

Test takers receive a score (Benchmark Level) for each section of the test. The Reading and Listening scores are scored automatically by the computer. The Writing and Speaking items are scored by Avant's Valid-certified raters who use a Scoring Rubric (a scoring tool that lists the criteria for meeting Benchmark Levels).

The rubric outlines in detail what expectations are placed upon the test takers for performance at any given level. Test takers who hope to score at a specific level know the characteristics of the work they must produce to achieve that level.

## Benchmark Scale

STAMP 4S reports assessment results using the CASLS (Center for Applied Second Language Studies) Benchmark Scale.

Scoring is done using Benchmark Levels of **Beginning**, **Transitioning** and **Expanding**. Within each Benchmark Level, there are sub-level designations of A, B and C. These designations identify test takers in the top third, middle third or bottom third of the range score for that particular level. Similar to ACTFL's low, mid and high designations, these designations will assist in seeing a further breakdown of the test taker's ability. While they are similar to ACTFL's levels, there is not a direct correlation (see Level Key below). Because it takes a great deal of time and practice for students to acquire the skills necessary to move from the Beginning Benchmark to the Transitioning Benchmark, teachers will be able to track student growth within the sublevels of A, B and C.

This Benchmark Scale is comparable to other well-known proficiency scales and relates to the ACTFL scale as shown below.

### Level Key

The letters A B C relate to the ACTFL scale in the following manner:

<b>Beginning</b>	<b>Transitioning</b>	<b>Expanding</b>
A-Novice-Low	A-Intermediate-Low	A-Advanced-Low
B-Novice-Mid	B-Intermediate-Mid	B-Advanced-Mid
C-Novice-High	C-Pre-Advanced	C-Advanced-High

[Score Interpretation Guide \(PDF\)](#)

## Reading and Listening Benchmarks

Each Reading and Listening test item has an associated level. Because the Reading and Listening sections are computer-adaptive, the test taker experiences questions at various levels. The algorithm used to determine the test taker's scores for Reading and Listening looks at the number of questions that were answered correctly at each test item level.

The chart on the next page shows topics typically associated with the Benchmark Levels for Reading and Listening test items. It also shows the Reading and Listening characteristics associated with the levels.

Benchmark Level	Topics	Characteristics
<b>Beginning</b>	Self Calendar/Time Colors/Shapes Pets/Animals School/Classroom Weather/Seasons Clothing Food/Beverage Family/Friends Home Places/Geography Community Daily Routines Shopping/Stores Leisure/Activities	Students who are reading or listening at Beginning proficiency are characterized by: <ul style="list-style-type: none"> <li>• Reliance of learned phrases and basic vocabulary</li> <li>• Ability to recognize the purpose of basic texts</li> <li>• Can understand a core of simple, formulaic utterances</li> </ul>
<b>Transitioning</b>	Health Holidays/Celebrations Occupations/Professions Transportation Travel/Vacations Future plans Culture Contemporary Issues <ul style="list-style-type: none"> <li>• Current events</li> <li>• Economics</li> <li>• Culture</li> <li>• Literature</li> <li>• Science</li> <li>• Social Science</li> <li>• History</li> </ul> <i>Plus more in-depth aspects of Beginning topics</i>	Students who are reading or listening at Transitioning proficiency are characterized by: <ul style="list-style-type: none"> <li>• In reading, ability to understand the main ideas and explicit details in everyday language</li> <li>• Ability to use language knowledge to understand information in everyday materials</li> <li>• Can follow short conversations and announcements on common topics and answer questions about the main idea and explicitly stated details</li> </ul>
<b>Expanding</b>	Arts Politics Religion Math <i>Plus more in-depth aspects of Beginning and Transitioning topics</i>	Students who are reading or listening at Expanding proficiency are characterized by: <ul style="list-style-type: none"> <li>• Can understand and use language for straightforward informational purposes</li> <li>• Can understand the content of most factual, non-specialized materials intended for a general audience</li> <li>• Can understand the content of most spoken factual, non-specialized language</li> </ul>

## The STAMP 4S Scoring Process for Writing and Speaking Items

Scoring for the STAMP 4S Writing and Speaking responses is done using the following three-step process and associated rubric.

