

Versant™ English Placement Test

Test Description

The Versant English Placement test is a 50-minute test designed to evaluate how well a person can understand English on everyday topics. Specifically, the test measures facility in spoken and written English – that is, how well a person can understand spoken and written English and respond appropriately in speaking and writing on everyday topics, at a native-like pace and in intelligible English.

The Versant English Placement test is intended for non-native speakers of English 16 years of age and older, and scores may be used for decisions such as placement testing, exit testing, and progress monitoring. The score report is comprised of an Overall score and four skill scores: Speaking, Listening, Reading, and Writing. Together, these scores describe the candidate's facility in spoken and written English.

Test Administration

The Versant English Placement test is administered via computer and can be delivered in both online and off-line mode. Once the computer has downloaded and installed Versant's test delivery platform (CDT), the Versant test can be taken at any time, from any location. Due to the automated administration, a human examiner is not required. However, depending on the test score use, a proctor may be necessary to verify the candidate's identity and/or ensure that the test is taken under exam conditions. To simulate test conditions, the test platform locks down the computer to prevent web-browsing, copying or pasting.

During test administration, an examiner's voice guides the candidate through the test, explains the tasks and gives examples. The candidate listens through a headset or speakers, and also sees instructions and examples on a screen. Test questions are responded to by speaking into a microphone or typing using the keyboard. The candidate's responses are sent to a remote server, from where the Versant testing system automatically analyzes them and posts scores to a secure website, usually within minutes of completing the test.

The Versant English Placement test consists of nine different tasks: Read Aloud, Repeats, Sentence Builds, Conversations, Typing, Sentence Completion, Dictation, Passage Reconstruction, and Summary and Opinion. Task descriptions and examples of each task are summarized in Table 1.

Table 1. Task Descriptions and examples of each task in the Versant English Placement Test.

Task	Task Description	Examples
Part A: Read Aloud	Read a passage aloud for 30 seconds. Candidates may not be able to finish reading the entire passage in 30 seconds, but this is not counted against them.	Australia is a very large country. It is the sixth largest country in the world. It is also a continent and is sometimes called the 'island continent'. It is surrounded by two oceans. Most of Australia is a desert so it is very flat and dry, but it also has rain forests and mountains. It is home to many different kinds of animals.
Part B: Repeats	Listen to a sentence and repeat the sentence word-for-word. The sentences are presented in approximate order of increasing difficulty.	<ol style="list-style-type: none"> 1. He's a great teacher. 2. It's not too late to change your mind. 3. People know how easy it is to get lost in thought.
Part C: Sentence Builds	Listen to a group of three short phrases presented in random order and then rearrange the phrases to make a sentence.	<ol style="list-style-type: none"> 1. my boss / to London / moved 2. of your family / any pictures / do you have 3. to their leader / listened carefully / the young men
Part D: Conversations	Listen to a conversation between two speakers and then answer a comprehension question.	<p>Speaker 1: Congratulations on graduating!</p> <p>Speaker 2: Thanks! It was a lot of work.</p> <p>Speaker 1: I know. You deserve a party.</p> <p>Question: Why does the man deserve a party?</p>
Part E: Typing	Type a given passage exactly as displayed in 60 seconds. This is an assessment of typing speed and accuracy and allows candidates to familiarize themselves with the keyboard.	Many people do not like public speaking. They are afraid to speak in front of a large group of people. There are many ways to get better at public speaking. First, it is good to know the room. You should know where to stand and where to set up your computer...(etc)
Part F: Sentence Completion	Read a sentence that has a word missing, and then supply an appropriate word to complete the sentence. Candidates are given 25 seconds for each item.	<ol style="list-style-type: none"> 1. Her favorite hobby is _____. She has so many books. 2. He arrives _____ and is often the first one here. 3. I asked a coworker to take over my _____ because I wasn't feeling well.
Part G: Dictation	Listen to a sentence and then type the sentence exactly as it is heard. Candidates are given 25 seconds for each item.	<ol style="list-style-type: none"> 1. I'll see you on Thursday. 2. How long can I keep this book? 3. She apologized to all her friends several times.

