

*Court Interpreter Oral Examination:*

# Test Rating Manual

**Consortium for Language Access in the Courts**

***Equal Justice: Bridging the Language Divide***

*The Mission of the Consortium is to inspire and enable its members to promote equal access to justice in courts and tribunals by eliminating language barriers for persons with limited English proficiency.*



Copyright © 2010 by the National Center for State Courts, on behalf and for the use of the Consortium for Language Access in the 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 Consortium members, staff and consultants to reproduce and distribute this publication for educational purposes if the copies credit the copyright holder.

For additional information, please contact:

Consortium for Language Access in the Courts  
National Center for State Courts  
300 Newport Avenue  
Williamsburg, VA 23185

This document is also available on the Consortium for Language Access in the Courts Web site:

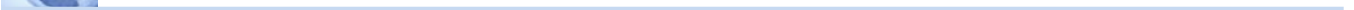
[http://www.ncsc.org/Web%20Document%20Library/EC\\_StateInterpCert.aspx](http://www.ncsc.org/Web%20Document%20Library/EC_StateInterpCert.aspx)

Publication Date: October 2010

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

---

*Consortium for Language Access  
in the Courts*



## TABLE OF CONTENTS

---

<b>Purpose, Authority and Applicability of this Manual</b> .....	1
[Section 1] <b>Selection of Raters</b> .....	2
[Section 2] <b>Rating Supervisors</b> .....	2
[Section 3] <b>Classification of Raters</b> .....	3
[Section 4] <b>Rating Consortium Exams for Which There Are Approved Lead Raters</b> .....	4
[Section 5] <b>Rating Examinations for Which There Are No Approved Lead Raters</b> .....	4
[Section 6] <b>Rating Examinations for Which There Are No Approved Raters</b> .....	5
[Section 7] <b>General Rating Principles and Confidentiality</b> .....	5
[Section 8] <b>Conflicts of Interest</b> .....	6
[Section 9] <b>Overview of the Test Rating Process Involving a Team of Raters</b> .....	7
[Section 10] <b>Criteria for Accepting Renditions of Scoring Units</b> .....	8
[Section 11] <b>Degree of Precision Required Per Category of Scoring Unit</b> .....	10
[Section 12] <b>Guidelines for Assessing Scoring Units as Correct or Incorrect</b> .....	11
[Section 13] <b>Mechanics of Scoring Exams as Correct or Incorrect</b> .....	12
[Section 14] <b>Scoring Unit Suggestion Forms and Rater Research</b> .....	15
[Section 15] <b>Rater Concerns or Questions about Tests</b> .....	17
[Section 16] <b>Training for Raters</b> .....	18
[Section 17] <b>Supervision of Rater Performance</b> .....	18
[Section 18] <b>Disqualification of Raters</b> .....	19
[Appendix 1] <b>Training Provider Disclosure</b> .....	23
[Appendix 2] <b>Training Provider Disclosure Agreement</b> .....	24
[Appendix 3] <b>Scoring Unit Suggestion Form</b> .....	25
[Appendix 4] <b>Test Results Report Form (200 Scoring Units)</b> .....	26
[Appendix 5] <b>Test Results Report Form (215 Scoring Units)</b> .....	27
[Appendix 6] <b>Marked Script (Sample)</b> .....	28



## Purpose, Authority and Applicability of this Manual

---

*The Court Interpreter Oral Examination: Test Rating Manual* was written by the Consortium for Language Access in the Courts' Technical Committee. The Technical Committee concerns itself with matters related to Consortium test development, test administration, oral test rating, test overviews, statistics, and psychometric analyses. It also considers, drafts, and recommends modifications to Sections 6 and 8 of the *Consortium Agreements*<sup>1</sup> which relate to Oral Test Instruments and Standards for Test Administration.

As required by testing industry standards,<sup>2</sup> this Manual frames the Consortium's process of developing and maintaining a sufficient supply of trained raters<sup>3</sup> and establishing standards to ensure they will rate oral exams in a valid and reliable manner. Consortium members who use these raters must be able to rely on them with confidence as part of the process of credentialing persons who wish to become certified spoken language court interpreters, regardless of who rates a given oral examination, when or where the examination is rated, or what the language combination may be.

