

Test Validation

Introduction

The Versant[™] Spoken English Test is designed to evaluate the level of oral proficiency of adult learners of English. Independent rating has been conducted in order to determine how the scores relate to the levels as defined in the Descriptors of the "Common European Framework for Languages".

This comparison was based on the interpretation of test results from a wide sample of candidates, and enabled us to identify the level of each test taker on the "European" scale. Descriptors from the framework were used by the independent experts while listening to test takers' performance, and deciding on their level of competence.

The results provide a clear and reliable comparison of the scores from the Versant[™] Spoken English Test and the six levels described in the "Framework for Languages".

Relating Versant Scores to Council of Europe (2001) Descriptors

The data presented in this guide are based on two consecutive research studies. The first study was conducted in 1999 and involved three independent raters rating 120 students on two open questions each (Bernstein, John De Jong, Pisoni, and Townshend, 2000).

In order to cover the full range of potential development as well as a wide range of aspects of conversational skill, a scale of oral interactional proficiency according to the European Framework was constructed. The scale was based on the Overall Spoken Interaction scale, but was rewritten to incorporate elements from the specific Strategic, Pragmatic, and Linguistic scales. The scale was written to contain the six basic levels used in the European Framework: A1, A2, B1, B2, C1, and C2 (See Summary Table 2.). Because the lowest level (A1) assumes some, albeit minimal, proficiency, a zero level was introduced to assign to subjects not reaching this minimal level or providing no evidence of proficiency at all (e.g., silence or avoiding behavior).

Three raters were contracted from three different European countries: the Netherlands, Switzerland, and the UK. All three raters were well acquainted with the Council of Europe Framework, and one rater is regularly involved as a professional rater in examinations of English as a foreign language. The raters were trained using two sets of examples. The first set consisted of video recordings of an oral task performed by subjects at each of the six levels of the European Framework. Thus raters were made aware of the meanings of the levels independent of the rating tasks to be performed in the experiment. The second set consisted of six digitized recordings of subjects performing one of the open-ended tasks (Part E) within the SET-10 test. Samples were selected to represent a range of levels, not all of them different. Raters played these samples using the Windows Media Player and were to rate them using the European Framework-based scale that was developed for the project. After rating them independently, raters discussed their results via e-mail and negotiated on discrepancies, thus acquiring a common feeling for the distinctive elements to be taken into account in assigning European levels to subjects.

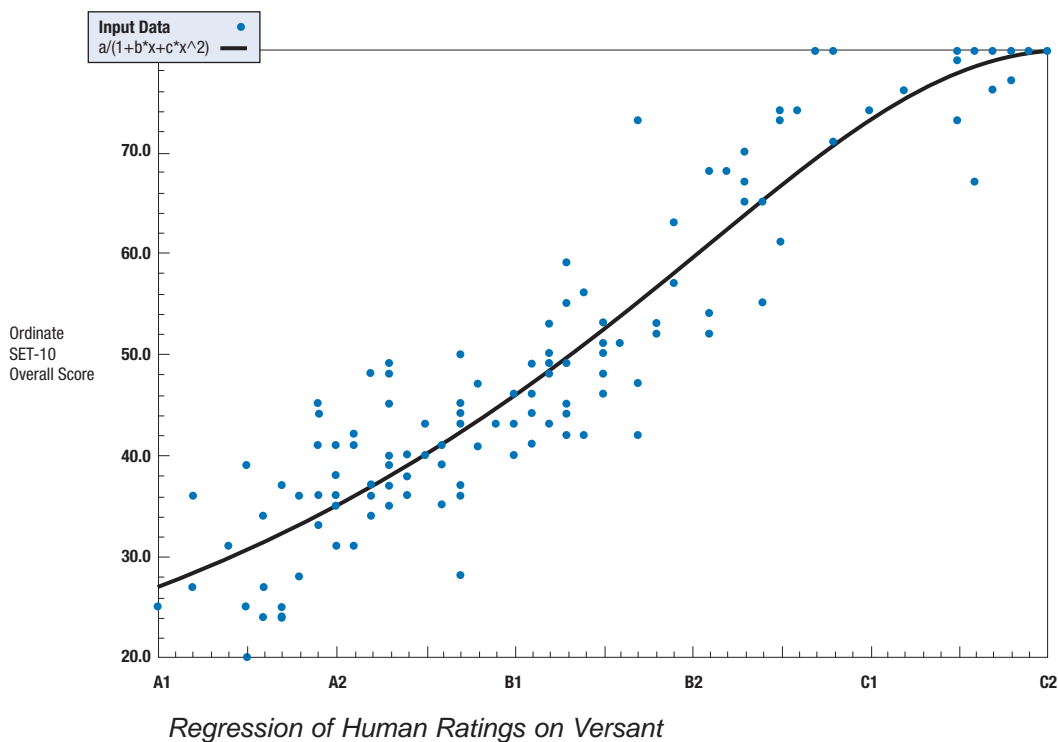
The rating sample for each rater consisted of 121 subjects responding to the open-ended questions. To control for intra-rater consistency, raters were presented with about 30% of the samples twice. The intra-rater reliability was .96 and the interrater reliability was .95. The correlation with the Versant Overall Score was .85.

