Scoring Process for TELPAS Speaking and Writing Responses

The Texas English Language Proficiency Assessment System (TELPAS) is an English language proficiency assessment aligned to the Texas English Language Proficiency Standards (ELPS). TELPAS is designed to measure the progress that emergent bilingual (EB) students make in learning the English language. TELPAS includes several types of open-ended questions that allow EB students to demonstrate their proficiency in the English language.

TELPAS listening and speaking assessments for EB students in grades 2–12 have test questions known as speaking prompts. This type of question requires students to orally record their 45- or 90-second responses to a speaking prompt. TELPAS reading and writing assessments for EB students in grades 2–12 have two types of open-ended questions—sentence rewrite and constructed-response. The sentence rewrite questions require students to identify errors in grammar and syntax and correctly rewrite a sentence. The constructed-response questions require students to draft written responses.

Student responses to open-ended questions are evaluated as a whole based on pre-established criteria identified and detailed in the scoring rubrics. Scoring rubrics are used to determine the effectiveness of each response. This document outlines the scoring process the Texas Education Agency (TEA) uses for students' responses to TELPAS speaking prompts, sentence rewrite questions, and constructed-response questions.

Overview of Scoring

A hybrid scoring model is used for all TELPAS speaking responses and for TELPAS writing constructed responses and sentence rewrites in grades 4–12. However, TELPAS writing constructed responses and sentence rewrite in grades 2 and 3, are 100 percent human scored.

TELPAS writing constructed responses, sentence rewrite and speaking responses are initially scored by an automated scoring engine (ASE) and at least 25 percent of student responses are then routed to human scorers. The scores generated by human scorers are used throughout the scoring window to monitor scores that are being generated by the ASE. This process is similar to a second reader in human scoring models.

Pre-Scoring Activities

Before scoring begins, a few activities occur which optimize how TEA assesses the ELPS on TELPAS.

- Developing New Questions
  Development activities begin with content and assessment experts developing the prompts, passages, and questions that will assess the ELPS. Committees of Texas educators (e.g., teachers, instructional specialists, district coordinators, principals, etc.) review passages and questions and provide input and feedback. Once passages and questions are reviewed and edited accordingly, they are approved by TEA and are ready for field testing.

- Field Testing New Questions
  Field testing questions is the process of gathering performance data on prompts, passages, and questions, but it does not impact students' scores on TELPAS. TEA conducts field-testing of newly developed TELPAS questions each year through embedded field-test questions.

Scoring Process for TELPAS Speaking and Writing Responses

Scoring Activities

The steps outlined below each have an important role in the overall scoring process.

```
Anchor Approval
Meetings
    → Training
        → Scoring
            → Reporting
                → Rescores
```

Step 1: Anchor Approval Meeting

A selected group of student responses are collected during field testing and are scored using scoring rubrics, i.e., the TEA-approved scoring criteria. Student responses from the field-tested questions are taken to anchor approval meetings. The purpose of anchor approval meetings is to apply the scoring rubrics to sample student responses and establish scoring boundaries. Student responses used to establish scoring boundaries are called "anchor responses", and these responses clearly exemplify each score point in the rubric. During the anchor approval meetings, educators identify anchor responses, provide consensus scores for these responses, and discuss the rationale for the scores these responses receive. These discussions are used to develop training notes or annotations for the responses. After the meeting, item sets intended for use as practice and qualification sets are created based on the approved anchor responses. These item sets are used to train and qualify human scorers.

Step 2: Training Human Scorers and ASE

Both human scorers and the ASE are trained with TEA-approved student responses.

- Training Human Scorers
  Scorers are assigned to score specific grades based on their experience and educational background. For each question type, human scorers go through a similar process. They study and review the anchor responses. Then scorers are given the opportunity to practice the different ways students are responding to the question. Scoring directors and scoring supervisors are available for clarification and guidance during the process and support the scorers in their training as needed.

  After completing practice sets, human scorers must demonstrate a complete understanding of the rubrics by successfully completing a qualification set. The qualification set is a quality measure to ensure that scorers are following the rubric and can accurately apply scores. Only scorers who

complete the training and pass the qualification set may score student responses. Scorers who do not pass the qualification set are not allowed to score student responses for TELPAS.

- o **Training for Speaking Responses**
  Two rubrics, a 2-point rubric and a 4-point rubric, are used to score responses to speaking questions. Human scorers are trained to apply the rubrics by using the approved anchor sets and then must successfully score at least two practice sets. Both the anchor and practice sets include annotations explaining why and how the score was assigned.

- o **Training for Sentence Rewrites**
  Students' sentence rewrites are scored using a generic 0- to 1-point rubric. Human scorers are trained to apply the rubric by using the approved anchor sets and then must successfully score at least two practice sets. Both the anchor and practice sets include annotations explaining why and how the score was assigned.

- o **Training for Constructed Responses**
  Students' constructed responses for grades 2 and 3 are scored using a 4-point rubric. Constructed responses for grades 4–12 are scored using a generic 12-point rubric that identifies scores based on three traits—vocabulary, usage, and completeness. Human scorers are trained to apply the rubric by using the approved anchor sets and then must successfully score at least two practice sets. Both the anchor and practice sets include annotations explaining why and how the score was assigned.

