliability can be measured easily and reported as a percentage from 1 to 100.

On the ACCESS for ELLs Paper Speaking Test, raters may award one of five different score points: Exemplary; Strong; Adequate; Attempted; No Response. In order to calculate the reliability of the raters, these score points can be converted into numbers, as shown in the table below.

Score Point	Numeric Score
Exemplary	4
Strong	3
Adequate	2
Attempted	1
No Response	0

Listen and Calculate Rater Reliability

Indicate the proficiency level in the **Task** column for each sample item. Listen to sample student responses and write your scores in the **Rater 1 scores** column. After scoring, write down a partner's scores in the **Rater 2 scores** column. Calculate the difference for each task and add the totals. To calculate the Adjacent %, take the sum of the differences and divide it by 0.12. Then, subtract the Adjacent % from 100 and this will equal the **Reliability %**.

Item	Task	Rater 1 scores	Rater 2 scores	Difference
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
Sum				
Adjacent %				
Reliability %				