Based on the ratings produced in the first study, a preliminary score transformation table from the SET-10 scale to the Council of Europe levels was computed. It was felt, however, that basing judgments only on the open-ended questions would most likely result in an underestimation for the test takers' Council of Europe levels, especially at the lower end of the scale. This belief resulted from the fact that (1) test takers, knowing that the

Test Validation

open-ended questions are not counted in the Versant Overall Score, might tend to skip these questions or put less effort into responding to them, and (2) the open-ended questions might seem intrinsically too hard for test takers at the lower levels.

The second research study was conducted in 2002 and involved four raters and 150 subjects. Raters were from the USA and from the Netherlands. In addition to the open-ended questions rated on the Council of Europe Scale, raters also rated test takers on pronunciation and fluency as exhibited in responses to other items in the Versant test. In total, raters produced 11,090 independent ratings. Interrater reliability was .94 and the correlation of ratings with the Versant Overall Score was .94. Ratings with subscores for fluency and pronunciation ranged over raters from .72 to .86.



The results of the second study indeed showed that test takers with Versant scores below 50 had been underestimated, and a new score transformation was computed. This Versant Test Validation Guide is based on this new transformation. The best projection of Versant Overall scores on the Council of Europe levels can be pictured as shown in the figure.

Two summary tables offer a quick reference for interpreting Versant scores in terms of Council of Europe (2000) descriptors. Summary Table 1 presents an overview relating the Council of Europe global scale (Council of Europe, 2001:24) to Versant scores. Summary Table 2 provides the more specific oral interaction scale of descriptors used in the studies mentioned above.

Summary Table 2

Summary Table 1: Aligning general level descriptors and test scores

© Council of Europe, 2001			SET-10
Level	Descriptor		Score
Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.	79
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.	69
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	58
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.	47
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.	36
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.	26
			20

Summary Table 2

Summary Table 2: Relation of Versant Overall Scores to Oral Interaction Descriptors Based on Council of Europe (2001)

SET-10	Council of Europe Levels and Descriptors	
80 79	C2	Conveys finer shades of meaning precisely and naturally. Can express him/herself spontaneously at length with a natural colloquial flow. Consistent grammatical and phonological control of a wide range of complex language, including appropriate use of connectors and other cohesive devices.
78 69	C1	Shows fluent, spontaneous expression in clear, well-structured speech. Can express him/herself fluently and spontaneously, almost effortlessly, with a smooth flow of language. Clear, natural pronunciation. Can vary intonation and stress for emphasis. High degree of accuracy; errors are rare. Controlled use of connectors and cohesive devices.
68 58	B2	Relates information and points of view clearly and without noticeable strain. Can produce stretches of language with a fairly even tempo; few noticeably long pauses. Clear pronunciation and intonation. Does not make errors that cause misunderstanding. Clear, coherent, linked discourse, though there may be some "jumpiness."
57 47	B1	Relates comprehensibly main points he/she wants to make on familiar matters. Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair may be very evident. Pronunciation is intelligible even if a foreign accent is sometimes evident and occasional mispronunciations occur. Reasonably accurate use of main repertoire associated with more predictable situations. Can link discrete, simple elements into a connected sequence.
46 36	A2	Relates basic information on, e.g., work, background, family, free time, etc. Can make him/herself understood in very short utterances, even though pauses, false starts, and reformulation are very evident. Pronunciation is generally clear enough to be understood despite a noticeable foreign accent. Uses some simple structures correctly, but still systematically makes basic mistakes. Can link groups of words with simple connectors like "and," "but," and "because."
35 26	A1	Makes simple statements on personal details and very familiar topics. Can manage very short, isolated, mainly prepackaged utterances. Much pausing to search for expressions to articulate less familiar words. Pronunciation is very foreign.
25 20	<A1	Candidate performs below level defined as A1.

If you would like a copy of the complete "Can-Do" guide, which provides a more detailed explanation of the alignment between individual scores and the "oral interaction" descriptors, please contact us on john.parsons@harcourt-uk.com.