

*Court Interpreter Oral Examination:*

# Test Construction Manual

**National Center for State Courts - Language Access Services  
Section**





Copyright © 2012 by the National Center for State Courts. All rights reserved.

The National Center for State Courts is an independent, nonprofit, tax-exempt organization in accordance with Section 501 (c)(3) of the Internal Revenue Code.

Except as permitted under the Copyright Act of 1976 and as otherwise expressly provided herein, no part of this publication may be reproduced in any form or by any means, electronic or mechanical, including the use of information storage and retrieval systems, without permission in writing from the copyright holder. Permission is hereby granted to NCSC members, Consortium staff and consultants to reproduce and distribute this publication for educational purposes if the copies credit the copyright holder.

For additional information, please contact:

National Center for State Courts  
300 Newport Avenue  
Williamsburg, VA 23185

This document is also available at: [http://www.ncsc.org/Web%20Document%20Library/EC\\_StateInterpCert.aspx](http://www.ncsc.org/Web%20Document%20Library/EC_StateInterpCert.aspx)

Publication Date: April 2010, Revision Date: July 2012

*Court Interpreter Oral Examination:*  
**Test Construction**  
**Manual**

---

***National Center for State Courts - Language  
Access Services Section***

The original version of this document was developed in 1996 under a grant from the State Justice Institute (SJI-95-12A-B164). Points of view and opinions stated in this report are those of the project staff and do not necessarily represent the official position or policies of the State Justice Institute or any member of the project advisory committee.



## TABLE OF CONTENTS

---

<b>Purpose of This Manual</b> .....	1
[Section 1] <b>Overall Design of Tests</b> .....	1
[Section 2] <b>General Management of Test Development Activities</b> .....	3
[Section 3] <b>Compilation of Materials for Test Development and Production of Base Texts</b> ..	4
[Section 4] <b>Qualifications of Test Development Teams</b> .....	6
[Section 5] <b>Test Development Process</b> .....	7
[Section 6] <b>Description of Scoring Units</b> .....	8
Table 1: Scoring Unit Descriptions and Testing Goals.....	9
[Section 7] <b>Selection and Classification of Scoring Units</b> .....	11
[Section 8] <b>Distribution of Scoring Units: Full Three-Section Examination</b> .....	13
Table 2: Scoring Unit Distribution for Standard Model.....	14
Table 3: Scoring Unit Distribution in the Consecutive .....	15
[Section 9] <b>Distribution of Scoring Units: Abbreviated Examination</b> .....	15
Table 4: Scoring Unit Distribution for Abbreviated Model .....	16
[Section 10] <b>Length of Test Sections</b> .....	16
Table 5: Number of Words in Test Sections .....	16
[Section 11] <b>Production of Test Materials for Test Administration</b> .....	17
[Section 12] <b>Production of Materials for Test Rating</b> .....	20
[Section 13] <b>Field Testing New Tests</b> .....	22
[Section 14] <b>Revising Existing Tests</b> .....	23
[Section 15] <b>Documentation of the Test Development Process</b> .....	25
[Section 16] <b>Statistical Analysis of Tests</b> .....	25
[Appendix 1] <b>Characteristics of Scoring Units</b> .....	29
[Appendix 2] <b>Sample Script</b> .....	31
[Appendix 3] <b>Sample Scoring Dictionary</b> .....	33



## Purpose of This Manual

---

This Manual frames the process used by the Language Access Services Section (LASS) of the National Center for State Courts (NCSC) for developing and maintaining valid and reliable oral performance examinations that can be used by state court systems to credential persons who wish to become spoken language court interpreters, regardless of when or where they take the examination. The exams are developed to measure a candidate's ability to faithfully and accurately interpret the range of English ordinarily used in courtrooms into another language, and to understand and interpret into English what is said by a native speaker of another language.

The level of performance the test measures is the minimum acceptable level for entry into the profession of spoken language court interpretation. These exams do not measure other aspects of the knowledge, skills, and abilities one must have to perform the duties of a court interpreter, such as engaging in appropriate forms of situational control, dressing and conducting oneself in a manner consistent with the dignity of the court, and delivering services via telephone or as a member of a team.

This *Test Construction Manual* provides a structure for the oral examination development process to ensure that all tests have similar levels of difficulty, both for all languages and all versions within the same language.

### [Section I]

## Overall Design of Tests

---

Every test is developed according to specific test models and includes a scoring dictionary.

1.1 Test instruments shall include a section for one or more of the modes of court interpretation as follows:

- *Sight Translation*: the mode of interpreting whereby a written document in one language is interpreted into another language. This mode includes two separate components: one single-page document in English to be interpreted into another language and another single-page document in that other language to be interpreted into English. Using a written script that does not identify the scoring units, candidates read and deliver an interpretation of the written test instrument. A separate copy of each script that identifies the scoring units is developed but not shown to the candidate. It is instead used during the rating process.
- *Consecutive Interpretation*: the mode of interpreting performed when interpreting the testimony of a witness or when parties are in direct conversation with the Court

on the record. Accordingly, this mode of interpretation is bidirectional. The test measures interpretation of some utterances from English into another language and other utterances from the other language into English. This section of the test is administered using a prerecording of the script. The written test instrument has the scoring units marked and is not shown to the candidate, but is instead used during the rating process.

- *Simultaneous Interpretation*: the mode of interpreting that renders into another language everything that is said in English during court proceedings while the speaker continues to speak. This section of the test is administered using a prerecording of the script. The written test instrument has the scoring units marked and is not shown to the candidate, but is instead used during the rating process.
- 1.2 Tests shall follow one of the two oral performance test models: The “standard model” is the original model used since the initial development of NCSC court interpreter exams and is sometimes referred to as a “full test” or a “whole test.” It includes all three sections: sight translation, consecutive and simultaneous interpretation.
- The “abbreviated model” was adopted in 2003. This model includes a simultaneous section and a measure of conversational proficiency in English to be chosen from a list of available standardized tests promulgated and maintained by the LASS.

1.3 Tests shall have a scoring dictionary. This is a reference document that the test development team prepares for test raters to use when scoring exams to ensure reliability in the rating process. It sets forth specific examples of how raters will determine whether certain renditions by candidates are to be counted as acceptable or unacceptable and is amended and updated from time to time according to established procedures (see Section 14). The following information will be included for each scoring unit on the test form: its number, classification, the actual word or words in the scoring unit, acceptable renderings, and unacceptable renderings (see sample in Appendix 2).