The responsibility of test writers is to develop oral exams that 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 Consortium tests measure 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.

The responsibility of *raters* is to rate oral exams taken by candidates in an objective, valid and reliable manner consistent with the nature and purposes of those tests. This *Test Rating Manual* provides guidelines for the oral examination rating process to ensure that all tests are so rated for all languages and all versions within the same language.

This *Test Rating Manual* is binding upon all raters of oral examinations administered by any member of the Consortium for Language Access in the Courts as well as upon those who administer the rating process, both program managers of member jurisdictions who enter into contracts directly with such raters as well as Consortium staff at the National Center for State

---

<sup>1</sup> See *Agreements for Consortium Organization and Operation*, Consortium for Language Access in the Courts, [http://www.ncsc.org/Web%20Document%20Library/EC\\_StateInterpCert.aspx](http://www.ncsc.org/Web%20Document%20Library/EC_StateInterpCert.aspx), Section 3.5.1.

<sup>2</sup> Standards 3.22-24, *Standards for Educational and Psychological Testing* (1999) issued by the American Educational Research Association, the American Psychological Association, and the National Council on Measurement in Education.

<sup>3</sup> The term for the individuals who grade exams preferred by the *Standards*, *ibid.*, is "scorers" and the process is termed "scoring" tests (see Standard 3.23). New Jersey's court interpreter program has historically used the term "examiners" and another term that could be used is "graders." Each of these terms may be used interchangeably, but the Consortium has historically preferred and used the terms "raters" and "rating."

Courts (NCSC) with whom program managers contract to coordinate rating services.<sup>4</sup> It sets forth how raters are selected, trained, classified, and supervised, as well as how they will perform their rating duties.

## [Section 1]

## Selection of Raters

---

- 1.1 Raters are recruited, selected, trained, and supervised by Consortium staff under the oversight of the Technical Committee.
- 1.2 Raters who have satisfactorily completed the rater selection process, continue to be in good standing, and are available to rate exams are identified on the Rater Contact List posted on the members-only section of the Consortium’s website under “Rater Contact Information.”
- 1.3 Only individuals who are on the posted list of raters (see Section 1.2) and whose rating services are contracted and managed in accordance with the provisions of Sections 2 through 5 below may serve as raters of Consortium examinations.
- 1.4 When Consortium members are prepared to administer exams that need to be rated, they may either contract with Consortium staff to coordinate the rating process or contract directly with raters who are on the Rater Contact List.
- 1.5 When any exam is rated, whether the rating process is managed directly by program managers or by Consortium staff, at least one Approved Lead Rater must be used if one exists in that language. If no Approved Lead Rater exists in that language, program managers and staff may contract with Approved Raters, provided their rating activities are overseen by a Rating Supervisor. If there are no Approved Raters in that language, Provisional Raters may be contracted, provided their rating activities are overseen by a Rating Supervisor.

## [Section 2]

## Rating Supervisors

---

- 2.1 Rating Supervisors are individuals who have extensive experience with the rating process and include the following groups: Approved Lead Raters, Consortium staff with extensive rating or rater supervision experience, and current or former program managers who have extensive rating or rater supervision experience.

---

<sup>4</sup> See Section 8.3 of the *Agreements*, *ibid.*

- 2.2 Rating Supervisors oversee the process of coming to consensus<sup>5</sup> for the Approved or Provisional Raters. This includes encouraging those raters to conduct the appropriate research when needed, consider the appropriate latitude allowed with each type of scoring unit, and develop the skills to become consistent and accurate raters.
- 2.3 In order to ensure the integrity of the rating process, a Rating Supervisor must be present at every consensus rating process undertaken by Approved or Provisional Raters.
- 2.4 Consortium staff will maintain a list of Rating Supervisors on the members-only section of the Consortium’s website.

## [Section 3]

## Classification of Raters

---

- 3.1 Consortium staff shall classify and, when appropriate, reclassify raters according to guidance from the Technical Committee.
- 3.2 There shall be two categories of raters: Approved and Provisional.
  - Approved Raters include Approved Lead Raters and Approved Raters.
    - Approved Lead Raters shall be highly experienced approved raters who have rated a substantial number of Consortium and/or federal examinations, served as an Approved Rater in good standing for at least one year, and demonstrated reliable, timely, professional performance of rating (and, if applicable, test writing) duties, and completion of continuing education requirements established by the Technical Committee.
    - Approved Raters shall be raters who satisfactorily complete the application process, demonstrate reliable, timely and professional performance of rating (and, if applicable, test writing) duties, and complete continuing education requirements established by the Technical Committee.
  - Provisional Raters are those individuals who served on the test writing team for a test in a new language. They serve as Provisional Raters until such time as they are eligible to be classified as Approved Raters.