<p>Part H: Passage Reconstruction</p>	<p>Read a short paragraph for 30 seconds. After 30 seconds, the paragraph disappears. Then, reconstruct the paragraph in 90 seconds, including the main points and as many details as possible.</p>	<p>Robert went to a nice restaurant for dinner. When the waiter brought the bill, Robert reached for his wallet, but it wasn't in his pocket. He remembered having his wallet when he came into the restaurant. The waiter looked around the floor near his table. He found the wallet under the table.</p>
<p>Part I: Summary and Opinion</p>	<p>Read a passage. Then, write a summary of the author's opinion in 25-50 words and give your own opinion on the topic presented in the passage in at least 50 words. Candidates are given 18 minutes to read the passage and write both responses.</p>	<p>Some children grow up in a big city while other children grow up in the countryside. Childhood experiences can be very different depending on where a person is raised. Although the countryside can be more peaceful than a big city, it is better for children to grow up in a big city... (etc)</p>

For each task, the Versant testing system presents a series of prompts in English and elicits spoken or written responses in English. Prompts that involve listening input are presented in various native and non-native voices that are distinct from the examiner’s voice.

In order to ensure that each skill score is reliable and generalizable, a multi-method multi-trait approach is taken. This means that, for example, the Speaking subscore is derived from performance on a number of different tasks. Figure 1 illustrates which sections of the test contribute to each of the four subscores. The typing task does not feed into the overall score, but the candidate’s typing speed and accuracy are provided separately on the score report.

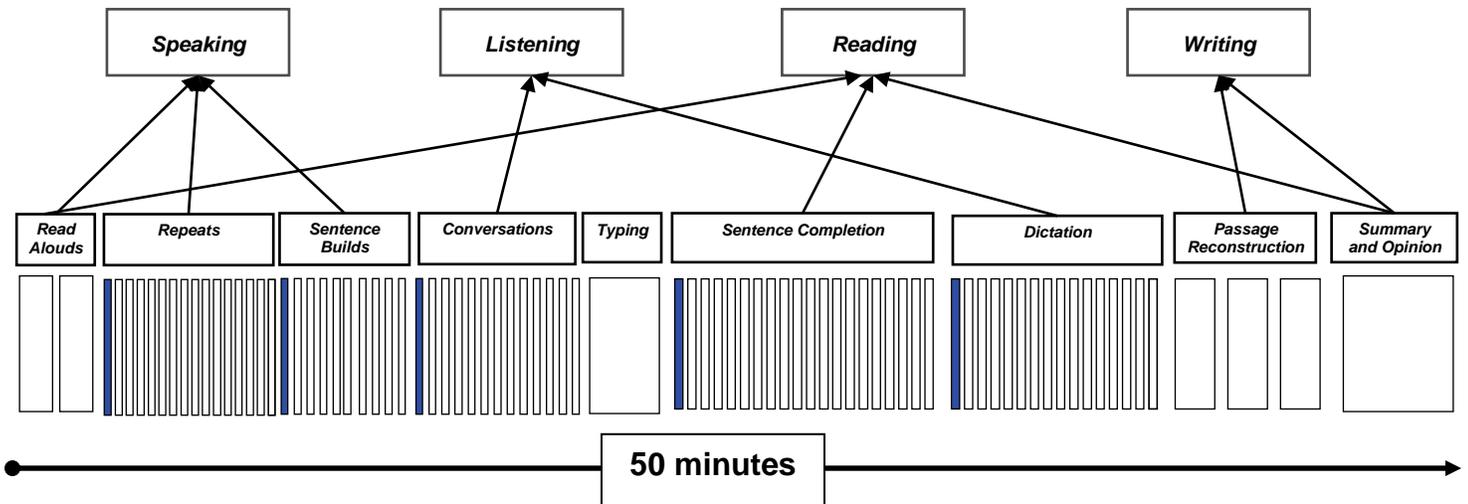


Figure 1. Scoring logic illustrating how tasks feed into subscores.

Validation Evidence for the Versant English Placement Test

A common question about any test is “how do we know this is a good, accurate, consistent test?” To answer that question, one needs to look at two factors that are critical to high quality tests:

1. Reliability and precision: Can you expect consistent results from the test? Would a person score the same if they took the test over and over again?
2. Accuracy: Does the automated scoring assign scores the same way an expert human assessor would?

For a test to be high quality, it must be precise, reliable, and valid for score use, as demonstrated by empirical validation investigations. A series of studies have been conducted on the Versant English Placement test as part of ongoing research, and the data show that the Versant English Placement test demonstrates high levels of validity and reliability.