- Step 1** Overall evaluation of the response to assess for Ratability — is the response on topic and appropriate for the given prompt/task? If the response is off topic, contains any profanity or includes any menacing or threatening language, the response will be treated as unresponsive or non-ratable and the test taker will receive a “0” score for that task.
- Step 2** Evaluation of each ratable response to determine a Benchmark Level based upon the Text-Type characteristics, followed by rating of Language Control factors (see below).
- Step 3** The system will then post a final rating for each of the tasks and an overall rating based upon an average of all individual task scores.

**Note:** 20% of all test taker scores are rated by a second rater to assist in tracking rater accuracy and reliability.

The same rubric is used for all speaking and writing items. Once rated, the average ratings across all Speaking and Writing items will appear on the summary report page and scores for each sample will appear on the detailed score page.

## STAMP 4S Scoring Rubric

Benchmark Level/ Sub-level	Text Type Characteristics	Language Control
<b>Beginning A</b>	<b>WORDS</b> – Produces words in target language with no connection. Does not have the extended vocabulary or the ability necessary to formulate even simple phrases to address the prompt.	Limited language control. Inability to create more than individual words. L1 (first language) influence may be strong. Errors are expected at this level, but the student must be able to produce at least two comprehensible words.
<b>Beginning B</b>	<b>PHRASES</b> – Language production is beyond simple words, but clearly shows the lack of ability to construct complete sentences.	May make frequent errors, but usually comprehensible to a sympathetic reader/listener. L1 (first language) influence may be present.
<b>Beginning C</b>	<b>SIMPLE SENTENCES</b> – Short, common expressions or memorized statements that may be combined together. Able to create at least two complete simple sentences.	Good accuracy for high frequency expressions. Usually comprehensible to a sympathetic reader/listener. Grammatical (syntax, spelling, conjugation) errors are expected at this level but sentences must make sense to be acceptable.

<b>Transitioning A</b>	<b>STRINGS OF SENTENCES</b> – Variety of sentences that utilize different verbs to create independent thoughts, mostly composed of a recombination of learned simple sentences with some added detail.	Good accuracy with formulaic sentences with some added detail. Errors may occur as student attempts higher level skills. Good Language Control is expected with the majority of the response.
<b>Transitioning B/C</b>	<b>CONNECTED</b> – Logical organization of ideas and flow of sentences or statements. The lower end must have at least 3 unique and non-formulaic sentences while the upper end is moving toward connected sentences with complexity, syntactical control and limited errors.	Good accuracy evident with possible errors that don't affect the overall meaning. Delivery may be somewhat choppy. May have repetitive use of concrete vocabulary with occasional use of expanding terms. Accuracy for complex sentences is emerging.
<b>Expanding A</b>	<b>PARAGRAPH</b> – A connectedness and a flow within the text that includes linking and transitional words and/or phrases. Demonstrates a connection of thoughts that create a coherent and extended discourse.	Language is error free a majority of the time with familiar topics. If errors exist, they are patterned and do not hinder overall meaning. Delivery is mostly fluent with only occasional hesitancy. Some abstract and precise use of vocabulary and terms with familiar topics is evident.
<b>Expanding B</b>	<b>EXTENDED PARAGRAPH</b> – Variety of cohesive devices and organizational patterns evident in response. Vocabulary is clear, specific and natural. Language is smooth and natural in delivery and without noticeable errors.	Language is fluent with limited errors. Ability to create complex language using precise and defined vocabulary. Control of the abstract as well as ease of use of idiomatic phrases and concepts. Clear sequential ordering evident (if required) and accurately follows target language conventions.

The resulting determination of the test taker's Benchmark Levels of Proficiency in Reading, Writing, Listening and Speaking are then made available on STAMP 4S Teacher Report Pages (see STAMP 4S Reporting Guide for more information).