---

<sup>5</sup> See Section 9 for a description of “consensus rating process.”

## Rating Consortium Exams for Which There Are Approved Lead Raters

---

- 4.1 Two raters are required, at least one of whom must be an Approved Lead Rater, in the following circumstances:
- The test to be rated is a complete oral examination, which includes both the standard model (all three sections) and the abbreviated model (simultaneous only).
  - The test is not a complete oral examination and the member administering the test program does not have a bifurcated examination process (see Section 4.2).
  - The candidate is a retest candidate.
- 4.2 Members who have a bifurcated examination process in which the simultaneous section is administered first, and the sight and consecutive sections are administered subsequently only to those who have qualified on the simultaneous, may contract with one rater to rate the simultaneous section of the examination given the following provisions:
- The rater is an Approved Lead Rater for that language combination who rates exams on a consistent and frequent basis.
  - If the Approved Lead Rater's score on the simultaneous section of the examination is in the range of 50 percent to 82 percent inclusive, that section is rated by a second approved rater and both raters conduct the consensus process through which they determine the final score for the candidate.
  - Two raters are used to rate the consecutive and sight translation sections of the exam for those candidates who received a passing or other qualifying score on the simultaneous section.

## Rating Examinations for Which There Are No Approved Lead Raters

---

When there is no Approved Lead Rater for a given language but there are Approved Raters for that language, program managers or Consortium staff should contract with two Approved Raters and a Rating Supervisor. The raters will independently rate each test and the Rating Supervisor will join them for the consensus process.

When examinations require rating and there are no approved raters on the Consortium's list, then two Provisional Raters may be contracted. This is likely to happen only when an oral examination is new and no test candidates have passed the examination. Program managers are encouraged to consult with Consortium staff<sup>6</sup> when this situation occurs.

- 6.1 Program managers or staff should contract with two Provisional Raters. The raters should evaluate the three-section test, abbreviated test, or test section independently. When they are ready to conduct the consensus process, a Rating Supervisor should observe or otherwise oversee the consensus process. The team shall report the final score to the program manager or staff and that score should be reported to the examinee.
- 6.2 The Rating Supervisor should reiterate the process of coming to consensus that is discussed during rater training, encouraging the Provisional Raters to conduct the appropriate research, consider the appropriate latitude allowed depending upon the type of scoring unit, and develop the skills to become consistent and accurate raters.

- 7.1 Rating court interpreting performance examinations is a challenging activity which requires extensive knowledge of the two languages involved and of the professional standards of court interpreting. Raters must also possess the following skills: the ability to focus and concentrate on test performances, an understanding and adherence to the principles of the rating process, a generous and humble grasp of the nature of human communication as it is expressed through language, as well as the ability to work efficiently, effectively and collegially with other raters.
- 7.2 Most importantly, raters must strive for complete neutrality, fairness, and objectivity throughout all rating activities.
- 7.3 Raters should discuss the performance of an examinee only with a Consortium program manager, Consortium staff, or another rater, and then only when it is necessary to do so.

---

<sup>6</sup>Go to: [http://www.ncsconline.org/D\\_Research/CourtInterp/Res\\_CtInte\\_ConsortCertStaffPub.pdf](http://www.ncsconline.org/D_Research/CourtInterp/Res_CtInte_ConsortCertStaffPub.pdf) to contact a staff member.

- 7.4 Raters must never discuss the outcome of an examination with an examinee, or have any contact with examinees about any examination. Should an examinee attempt to communicate with a rater about the examination or the results thereof, the rater should not permit the communication to go forward and should report such attempts to the office that contracted such rating services, i.e., either the member program manager or Consortium staff.
- 7.5 Raters must be open to new language usage, as well as geographical and social variations in usage.
- 7.6 Raters must be capable of recognizing when an interpretation of a given scoring unit rendered by an examinee is correct and marking that scoring unit as correct, even when the rendition is not the best one available, is not one that the rater prefers, or is not one that the rater would have employed.
- 7.7 Raters must never disclose any of the content of any exam to anyone.