Reliability and Precision

To understand the precision and reliability of the Versant English Placement test Overall scores and the subscores, the following were examined: the Standard Error of Measurement of the Overall score and the split-half reliability. Estimates of reliability such as these are the foundation of any valid test.

The Standard Error of Measurement (SEM) provides an estimate of the confidence around an assigned score. The SEM of the Overall score on the Versant English Placement test is 2.7 points. This is a very narrow SEM for a language proficiency test, and it means that a candidate's *true score* is +/- 2.7 points within their reported score. For example, a candidate who scored 30 on the test would normally score anywhere between 27.3 and 32.7 if that candidate were to take the test over and over again.

The split-half method is another measure of reliability which shows the consistency of test scores. It splits the candidate's responses in half (usually odd and even numbers) and correlates the two sets of scores to provide a measure of structural reliability. The machine scoring reliability of the Versant English Placement test was estimated in a study of 90 non-native candidates. The results are summarized in Table 2.

Table 2. Split-half reliability coefficients using machine scores on the Versant English Placement Test (n=90).

	Machine split-half Reliability
Overall Score	0.96
Speaking	0.89
Listening	0.90
Reading	0.80
Writing	0.85

The Versant English Placement Test is closely aligned with other spoken and written tests in the Versant suite, such as the Versant English Test, Versant Pro – Speaking, and Versant Pro – Writing. These tests all share many elements in common: test specifications; tasks; rating criteria; scoring models; and, because the test items are linked through Item Response Theory (IRT) statistical modeling, the scoring reporting scale is equivalent. Thus, previously established reliability data on the Versant suite can be legitimately applied to the Versant English Placement Test. For example, when split-half reliability was estimated from test responses using careful human transcriptions and human judgments, it was found to be 0.99 (n=143) and 0.99 (n=124) on the Versant English Test and Versant – Pro Speaking respectively. Human-scored split-half reliability for the Versant Pro – Writing was 0.93 (n=124). Due to the closeness of test construct and test design, the Versant English Placement Test is comparable in terms of split-half reliability as estimated from human scores. Reliability evidence on the Versant English Test, Versant Pro – Speaking, and Versant Pro – Writing is available from Pearson by request.

Accuracy

Another important consideration with any machine-graded test is the need to ensure that automatically generated scores are as accurate as those of trained human evaluators. In one study, 29 candidates who took the Versant English Placement test were evaluated by both machine scoring and by expert human raters. This study was performed to answer one of the most important questions: “How closely do the machine scores correspond to those generated by the human judges?” The results are summarized in Table 3.

Table 3. Human-to-Machine Score Correlation Coefficients for the Versant English Placement Test (n=29).

	Machine-to-Human Correlation
Overall Score	0.99
Speaking	0.97
Listening	0.98
Reading	0.95
Writing	0.94

The correlation coefficient between human-generated and machine-generated Overall scores was 0.99. This means that the automated scoring system provides scores that accurately predict scores given by human experts.

Validation of Summary and Opinion Task

The Versant English Placement test introduces a new task, Summary and Opinion, to the Versant suite of tests. Because of this, several validation studies were performed on this task in particular to ensure that it reliably and accurately assesses candidates.

In the Summary and Opinion task, candidates are presented with a passage. They are given 18 minutes to read the passage, write a summary of the author’s opinion in 25-50 words, and give their own opinion on the topic presented in the passage in at least 50 words. The passages contain an opinion on an everyday topic. All passages consist of an introduction, two body paragraphs, and a conclusion. All passages are relatively simple in structure, use vocabulary from the most frequently-occurring 1,200 words in English, and range in length from 275 to 300 words.

In the Summary response, candidates are expected to demonstrate a clear understanding of the author’s opinion and to identify the supporting points without including unnecessary details or repeated information. In order to do so, the candidate must read the passage presented, understand the concepts and details, and evaluate the information presented in order to identify the most important points. Candidates must construct an informative and succinct response with appropriate spelling, punctuation, capitalization, syntax, and grammar. Responses are scored on the quality of the summary and writing conventions.

In the Opinion response, candidates are expected to provide their own opinion on the topic presented and to provide clear and appropriate supporting ideas and/or examples. Candidates must construct an informative response with appropriate spelling, punctuation, capitalization, syntax, and grammar. Responses are scored on the quality of the opinion and writing conventions.