- 2.1 The LASS is responsible for the development of new and modification of existing oral examinations, and ensures that test forms comply with the standards described in this Manual. LASS staff may serve as the Test Development Manager (hereinafter “TDM”), or contract with a third-party consultant who meets the criteria required to serve in that capacity and who would operate under the general oversight of the LASS.
- 2.2 The TDM must have experience writing court interpreter examinations and possess an understanding of and follow the policies for test development articulated in this Manual.
- 2.3 The TDM is responsible for assembling and supervising the test development team as outlined in Section 4, arranging for the professional review required per Section 3.5, and arranging for the linguistic review required under certain circumstances per Section 5.4. LASS staff should solicit from state court interpreter program coordinators and other experts suggestions for:
- Members of test development teams and consultants hired to perform the reviews mentioned above; and
  - Other forms of assistance, advice, and resources as LASS staff deems appropriate.
- 2.4 The test development team shall conduct a preliminary assessment of the availability of bilingual legal glossaries of English and the corresponding test language. A list of such bilingual resources should be compiled by the team members and given to staff for distribution to state court interpreter program coordinators. If it is determined that there are no credible bilingual legal glossaries for the new test language, the team members will develop one as a component of the test development process.
- 2.5 The process should include a thorough written explanation of any special characteristics of the new test language that may:
- Present challenges for the provision of court interpreting services into or from that language,
  - Require deviation from the test design described herein, or
  - Require special attention during the course of developing the test (e.g., how to treat issues of dialect variation and clarify what dialect or form of the language the test would be written in).

Furthermore, the process should also analyze what the name of the language being tested should be. Ordinarily the names of languages for tests should follow the nomenclature of Ethnologue, but variations may be allowed upon demonstration that there are substantial and persuasive reasons for departing from Ethnologue.

The results of the research and conclusions reached on such language issues should be written up and kept in the official test file for each new language.

[Section 3]

## Compilation of Materials for Test Development and Production of Base Texts

---

- 3.1 LASS staff will compile a collection of transcripts of court proceedings and other documents that may serve as base texts for new exams. State court interpreter program coordinators may submit such materials to staff for addition to this collection at any time.
- 3.2 The base texts which serve as the starting point for developing the two parts of the sight translation section of an examination should be selected, if possible, from documents actually used by a court and originally written in the source language, such as:
- For English to other language sight translation:
    - Police and other law enforcement reports.
    - Investigative reports (e.g., pre-sentence report, child custody report) prepared for a court or a court support office.
  - For other language to English sight translation:
    - Correspondence to a judge and/or other court staff (e.g., character reference letters to a judge).
    - Documents prepared for or by a court to which parties are entitled access (e.g., stipulation, affidavit, court order, police report).

Documents that include text that is more or less established and non-variable should be avoided as they are readily available in the public domain and therefore could compromise test security. Examples include the Miranda warning, pre-printed forms, form letters, standard conditions of probation and other standard documents whose text does not vary.

The final script of the sight translation document in the other language shall be translated to English by the test development team and maintained in the test file.

- 3.3 The base texts which serve as the starting point for developing the consecutive and simultaneous sections of an examination should be selected from transcripts of official court proceedings that involve general legal concepts and procedures, and from types of legal situations that require large volumes of interpreting services. They should also represent a wide range of:

- Stages of legal proceedings (e.g., preliminary hearings, trials, sentencing).
- Types of proceedings and legal subjects heard in state courts (e.g., criminal, family, and civil courts).<sup>1</sup>
- Levels of proceedings (e.g., state, county, and, to a lesser degree, municipal courts).

3.4 The collection of base texts should not include materials that are:

- Unique to a particular state or a small subset of states, except where they can be easily adapted to meet LASS standards.
- Specialized, arcane, or highly technical (e.g., reports or testimony of an expert witness, sophisticated legal argumentation).
- Replete with discourse that contains an inordinate number of legal terms (e.g., motions and legal arguments on technical points of law).
- Shocking in terms of content (e.g., containing graphic testimony in cases such as murder and sexual assault).

3.5 The base texts should be reviewed and edited by LASS staff to ensure that the discourse is clear, consistent and tells a coherent story. These edits may include, but are not limited to, the deletion of repetitive utterances, false starts by speakers, and nonsensical words or phrases. In addition, the base texts may be edited in the test development process as necessary to meet the requirements outlined in Sections 7 through 10.

3.6 Every base text edited by LASS staff should be reviewed by professionals with the appropriate expertise to assess and confirm that the discourse of the text is consistent with acceptable legal practice and procedure.

3.7 The test development team shall read and consider the base texts for content, cultural appropriateness and sensitivity. The TDM, in consultation with LASS staff, shall select from the collection of reviewed base texts the texts to be used by the test development team.

3.8 When choosing a base text to be used for any new exam being written, the TDM shall review the languages for which each base text has already been used, analyze in consultation with the specialists selected for the new exam the likelihood that a candidate taking the test in the new language would also take the test in any of the languages for which any section of that test form is already in use, and offer other base texts for the new test's development. For example, the same base test text should not be used for two languages in the same family of languages (e.g., Romance or Slavic languages), two languages that are native languages within the same country and

---

<sup>1</sup> Over time, the LASS will increase the use of civil and family transcripts while maintaining a significant use of criminal transcripts.

spoken by large numbers of persons (e.g., Cantonese and Mandarin, Gujarati and Hindi, Haitian Creole and French), or two languages for which there are large numbers of bilingual speakers (e.g., Arabic and French). The TDM shall maintain a list that identifies all such potential conflicts and update it from time to time.

- 3.9 Before being given to a test development team, a base text should be edited to conform with the parameters for test construction as follows:
- The number of words (see Table 3) for each section and, in the sight, each part;
  - In the consecutive, the range of utterances in each source language (see Table 5) in which scoring units will be embedded; and
  - Linguistic phenomena constituting potential scoring units for all ten categories as distributed in each test section and, for the sight section, in each part (see Table 2).

## [Section 4]

## Qualifications of Test Development Teams

---

- 4.1 Under the direct oversight of the TDM, each test should be written by a team consisting of at least two specialists as described below, one of whom must be a native speaker of the non-English language:
- One applied linguist, preferably a practicing professional interpreter who possesses the highest credentials available in the field, as confirmed by the TDM<sup>2</sup>; and
  - One theoretical, scholarly linguist who has the most formal academic training possible in the linguistics of the language or, if such an individual is not available, the literature of the language, or, when no such trained linguist is available, a highly educated bilingual professional such as an attorney or teacher who is a native speaker of the language. Any such specialist selected to serve on the test development team will be trained to recognize the demanding specifics of court interpreting (e.g., maintaining the register of the source language; no editing, paraphrasing or embellishing).
- 4.2 When it is not possible to assemble a team of professionals as described above, the team should consist of persons who come as close as possible to possessing these skills and credentials and who are deemed to be the most competent persons available for the purpose of writing a test.