Rubrics and sample scoring guides are available on the TELPAS Resources webpage. In addition, the scoring guide used for some speaking questions and each constructed-response question on a TELPAS released test is available on the TELPAS Released Test Questions webpage.

- **Training the ASE**
  The ASE is trained on student responses and human scores from the field-test data. For speaking responses, the ASE is trained to emulate how humans would score student responses for each speaking question based on the content and linguistic features of the responses. It is trained to accurately recognize students' speech based on an ample number of spoken responses produced by EB students with a variety of accents who are representative of the demographics of TELPAS test takers. For writing responses, the ASE is trained to emulate how humans would score student responses for each question.
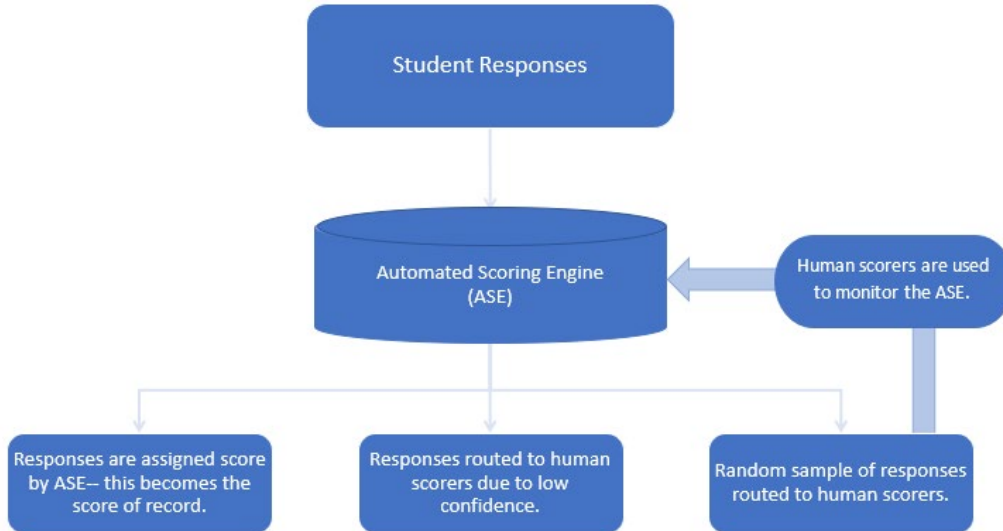
TEA evaluates the performance of the ASE for each question using a subset of student responses and corresponding scores that the ASE has not processed. Performance is examined relative to how humans would score the responses to each question; TEA requires the ASE to agree with human scorers at the same rate human scorers agree with one another and that the distribution of ASE scores is similar to the distribution of human scores. As part of the training process, the ASE calculates confidence values that indicate the degree to which the ASE is confident the score it has assigned matches the score a human would assign.

Scoring Process for TELPAS Speaking and Writing Responses

Step 3: Scoring

The process for the hybrid scoring model is outlined below.

- All grade 4-12 student responses are scored for the first time by the ASE.
- At least 25 percent of student responses for each grade and language domain are then routed to trained human scorers. This set of second scores by human scorers is used to monitor the ASE scoring and is similar to what is done in human scoring to examine how often scorers agree with one another.

TELPAS Writing Constructed Responses and Sentence Rewrite (4-12)	TELPAS Speaking Responses
- A random sample of student responses are routed to hand scoring and used to the monitor the ASE, while providing the score of record. - Student responses that the ASE identified as "low confidence" are routed to trained human scorers. The low confidence responses are often those responses that are on the border between two score points. The purpose of this routing is to ensure that unusual or borderline responses receive fair and accurate scores. - Student responses that the ASE assigns a "condition code" for are routed to the ASE for verification. "Condition Codes" for TELPAS Writing include responses that are	- During the scoring window, checks are conducted to ensure that the ASE is scoring reliably and begins the reliability and validity check of at least 15 percent of the responses. - Responses that are "not scorable" are identified and assigned to scorers. These are responses the system does not have confidence in assigning a score. The "not scorable" responses may have loud background noise (e.g., school bell rings, static sound in recording), mumbled or faint speech, the presence of a language other than English, the volume of the recorded response is too low and difficult to score or are responses that did not have

written in a language other than English or are identified as needing score verification.	similar enough representation in the materials used to train the scoring engine.

Step 4: Reporting

Results from the TELPAS administration are sent to districts after all testing is complete. Reports for each student include the total number of points achieved, the scale score, and the proficiency rating for each of the four language domains. The reports also include the student's composite rating—his or her overall level of English proficiency—for the current and the previous school year. For more information about reporting of TELPAS results, refer to the [Interpreting Results](#) page in the *District and Campus Coordinator Resources*.

Step 5: Rescores

If district personnel or a parent or guardian has concerns about a student's score on his or her responses to these open-ended questions, district testing personnel can request that the student's responses be rescored for a fee. Rescore requests for speaking assessments are separate from rescore requests for writing. All rescore requests are sent for human scoring. The results of the rescore are provided to district testing personnel. If the score on the open-ended questions changes, the fee is waived.

The rescore request must be submitted in the Test Information Distribution Engine (TIDE) by district testing personnel within a specific window. Refer to the [Calendar of Events](#) for specific dates. For assistance in submitting rescore requests in TIDE, contact Texas Testing Support at (833) 601-8821 or TexasTestingSupport@cambiumassessment.com.