During the field test of the Versant English Placement test, more than 1,453 Summary and Opinion responses were collected from natives and English learners. In order to develop the automated scoring models, responses to Summary and Opinion items were presented to ten educated native English speakers to be judged for quality of the summary, quality of the opinion, and writing conventions. Before the native speakers began rating responses, they received training in how to evaluate responses according to analytical and holistic rating criteria. All raters held a master’s degree in either linguistics or TESOL. The raters logged in to a web-based rating system and evaluated the written responses to Summary and Opinion items. Rating stopped when each item had been judged by at least two raters.

Inter-rater reliability was calculated for each trait rated in the non-native field test data (n=476). Inter-rater reliability estimates were obtained by calculating the Pearson product moment correlation coefficient for a given candidate's scores. These reliability coefficients are displayed in Table 4.

Table 4. Inter-rater reliability coefficients for the traits rated in the Summary and Opinion task (n=476).

Trait	Inter-rater <i>r</i>
Summary	0.91
Opinion	0.94
Conventions	0.87

These high levels of inter-rater reliability demonstrate that candidates can be consistently rated on the given traits based on performances elicited by the Summary and Opinion task.

Human scoring of the 476 non-native candidates was then compared to the machine scoring of the same candidates. Results are summarized in Table 5.

Table 5. Human-to-Machine score correlation coefficients for responses to the Summary and Opinion task (n=476).

	Machine-to-Human Correlation
Summary	0.89
Opinion	0.92
Conventions	0.90

The correlation coefficient between human-generated and machine-generated scores range from 0.89 to 0.92. This means that the automated scoring system provides scores that accurately predict scores given by human experts.

Another common estimate of a task's reliability is the test-retest method, where scores on the first Summary and Opinion item are compared with scores on a second Summary and Opinion item. This reflects the consistency of candidate's performance on the task. The test-retest reliability of the Summary and Opinion task was calculated in a study of 90 candidates who took the task twice, a subset of the 476 non-native candidates. The test-retest correlation coefficient from machine scores is 0.88 and from human scores it is 0.87. The data suggest that there is sufficient information in the Summary and Opinion task to extract reliable estimates of a candidate's performance. The difference between the Versant automated and careful human rating is small.

Versant English Placement Test and Common European Framework of Reference (CEFR) Level Estimates

The CEFR is published by the Council of Europe, and provides a common basis for describing language proficiency using a six-level scale from A1 (lowest), A2, B1, B2, C1, to C2 (highest). Because the Versant English Placement test shares many tasks with the Versant Pro – Speaking test and the Versant Pro – Writing test, the results from previous studies mapping Versant Pro - Speaking and Writing scores onto the CEFR levels has been applied to the Versant English Placement test. That is, the established Versant Pro score ranges aligned with the CEFR levels have been used for the Versant English Placement test, as shown in Table 6. (Note that the Versant English Placement test assesses only up to CEFR level C1).

Table 6. Mapping of CEFR Levels with Versant English Placement Test Scores.

Versant English Placement Test 20 – 70	CEFR A1-C1
20-23	<A1
23-32	A1
33-45	A2
46-55	B1
56-67	B2
68-70	C1

Score Use

Once a candidate has completed a test, the Versant testing system analyzes the spoken and written performances and posts the scores at www.VersantTest.com. Test administrators and score users can then view and print out the test results from a password-protected section of the website.

Score users may be educational and government institutions as well as commercial and business organizations. Pearson endorses the use of Versant test scores for making decisions about the English skills of individuals, provided score users have reliable evidence confirming the identity of the individuals at the time of test administration. Score users may obtain such evidence either by administering the Versant English Placement test themselves under secure conditions or by having trusted third parties administer the test. In several countries, education and commercial institutions provide such services.

Versant English Placement test scores can be used to assess how well and efficiently a candidate can process and produce spoken and written English on everyday topics. Scores may also be used effectively in evaluating whether an individual's level of spoken and written English is sufficient to perform certain tasks or functions requiring mastery of spoken and written English.

The Versant English Placement test score scale covers a wide range of abilities in spoken and written English communication. Score users must decide what Versant English Placement test score can be regarded as a minimum requirement in their context. Score users may wish to base their selection of an appropriate criterion score on their own localized research. Pearson can provide further assistance in establishing criterion scores.

For more information about the Versant English Placement test, please email sales@pearsonkt.com