---

<sup>2</sup> For example, for Spanish, the highest credential available is certification by the Administrative Office of the U.S. Courts or reaching a score on a NCSC Spanish examination of 80 percent or more; for some other languages, the highest credential available is reaching a score on the NCSC examination of 80 percent or more. In the case of a language for which no examination exists, there may be no credential available in the field. The TDM will confirm what the “highest credential available in the field” is, depending upon the language for which the exam is being developed.

- 4.3 If, prior to the development of an examination in a new language, there is no avenue through which a member of the test development team can become credentialed (certified, licensed, approved, etc.) and that member wishes to be so credentialed, he or she may notify the TDM of his or her desire to obtain the credential prior to commencing the test development process. Once LASS staff and the TDM have decided which simultaneous base script will be used, the TDM will administer the simultaneous section of the test and hold the recording until the new test has been completed.

Once the test is developed, the TDM will have the second member of the test development team rate the exam under his or her supervision. Any member of a test development team who passes the simultaneous test should be considered as having passed the test as if it had been taken under the usual circumstances and the TDM will recommend to that team member's home state that these test results be accepted for credentialing purposes.<sup>3</sup>

## [Section 5]

## Test Development Process

---

- 5.1 The TDM shall develop a comprehensive training program that all new members of test development teams will complete that consists of the following:
- The TDM shall review the concept of scoring units as described in Section 6 with the test development team, and ensure that the team members have an understanding of how the examination, when completed, will be administered and scored.
  - The TDM shall review this Manual and the *Test Writing Handbook* with team members prior to beginning the writing process.
- 5.2 The TDM shall guide the test development team in the production of scoring dictionaries.
- 5.3 When the test development team produces translations of any English script into another language, an independent linguistic review should be completed to ensure that the translated script is dialect-free and does not contain regional varieties that might not be recognized by all speakers of the language. The TDM shall explain to the test development team that they should produce a document that is a true dynamic equivalent of the English text and does not sound like a translation. The story, concept, or substance of the original text should be reproduced in the other language in a natural way that examinees who are native speakers of that language will find to be authentic speech or text that a native speaker might say or write.

---

<sup>3</sup> This provision was adopted in view of the fact that, for many languages in which the NCSC will develop tests in the future, there is likely to be a very small pool of experts from which to select test developers and raters. It is essential to offer this option since, in its absence, prospective test developers and raters may decline to provide their services.

- 5.4 When developing the consecutive section of the test, the TDM must understand and review with the test development team the meaning of the utterances that are in the non-English language to ensure that the substance of the questions and answers make sense and flow in a logical way. Team members shall provide the TDM with a final English version of the witness' answers.
- 5.5 A complete set of the test scripts developed per Section 11 shall be retained for each test form identifying the number of words for each paragraph in the sight translation and simultaneous sections and for each utterance in the consecutive section.

[Section 6]

## Description of Scoring Units

---

- 6.1 Scoring units represent objective characteristics of language that the interpreter must understand and render appropriately during the interpretation. Each scoring unit is a word or phrase that captures a logically complete linguistic unit.<sup>4</sup>

Tests are scored on the basis of scoring units, which are specific linguistic phenomena that interpreters must be able to deliver for a complete and accurate interpretation. A scoring unit is a pre-selected word or set of words in the exam material that is clearly identified by bolding and underlining the word or words comprising the scoring unit and then identifying the scoring unit in bolded superscript with a number and letter identifying the category of scoring unit to which it has been assigned. The following table lists the ten categories of scoring units and their corresponding testing goal(s).

---

<sup>4</sup> William E. Hewitt, *Court Interpretation: Model Guides for Policy and Practice in the State Courts*, National Center for State Courts, 1995, 104.

Table 1: Scoring Unit Descriptions and Testing Goals

SCORING UNIT CATEGORY	DESCRIPTION	TESTING GOAL(S)
A: Grammar	“Grammar is a system of principles that govern the way a language works. Grammar describes how words relate to each other, particularly how they function in sentences.” <sup>5</sup>	Measure recognition of and, within the limits of the source and target languages, satisfactory handling of the interpretation of grammar, especially verbs.
B: Language Interference	Terms or phrases that may invite misinterpretation due to interference of one language on another (e.g., false cognates, awkward phrasing, terms or phrases susceptible to literal renditions resulting in loss of precise meaning).	1-Measure the ability to keep languages separate, speaking them as an educated native speaker would, with no interference from the other language, and 2-Measure the ability to avoid being constricted unnecessarily by the source language resulting in interpretations that are literal or verbatim.
C: General Vocabulary	Vocabulary that is widely used in ordinary parlance and could be spoken by native speakers appearing in any courtroom.	1-Measure the ability to preserve lexical content of general source language terms when interpreted into the target language, 2-Measure the depth and range of vocabulary, and 3-Measure the ability to tap into a deep reservoir of vocabulary without hesitating or stumbling.
D: Legal Terms and Phrases	Any word or phrase of a legal or technical nature, or which is not common in everyday speech, but is commonly used in legal settings.	Measure range of knowledge and recognition of common legal terms and styles of language used in courtrooms and the ability to faithfully interpret them into the target language, going into both languages, but especially from English into the other language.

<sup>5</sup> DiYanni, Robert, and Pat C. Hoy II. *The Scribner Handbook for Writers*. Boston: Allyn and Bacon, 1995, p. 221.