## [Section 8]

## Conflicts of Interest

---

- 8.1 When a rater clearly recognizes the voice of an examinee, he/she must contact the program manager or Consortium staff for instructions about how to proceed. In the event a rater actually rates an exam by someone whom the rater can identify, the rater may not disclose any information about the candidate or the exam to anyone other than may be necessary within the context of the rating process itself.
- 8.2 Insofar as is possible, no identifying information about any candidate or group of candidates shall be provided to the raters other than the candidate ID number assigned by Consortium staff.
- 8.3 There are occasions when a rater also serves as an interpreter trainer, e.g., delivering training at institutions of higher education, orientation programs, skill-building seminars, or professional conferences. Although serving as a trainer does not automatically disqualify an individual from serving as a rater, such information must be disclosed to avoid the appearance or perception of a conflict of interest.
- Appendix 1 provides a form which program managers should use when hiring or contracting with an expert or vendor for training purposes. The form requires disclosure of any rating and training activities engaged in by the expert or vendor, or any employee of the expert or vendor, that may be perceived as a conflict of interest.
  - Appendix 2 is a sample oath form that raters who are also trainers should execute and turn in to program managers. It certifies that if a training expert or vendor

serves as a rater or test writer, or is in any way familiar with Consortium test content, the trainer will not use or disclose test content in any way in their training materials or presentations. In addition, the oath form includes a promise by any trainer who serves as a rater or test writer or is in any way familiar with test content not to market the training session as “test preparation” training, nor advertise an increase in passing rates as a result of the training.

[Section 9]

## Overview of the Test Rating Process Involving a Team of Raters

---

The major steps of the rating process shall be as follows:

- 9.1 The entity coordinating the rating process (the program manager, or, in some cases, Consortium staff) shall supply the raters with the documentation and information they need to rate the exam(s).
  - The documentation includes the following, at a minimum:
    - a recording of each exam
    - test scripts and the corresponding scoring dictionaries
    - *Scoring Unit Suggestion Forms*
    - *Test Results Report Form*
  - The program manager or staff shall identify which rater is responsible for the rating process. That individual will typically be the Approved Lead Rater for that language combination. If there is no Approved Lead Rater, it will be an Approved Rater designated to be responsible for filling out the master copy of the rating results and facilitating the rating process.
- 9.2 Each rater individually listens to and preliminarily scores each section of the exam on the appropriate test script(s) for each exam.
- 9.3 Raters go over each exam together, reviewing each scoring unit on which they had preliminarily arrived at a different conclusion. This is referred to as the “consensus process” as the raters must agree with each other for any scoring unit to be assessed to be incorrect. The process involves making sure what the candidate actually said (which sometimes requires listening to that portion of the test again), assessing that rendition in the context of the scoring dictionary, reaching a consensus on whether the rendition is or is not acceptable, marking the test script accordingly, adding up the number of correctly rendered scoring units, and filling out scoring grids. If there are any scoring units about which the raters cannot agree initially, then the raters conduct research (see Section 14). Teams may conduct the consensus process work face to face,

telephonically, or electronically so long as the communication technique used is secure and does not compromise test content. A Rating Supervisor must be involved in the consensus process when no Approved Lead Rater is involved.

- 9.4 When research yields agreement and proposed revisions to the Dictionary, the lead rater completes and separately submits to Consortium staff a *Scoring Unit Suggestion Form* (see Appendix 3). However, when all the effort and review cannot result in agreement on whether a particular rendering for a given scoring unit should be rated as acceptable or unacceptable, the benefit of the doubt shall be given to the candidate and the scoring unit shall be counted acceptable.
- 9.5 For each section (or, in the case of the sight section, each part), the score shall be calculated using the scoring grid provided (see Section 13.7).
- 9.6 Scores will be compiled for all remaining sections, the Test Results Form will be completed, (see Section 13.9 and Appendices 4 and 5), any concerns that may have arisen with the test itself or the rating process are identified, and any other work should be completed that may be appropriate.
- 9.7 Raters shall return all materials as directed.