SCORING UNIT CATEGORY	DESCRIPTION	TESTING GOAL(S)
E: Idioms and Sayings	An idiom is “a speech form or an expression of a given language that is peculiar to itself grammatically or cannot be understood from the individual meanings of its elements” <sup>6</sup> . Sayings are short expressions such as aphorisms and proverbs that are often repeated and familiar setting forth wisdom and truth.	Measure breadth of knowledge and understanding of a language’s common idioms and sayings, and the ability to interpret the meaning or an equivalent idiom or saying in the target language.
F: Register	Style of language drawn upon in various social settings; a key element in expressing degrees of formality, including curses, profanity, and taboo words. Register shows, through a pattern of vocabulary and grammar, what a speaker or writer is doing with language at a given moment.	Assess ability to preserve the level of language so that others’ impression of the speaker is not raised or lowered by the interpreter, and assess ability to interpret offensive terminology.
G: Numbers and Names	Any number, measurement, or proper name.	Measure ability to be precise and accurate with all numbers, maintain weights and measures as stated in the source language without converting them to another system (e.g., from the metric system to United States customary units), preserve names of businesses, streets, etc., without interpreting them (except that “Avenue,” “Street,” etc., may or may not be interpreted, but the actual name is not to be interpreted), and conserve every letter of a spelled name in the order uttered.
H: Markers, Intensifiers, Emphasis and Precision	Any word or phrase giving emphasis or precision to a description (e.g., adverbs, adjectives) or statement (e.g., can be grammatical in form), including time (e.g., the day after tomorrow, last night, next week).	Ensure that the various ways of marking speech are preserved so the same degree of impact and precision is conveyed to the listener of the interpretation.
I: Embeddings and Position	Words or phrases that may be omitted due to position (at the beginning, middle, or end of a long sentence or in a string of adjectives or adverbs) or function (tag questions).	Ensure that all elements of the source language are preserved, especially those that candidates may deem to be “unimportant” or forget due to their location or function in the utterance.
J: Slang and Colloquialisms	Slang and colloquialisms are informal, nonstandard words or phrases that are used in informal, ordinary conversation but not in formal speech or writing and are identified in standard dictionaries as “slang,” “colloquialism,” or “informal” or are listed in published dictionaries of slang and/or colloquialisms or in scholarly articles and books so identifying them. Slang items, which are coined by social groups, may be used in scripts only when they have passed into widespread usage.	Measure the range of knowledge of nonstandard, informal forms of speech and the ability to interpret the meaning of such words and phrases without being bound to preserve their low register.

<sup>6</sup> *The American Heritage College Dictionary*, Third Edition, Boston: Houghton Mifflin Co., 1997, p. 674.

The TDM will guide the test development team in understanding and abiding by the following principles of assigning scoring units during the course of developing a new test, when reviewing existing tests during the field testing stage, or as part of ongoing assessments of existing tests. The TDM and the team members should also closely examine and use the *Test Writing Handbook* as a guide.

- 7.1 The test development team, with oversight of the TDM, should select scoring units using the grids provided in the *Test Writing Handbook* as an aid. Specific types of scoring units should follow, as much as possible, the systematic structure provided therein.
- 7.2 A scoring unit may be any word or set of words in the base text regardless of their perceived importance in the discourse. It does not have to be a critical or important part of the discourse.
- 7.3 The discourse in which a scoring unit is embedded shall provide a clear context so the meaning of the scoring unit is unambiguous.
- 7.4 The ways the scoring unit can be correctly rendered in the target language should be clear and unambiguous.
- 7.5 The test development team should remember that there is often, if not usually, more than one right way to interpret a word or phrase and there are always many ways to incorrectly interpret a word or phrase.
- 7.6 Words and phrases being considered for scoring units in English source materials should be used in contemporary parlance by a wide range of native English speakers in the United States and its possessions, including lawyers, judges, and witnesses, and may not be obscure, archaic, outdated, or unique to a particular region, dialect, or other subset of speakers.
- 7.7 Words and phrases being considered for scoring units in a source language other than English should be used in contemporary parlance by a wide range of its native speakers in all countries where the language is spoken, including witnesses and parties, and may not be obscure, archaic, outdated, or unique to a particular country, region within a country, or other subset of speakers of that language.
- 7.8 Loan words and other types of interference in a mother tongue that arise due to influence of the English language should not be included in the test materials, but if they are, they should never be selected as scoring units.
- 7.9 The same or nearly same text should not be used as a scoring unit more than once in an examination. In addition, a particular verb form (i.e., tense, mood) should appear as a grammatical scoring unit no more than once in a test section.

- 7.10 Whereas many words or phrases can readily be classified in more than one of the categories of scoring units, any words or phrases that are arguably to be classified in the categories of *Legal Terms and Phrases* and *Numbers and Names* may not be selected as scoring units for any other categories. For example, a word or phrase that is arguably a legal term or phrase cannot be classified as a General Vocabulary scoring unit.
- 7.11 The number of legal terms that are not scoring units should not exceed the number of legal terms that are scoring units in any section of the test.
- 7.12 Care should be taken to distribute scoring units so they appear throughout the test as described in Sections 8 and 9 and have adequate space between them to facilitate the rating process and to ensure that they do not appear in clusters at the beginning or at the end of a test section. Scoring units may be adjacent to one another when there is sufficient spacing between the pair of scoring units and the scoring units preceding and following them, but should otherwise be avoided.
- 7.13 A scoring unit should be clearly associated with one of the ten categories of scoring units defined in Section 6. Any word or phrase in a test script could be selected as a scoring unit and many could likely be classified in more than one of the categories of scoring units. When a word or phrase is under consideration to be selected as a scoring unit, and the test development team is not sure how to classify the item, it should consider the guidance provided in the *Test Writing Handbook*. If there is still confusion or disagreement, the members should either avoid making the item a scoring unit or make a clear decision as to how to classify it and note the reasons why it has been classified as such in “Notes” column of the scoring dictionary.
- 7.14 A scoring unit should consist of as few words as linguistically appropriate, rarely exceeding four words. The test development team should avoid phrases that include enough words to confuse or diffuse the testing goal as outlined in Table 1 or that can be reasonably viewed to include two different types of scoring units.

For example, “*the tall gentleman,*” which includes very few words and might be intended as a single scoring unit, is not an acceptable scoring unit because it includes multiple types of scoring units, among them the following examples:

- *Grammar* to measure article/noun or adjective/noun agreement of number and/or gender;
- *Markers and intensifiers* to assess accuracy of interpretation of the adjective “tall;”
- *General vocabulary* to test range of vocabulary with respect to “gentleman” and perhaps even “tall;”
- *Register* to ensure that “gentleman” is not lowered in register with renditions such as “man,” “guy,” “fellow,” or another similar term.

7.15 Suggested steps for identifying scoring units:

**Step One:** Read through a test script and, following the *Test Writing Handbook*, preliminarily assign scoring units.

**Step Two:** Analyze all the scoring units that have been provisionally selected and evaluate them against the principles for selecting scoring units identified in the preceding subsections of Section 7, as well as Appendix 1.

**Step Three:** Using the parameters set forth in Section 8 (for a full 3-section examination) or in Section 9 (for an abbreviated examination), calculate the types of scoring units that are generated in step one. Identify those which are over-represented and those which are under-represented in each test section.

**Step Four:** For those that are over-represented, select the units that appear best to keep (based on perceived strength of the scoring units themselves as well as on their proximity to other scoring units) and eliminate as many as may be necessary to approximate the target number for each type.

**Step Five:** For those that are under-represented, it may be necessary to edit the script in such a fashion as to add material that includes these types of scoring units. The test development team should be careful when adding text to keep the total number of words within the limits allowed in Table 5.

[Section 8]

## Distribution of Scoring Units: Full Three-Section Examination

---

- 8.1 When a full test is being developed, the scoring units should be distributed among the test sections so that each test follows the presumptive distribution outlined in Table 2. The test development team should review the appropriateness of these percentages for each language combination and recommend adjustments to them when the linguistic properties of English and the other language make an adjustment necessary. It is acceptable for the TDM to approve deviation from the target indicated in the “Target %” column in the table below by plus or minus 10 percent, but if the deviation is significant, the TDM should consult with LASS staff for input and direction.

Table 2: Scoring Unit Distribution for Standard Model

SCORING UNIT CATEGORY	SIGHT - ENGLISH TO OTHER LANGUAGE	SIGHT - OTHER LANGUAGE TO ENGLISH	CONSECUTIVE	SIMULTANEOUS	UNIT TOTAL	TARGET %
A: Grammar	4	4	15	10	33	15
B: Language Interference	3	3	9	6	21	10
C: General Vocabulary	8	8	15	13	44	20
D: Legal Terms and Phrases	3	3	11	16	33	15
E: Idioms and Sayings	0	0	7	4	11	5
F: Register	1	1	5	3	10	5
G: Numbers and Names	1	2	6	5	14	7
H: Markers, Intensifiers, Emphasis and Precision	3	3	9	7	22	10
I: Embeddings and Position	1	1	9	8	19	9
J: Slang and Colloquialisms	1	0	4	3	8	4
<b>Total Number of Units</b>	<b>25</b>	<b>25</b>	<b>90*</b>	<b>75</b>	<b>215</b>	<b>100</b>

\* Some early versions of the NCSC's exams included 75 scoring units in the consecutive section. Subsequent exams include 90 scoring units.

## 8.2 Distribution of Scoring Units in the Consecutive Section

When drafting the consecutive section of an examination, the utterances to be interpreted should meet the following standards:

- They should be of varied lengths, ranging from one word to no more than fifty words.
- When drafting the consecutive section, the test development team shall abide by the following provisions at the beginning of the script in order to give examinees a chance to get started and into the rhythm of consecutive interpretation:
  - No more than twenty words and no scoring units should be used in the first two utterances.
  - In the next two utterances, the number of scoring units should be relatively small, and whatever scoring units are included should not be especially difficult or challenging.
- The scoring units must be embedded in utterances that vary in length approximating the distribution shown in Table 3.
- The number of scoring units within utterances should generally be proportionate to the length of the utterance so that short utterances have fewer scoring units and long utterances have more scoring units (e.g., utterances in the range of 1-10 words should have no more than two scoring units and utterances in the range of 40-50 words should have four or five scoring units). The maximum number of scoring units that may be assigned to any English source-language utterance is four and the maximum for any utterance in the other source language is five.

- As illustrated in Table 3, the percentage of scoring units found in utterances ranging from 21 to 30 words must always be greater than the percentage of scoring units found in utterances ranging from 11 to 20 or 31 to 40 words.
- The percentage of scoring units in utterances ranging from 11 to 40 words may deviate by up to 10 percent from the percentages indicated in the table below.

Table 3: Scoring Unit Distribution in the Consecutive

<b>UTTERANCE LENGTH</b> <i>(In Number of Words)</i>	<b>TARGET %</b>	<b>TARGET N</b>	<b>TARGET %</b>	<b>TARGET N</b>
<i>1-10</i>	10%	4	10%	5
<i>11-20</i>	25%	10	25%	12 or 13
<i>21-30</i>	30%	12	30%	15
<i>31-40</i>	25%	10	25%	13 or 12
<i>41-50</i>	10%	4	10%	5
<b>Total %</b>	<b>100%</b>		<b>100%</b>	
<b>Number of Scoring Units</b>		<b>40</b>		<b>50</b>

- 8.3 Distribution of Scoring Units in the Beginning of the Simultaneous Section  
 When drafting the simultaneous section of any examination, no scoring units should be included in the first ten seconds of the recording (i.e., approximately the first twenty words) in order to give examinees a chance to get started.

[Section 9]

## Distribution of Scoring Units: Abbreviated Examination

- 9.1 When an abbreviated test is being developed, the scoring units should be distributed within the test as outlined in Table 4. The test development team should review the appropriateness of these percentages for each language combination and recommend adjustments to the TDM when the linguistic properties of English and the other language appear to make an adjustment necessary. It is acceptable for the TDM to approve deviation from the target number of scoring units by plus or minus 10 percent, but if the deviation exceeds 10 percent, the TDM should consult with members of the Technical Committee for input and direction.

**Table 4: Scoring Unit Distribution for Abbreviated Model**

SCORING UNIT CATEGORY	TOTAL # OF UNITS	TARGET % OF UNITS
A: Grammar	16	21%
B: Language Interference	0	0
C: General Vocabulary	14	19%
D: Legal Terms and Phrases	13	17%
E: Idioms and Sayings	4	5%
F: Register	4	5%
G: Numbers and Names	5	7%
H: Markers, Intensifiers, Emphasis and Precision	8	11%
I: Embedding and Position	8	11%
J: Slang and Colloquialisms	3	4%
<b>Total Number of Units</b>	<b>75</b>	<b>100</b>

9.2 The length and assignment of scoring units at the beginning of the simultaneous section in the abbreviated model shall comply with the requirements for the same test section in the full three-section examination and the words should be counted as described in Section 10.2.

[Section 10]

## Length of Test Sections

10.1 The number of words in each test section must be within the ranges shown in Table 5.

**Table 5: Number of Words in Test Sections**

SECTION OF TEST	RANGE IN NUMBER OF WORDS
<b>Sight translation</b>	400-450
English → Other Language	200-225
Other Language → English	200-225
<b>Consecutive</b>	850-950
English → Other Language	400-450
Other Language → English	450-500
<b>Simultaneous</b> (for both standard and abbreviated models)	800-850
<b>Total</b>	<b>2050-2250</b>

10.2 When English is the source language of the test script, words should be counted as follows:

- Any word that would be separated from other words by a space or punctuation when written constitutes one word.
- Each portion of a hyphenated word counts as one word (e.g., “twenty-three” counts as two words).
- Numbers should be counted the way they would be spoken aloud (e.g., “Two thousand six,” or “Two thousand and six”).
- If there is a succession of letters and/or numbers as in a license plate (e.g., FLN18P) or as in a case number (e.g., DV-1998-23), each letter should be counted as one word and the numbers would be counted as indicated above.

10.3 When a language other than English is the source language, the TDM shall obtain guidance from the specialists for that language during the writing of the first test on whether the counting methodology set forth in Section 10.2 works in that language.

- When specialists for a given language indicate that those basic principles work for that language, the TDM shall make a written notation to that effect in the master file for that exam. If there are any exceptions to those principles, those exceptions shall also be noted.
- However, when specialists for a given language indicate that those basic principles do not work in that language, the TDM shall guide them in writing a set of operating principles for the counting of words that are functionally equivalent to the general principles set forth for English. The TDM shall file a written report of these principles in the master file for that exam.

[Section II]

## Production of Test Materials for Test Administration

---

The TDM, test development team, and LASS staff are responsible for producing test scripts that are ready for administration. This section describes how the script for each section of an examination should be prepared for that purpose.

11.1 The sight translation documents to be handed to the candidates during the first section of the examination should be produced as follows:

- Both parts should include only the text of the test material and not include any headers, footers, or other writing that is not test content. The margins of the document should be set at one inch at the top, bottom, left, and right of an 8.5 x 11 standard white sheet with no watermarks.

- The text should fit on a single sheet of paper. In the event that the text is too long to fit on a single sheet, the font size should be adjusted as little as is necessary to fit the text on one sheet.
- No scoring units should be identified by bolding, underlining, or including identifiers in superscript.
- The text should be arranged in paragraph format, with the first line of each paragraph indented.
- The document written in English should be typed in Times New Roman font, size 12, double spaced, and left justified in portrait layout.
- The document written in the other language should follow the rules for the language's writing system per the leading authority for writing the language, and follow the same font requirements as mentioned above to the degree that they apply to the orthography of the language.
- The document written in the other language should follow the appropriate orthographic rules of the language, including the production of more than one version if the language has more than one written form (for example, traditional and simplified for Cantonese and Mandarin).

11.2 Two versions of each test script shall be developed:

- One "unmarked" version for proctors, which does not include the identification of scoring units, but does include the headers and footers and other requirements referred to in Section 12.1, and
- One "marked" version for raters, which includes the identification of all scoring units and conforms to all the requirements referred to in Section 12.1.

11.3 Production of CD recordings for the consecutive and simultaneous sections of the examinations:

- The recording for both the consecutive and simultaneous sections should be produced as follows:
  - The speakers should strive to emulate the quality of broadcast voice.
  - The recordings should be produced so that they are clear, free from extraneous noise (including hiss), and easily understood.

11.4 The recording by which the consecutive section is administered should be produced as follows:

- It should be recorded at a conversational rate of speech, neither fast nor slow, clearly enunciated, and with enough expression to appropriately emote the text.
- The person who reads the English portion of the script should be a native speaker of English and the person who reads the non-English portion of the script should be a native speaker of that language.
- A standardized amount of pause time between tracks should be inserted by appropriate software so that individuals proctoring the exams expect the same amount of pause time and can seamlessly and consistently administer the consecutive test regardless of the language.
- The gender of the person reading each role should match the gender of the role he or she is reading.
- The test development team should review the initial CD recording to ensure that the recording and the test script are consistent and that there are no errors introduced by the readers whose voices were recorded or in the process of producing the recording itself. Any problems discovered by this review should be corrected before the recording is considered ready for administration.

11.5 The recording by which any simultaneous section is administered should be produced as follows:

- It should include a standardized, 45- to 60-second introduction, which includes identification for the examinee of the first sentence or phrase of the script followed by five seconds of silence before the test begins.
- It should be recorded throughout at a speed of 120 words per minute, as far as is possible. No single minute should be slower than 110 words per minute or faster than 130 words per minute. The introduction should be recorded at the same volume as the test content.
- The TDM should review the initial CD recording to ensure that the recording and the test script are consistent and that there are no errors introduced by the readers whose voices were recorded or in the process of producing the recording itself, and that the pacing complies with the requirements articulated in the previous bullet. Any problems discovered by this review should be corrected before the recording is considered ready for administration.

12.1 The final versions of the test scripts to be used in the rating process should conform to the following editorial guidelines, which are illustrated in Appendix 2:

- Margins should be set at 1 inch at the top, bottom, left, and right of an 8.5 x 11 standard white sheet, with no watermarks.
- Scripts should be typed in Times New Roman font, font size 12, double spaced, and left justified in portrait layout.
- The consecutive and simultaneous test scripts should be paginated using page numbers of the same font and font size as the text in the center of the bottom margin. Since each sight part is a single page, the sight scripts need not be paginated.
- Consecutive scripts should be formatted so that each utterance in English is preceded by a “Q.” and each utterance in the other language is preceded by an “A.” with the Q or A appearing at the left margin and the corresponding text indented by one-half inch.
- All scoring units should be bolded and underlined, and each scoring unit should be identified in superscript by a number and type identification, which should be bolded, but not underlined. For example: **construction manual**<sup>1C</sup>

If a scoring unit begins on one line of the script and wraps to the next line, a hard return should be inserted immediately prior to the first word of the scoring unit to artificially push it to the next line, for example:

“We expect to hear that witness testify **the day after tomorrow**<sup>29H</sup>.”

Changed to:

“We expect to hear that witness testify **the day after tomorrow**<sup>29H</sup>.”

- When a word or words that are not part of a scoring unit are located between words that constitute a scoring unit, the script should be marked as follows: “The suspect clearly intended to steal **not only** the car stereo **but also**<sup>12A</sup> the GPS.” No other scoring unit may be assigned between the separated components of such scoring units.
- Sometimes punctuation precedes or immediately follows a scoring unit. When this happens, the following guidelines shall be followed in producing the script:
  - Punctuation markers (e.g., comma, period, question mark, exclamation point, open or close quotation mark) should not themselves be bolded or underlined.

- When there is a punctuation marker immediately following the last letter of a word in a scoring unit, the superscripted identifying information should be inserted before that punctuation marker, as illustrated in the following example: “...his eyes rolled back in his head and he **collapsed**<sup>25C</sup>.”
- Each script shall be identified in the header of the document, in Times New Roman font, bolded, font size 14. The header should contain:
  - at the left margin, the test form identification (language and form number),
  - in the center, identification of the test section and, if applicable, part; and
  - at the right margin, the script identification.
- At the bottom right of the last page of each test section and the two parts of the sight translation section, there should be a grid for raters to complete, with the summary of the point score for that test section. See Appendix 2 for an example.
- Numbering corresponding to recorded tracks of the consecutive section of the examination shall be included in the left margin of both marked and unmarked consecutive scripts.

12.2 The scoring dictionaries that are created for each test section should adhere to the following editorial guidelines, which are illustrated in Appendix 3:

- Scoring dictionaries should ordinarily be created and maintained as an Excel document in landscape layout. However, Word table format may be used instead when the test development and maintenance functions would be easier or more efficient, or when there are special needs for a given language’s font that make using Word either necessary or significantly more efficient. In instances where Word table format is selected, it shall also be in landscape layout and follow the same editorial guidelines provided in this section.
- The header should be typed in Times New Roman font, bolded, font size 14, and should identify the language, test form, and script Identification.
- Column headings should be as follows:
  - Column 1: No.
    - This column should contain the number of the scoring unit within the test sections (i.e., 1 through 25, or 1 through 90, etc.)
  - Column 2: Scoring Unit Type and Description
    - This column should contain the alpha identification of the type of scoring unit (A, B, C, D, E, F, G, H, I, or J), a colon, the name of the scoring unit category (e.g., Legal Terms and Phrases or Grammar).
  - Column 3: Scoring Unit Texts
    - The word(s) or numbers of the actual scoring unit should be provided in this column.

- Column 4: Acceptable Response(s)
  - This column should contain all of the acceptable interpretations of the scoring unit, separated by semi-colons.
- Column 5: Unacceptable Response(s)
  - This column should contain all of the unacceptable interpretations of the scoring unit, separated by semi-colons.
- Column 6: Notes
  - This column should contain any instructions, hints, or explanations for test raters about why the test development team classified a particular scoring unit as they did or other guidance on how to rate it. This column will later contain additional notes from the raters.

13.1 Some form of field testing shall be conducted on every new test before it is released to members for routine administration and rating. Relying on the advice and expertise of the specialists serving on the test development team and any other expert the TDM selects, the TDM will coordinate the form of field testing listed below that is most appropriate for a given language. Field testing is a simulation of operational test administration that identifies scoring units that do not function as intended or pose other problems for candidates or raters, identifies additional acceptable and unacceptable responses for the scoring dictionary, and uncovers other problems in the structure, content or mechanics of a particular test. After field testing, the TDM and specialists must make needed adjustments identified in the field testing to the test instrument before it is considered an operational product. The options for this phase of test development are as follows:

- For languages which are spoken by two or more distinct speech communities around the world and for which there is a large number of prospective candidates, field testing should be administered to a sample of test takers who may or may not be taking the test in order to obtain a court interpreting credential, come from at least two geographically separate parts of the United States, and represent as many different speech communities of the language as possible. Any of these components may be relaxed if the TDM finds that it is not feasible and documents the reasons for reaching that conclusion.
- For languages which are spoken by two or more distinct speech communities around the world and for which there is a small number of prospective candidates, field testing should ideally be administered to a sample of test takers who may or may not be taking the test in order to obtain a court interpreting credential and

represent at least the language’s largest speech communities represented by residents of the United States. If the sample includes test takers who seek a court interpreting credential through the field test, the results shall be considered final only after all field test candidates have been tested, all field tests have been analyzed, and the scoring units and other necessary adjustments have been made, and the field tests are then graded on the basis of the final script(s) of the test instrument(s).

- For languages which are spoken by essentially a single speech community, field testing should ideally be administered to a sample of test takers who may or may not be taking the test in order to obtain a court interpreting credential. If the sample includes test takers who seek a court interpreting credential through the field test, the results shall be considered final only after all field test candidates have been tested, all field tests have been analyzed, and the scoring units and other necessary adjustments have been made, and the field tests are then graded on the basis of the final script(s) of the test instrument(s).
- When a test is being developed in any language for which the foregoing options are not feasible, field testing must be given to a minimum of two candidates and the results shall be considered final only after all field test candidates have been tested, those field tests have been analyzed, and the scoring units and other necessary adjustments have been made, and the field tests are then graded on the basis of the final script(s) of the test instrument(s).

## [Section 14]

## Revising Existing Tests

---

- 14.1 All tests shall undergo several types of ongoing review and evaluation. Raters who grade tests sometimes discover problems in test scripts such as scoring units that do not perform well over time or shortcomings in the selection and/or classification of scoring units. In addition, raters often find that renderings listed in the scoring dictionary should be transferred from acceptable to unacceptable, or from unacceptable to acceptable, or that examinees present new renderings not contained in the scoring dictionary that should be added as acceptable or unacceptable. Finally, LASS staff, whether on their own or through expert consultants, may find through standard psychometric analyses that scoring units are not performing well and need to be replaced.
- 14.2 In order to ensure that all test scripts and scoring dictionaries are revised as needed in a timely manner, LASS staff shall designate two experts to serve as a test maintenance team for each language and maintain and update as needed a list of the members of all such teams. Such teams should consist of the experts who wrote the test or, when neither or only one is feasible, other experts as defined below. Such test maintenance teams will be responsible for revising each test form as well its corresponding scoring dictionary. Other experts who may serve on test maintenance teams are professional

linguists whose totality of experience as test raters, linguists, and, if applicable, court interpreters, indicates they possess the expertise needed to perform these duties. The TDM shall designate one member of the team to be the editor who updates the electronic file of the test form as well as the scoring dictionary.

- 14.3 Under the oversight of the TDM, test maintenance teams shall analyze and determine how to handle all proposed revisions of test scripts offered by themselves (which may arise out of their own work as raters or when asked to review or evaluate an existing test form) or by others (which would come from raters other than themselves or from staff or consultants as a result of psychometric or content analysis). Such teams shall then determine the final disposition of all proposed revisions. When the nature of the revisions is deemed by the TDM, in consultation with a test maintenance team, to be sufficient to issue a new version, it should be reissued under a new number (e.g., 1.2 would be the second version of test #1, 1.3 would be the third version of test #1). The TDM will then withdraw all copies of the previous version of the test form and commence distribution of the new test form in whatever manner is deemed appropriate to ensure that no member administers the superseded version.
- 14.4 Under the oversight of the TDM, each test maintenance team shall also be responsible for all updates of that test form's scoring dictionary. Raters should identify possible revisions or additions to a scoring dictionary when any of the following occur:
- A candidate correctly or incorrectly interprets a scoring unit in a way that does not appear in the scoring dictionary and, in its opinion, may be used by future candidates.
  - A rendering appears in the scoring dictionary as acceptable which should actually be listed as unacceptable.
  - A rendering appears in the scoring dictionary as unacceptable which should actually be listed as acceptable.
- 14.5 When the team of raters consists of the same persons as the test maintenance team, they shall ordinarily revise a test script or scoring dictionary as soon as they have completed their analysis and determined that the revision is warranted as directed by the TDM. However, when the team of raters includes one or more raters who are not a member of the test maintenance team, they must submit in writing to the test maintenance team the proposed revision to a test script or scoring dictionary. The test maintenance team shall review all such submissions in a timely manner and revise a test script and/or scoring dictionary as directed by the TDM.
- 14.6 If the test maintenance team reviewing the suggested revisions disagrees with the suggestion or is uncertain of the accuracy or validity of the suggestion, it may confer with other raters or members of the test development team as part of its research and analysis.

- 15.1 LASS staff should collect and preserve information about the development and maintenance of each test, including the names and qualifications of every person involved in the writing and review of the test instrument.
- 15.2 The TDM will document for staff any deviations from the construction standards as described in this Manual, articulating why such deviation was necessary, and noting any specific challenges or problems faced by the test development team and reviewers.

- 16.1 After an adequate sample of examinees has taken a given test version, LASS staff should conduct an analysis of the test to address any reliability and validity issues that may arise. Information obtained through the analysis will be used to revise or delete poor performing items from future test versions.



# Appendices

---



CRITERIA	DESIRABLE	UNDESIRABLE
<b>Be clear</b>	Unit's testing purpose must be clear and unambiguous	Unclear, ambiguous, permitting raters to disagree on how to rate
	Discourse in which a SU is embedded must be clear	Context is ambiguous or unclear
	<i>Note, however, that certain words may be ambiguous in the sense that they are generic terms and are not specific. For example, "gun" is a general term for "weapons that shoot bullets" and it is not precise, but it can be a scoring unit only if the target language has a similarly imprecise term for "gun."</i>	
<b>Promote objective and reliable assessments</b>	Must be amenable to being graded objectively and without variation from rater to rater or rating team to rating team	Requires subjective consideration and guessing by raters, open to many understandings, arguments among raters, opens door to raters' personal preferences
<b>Be in current usage</b>	Words/phrase used in contemporary parlance; could appear in today's courtroom	Esoteric, outdated, archaic, not likely to be heard in today's courtroom
<b>Be in widespread usage (except for certain legal terms or concepts which are more typical of the register of law)</b>	Words/phrases used by wide range of the language's native speakers Legal terms that are fairly common and universal across the United States and its possessions	Words/phrases of regional usage or, if spoken in multiple countries, a single country or small subset of countries Legal terms that are highly technical, infrequently used, or peculiar to a state or region
<b>Be unique</b>	Appears only once within a test (except that verb forms [e.g., tense, mood] may be used only once per test section)	Appears two or more times as scoring units (e.g., perhaps as a verb once and as a general vocabulary item elsewhere)
<b>Be concise and focused</b>	Consists of the smallest linguistic unit possible: a single word if possible and, if not, a small set of words that constitute a meaningful unit (such as idioms, proverbs, frozen legal language)	Consists of too many words so that it's not clear what's really being measured or requires raters to assess too many variables within the scoring unit

CRITERIA	DESIRABLE	UNDESIRABLE
<b><i>Be appropriate to the court context</i></b>	Must be something that could reasonably be said by anyone in a courtroom, whether on or off the record (does not have to be common or frequent), or in court-related contexts	The word or phrase would have little or no chance of ever being used by anyone speaking on the record in a court of law or in a related context such as arbitration, mediation, & during contacts with court support staff
<b><i>Be appropriately distributed in the test text</i></b>	Must be spread out as much as possible May be next to each other so long as there is space between those preceding and following Must follow specifications regarding distribution in §§8-9 of the TCM	Too many chunked together within one sentence or utterance Appear too early in the consecutive or simultaneous section Distribution deviates beyond the permitted range set forth in the TCM
<b><i>Be reasonable</i></b>	Must reflect ordinary, actual usage of any type of typical speaker Some may even be difficult or challenging	Cute, tricky, or sneaky—a function of test writer’s joy in playing games with language rather than a measure of actual usage Unrealistically difficult terms that entry-level interpreters have little or no chance of accurately interpreting, especially in the simultaneous section

Good morning. My name is Janet Smith and I am an assistant **state’s attorney**<sup>1D</sup> in our **county**<sup>2C</sup>. Ladies and gentlemen of the **jury**<sup>3D</sup>, we believe that the evidence will show that during the morning **hours**<sup>4I</sup> on **July 2<sup>nd</sup>, 2004**<sup>5G</sup>, a woman by the name of **June Jones**<sup>6G</sup> was brutally **beaten**<sup>7A</sup> and assaulted in the living room of her home here in the city of Madison. She was held down on her living room floor for two hours, and her **assailant**<sup>8B</sup> took her **purse**<sup>9C</sup> when he **finally**<sup>10I</sup> ran away. **We will prove**<sup>11A</sup> that the **defendant**<sup>12D</sup>, **Omar Butler**<sup>13G</sup>, is the person who assaulted, restrained and **stole from**<sup>14A</sup> June Jones. The State has **charged**<sup>15D</sup> Mr. Butler with **attempted sexual assault**<sup>16D</sup>, **false imprisonment**<sup>17D</sup> and theft.

June is a 42 year old **divorcée**<sup>18C</sup> who lives **alone**<sup>19I</sup> at **1729 Rosewood Avenue**<sup>20G</sup>. On the night of **Friday, July 1<sup>st</sup>**<sup>21G</sup>, June was at her house with a friend, **hanging out**<sup>22J</sup> and drinking. **Late**<sup>23I</sup> in the evening, the victim told her friend that she wanted to go over to Harry’s Bar. Her friend said that he wanted to **call it a night**<sup>24E</sup>, so June decided to go by herself.

Possible Points	
(-) # Incorrect	
Total Correct	



<b>No.</b>	<b>SCORING UNIT CATEGORY</b>	<b>SCORING UNIT TEXTS</b>	<b>ACCEPTABLE RENDITION(S)</b>	<b>UNACCEPTABLE RENDITION(S)</b>	<b>NOTES</b>
1	D: Legal Terms/Phrases	state’s attorney	<i>Acceptable</i>	<i>Unacceptable</i>	<i>Notes from members of the test development team to test</i>
2	C: General Vocabulary	County	<i>interpretations</i>	<i>interpretations</i>	<i>ratets (or notes by test ratets</i>
3	D: Legal Terms/Phrases	Jury	<i>of the various</i>	<i>of the various</i>	<i>for the benefit of subsequent</i>
4	I: Embeddings/Position	Hours	<i>scoring units would</i>	<i>scoring units would</i>	<i>ratets) would be entered in</i>
5	G: Numbers & Names	July 2 <sup>nd</sup> , 2004	<i>be entered in these</i>	<i>be entered in these</i>	<i>these spaces.</i>
6	G: Numbers & Names	June Jones	<i>spaces.</i>	<i>spaces.</i>	
7	A: Grammar	Beaten			
8	B: Language Interference	Assailant			
9	C: General Vocabulary	Purse			