[Section 10]

## Criteria for Accepting Renditions of Scoring Units

---

Raters should keep several things in mind when rating examinations, including the guidance provided in this Manual, in the *Court Interpreter Oral Examination: Test Construction Manual*, and its accompanying *Court Interpreter Oral Examination: Test Writing Handbook*.

### 10.1 Grammar (A)

If the scoring unit is a verb form, it should be the same verb form in the target language where such exists and is considered appropriate usage. There is no latitude here. The same is true for grammatical constructions, e.g., if the scoring unit is testing for agreement with respect to gender and number, then each of the two elements must be precise in the target language. There is some latitude here with respect to lexical accuracy, as the primary testing goal is grammaticality, not lexical equivalence. If the rendition is within the semantic range of the source term (for example, “run,” “jog,” and “sprint” are all some form of “movement that is faster than walking”), it should be considered correct so long as the grammatical feature being tested is rendered correctly.

## 10.2 Language Interference (B)

There is some latitude here if the candidate does not fall for a false cognate or other form of language interference so long as the candidate gives an acceptable rendering in the target language, even though it is less than ideal or not completely on the mark.

## 10.3 General Vocabulary (C)

These renditions must be pretty much on the mark in terms of lexical or semantic faithfulness. Sometimes usage dictates whether or not a lexical item is acceptable. For example, common usage has “illegal immigrants” often referred to as “undocumented immigrants,” and both versions would be acceptable.

## 10.4 Legal Terms and Phrases (D)

These renditions must be pretty much on the mark in terms of lexical or semantic faithfulness. For example “Beyond a reasonable doubt” must be just that in the target language. “Beyond any doubt” would not be correct.

## 10.5 Idioms and Sayings (E)

Renditions should be considered correct if: a) an equivalent idiomatic expression is given in the target language, or b) the source language idiomatic expression is interpreted faithfully, but not in the form of an idiom. For example, if the scoring unit is “It was raining cats and dogs,” the best interpretation is into the target language using an equivalent idiom that means it was pouring or raining hard. But it is also acceptable if the interpretation in the target language is merely “It was pouring,” or “It was raining hard,” even if there is an available idiom with this meaning that the interpreter could have used but did not.

## 10.6 Register (F)

The renditions must be precise when there is a same-level register word or phrase in the target language. If the scoring unit is “uh-huh” the rendition must be equivalent in register to “uh-huh,” not raised in register to, for example, “yes,” or “of course.” Likewise, profanity or vulgarity must not be sanitized or cleaned up and, at the other end of the spectrum, high register scoring units must not be simplified.

## 10.7 Numbers and Names (G)

Interpretations of numbers, weights, and names must be precise and dates must be exact. For example, if the address is 1234 Smith Street and the examinee says, “1234 Smith Avenue”, it is wrong (except that “Avenue,” “Street,” etc., may or may not be left in the source language, but the actual street name is not to be interpreted). When referring to a year, the last two numbers are acceptable, e.g., “99” is interchangeable with “1999”; “06” with “2006.” Also, any form of a number is acceptable, e.g. one, one, three for 113, or one hundred thirteen, etc., as long as the number is accurate.

- 10.8 Markers, Intensifiers, Emphasis and Precision (H)  
For this classification of scoring unit, it is important that the force and precision of the source language scoring unit be preserved. For example, if the speaker says, “he was walking very quickly,” and the scoring unit is “very,” the candidate must include the equivalent of “very” and not say only “quickly” without the intensifier.
- 10.9 Embeddings and Position (I)  
Raters have lots of latitude here with respect to semantic and grammatical precision. If the unit is a word or phrase that might be left out because of its position, as long as something within the semantic range for the scoring unit is present, even if it is on the outside fringes of that range, it must be accepted. For example, if the speaker says, “So then what happened?”, raters should accept “And then what happened?” or “But then what happened?”
- 10.10 Slang and Colloquialisms (J)  
The same rationale as for idiomatic expressions applies here. For example, if the English source language scoring unit is “to croak,” croak may be rendered either with an equivalent colloquial expression, or convey the meaning “to die,” using a word or phrase that means “to die” but is not colloquial in form.

[Section II]

## Degree of Precision Required Per Category of Scoring Unit

---

- 11.1 While rating exams it is important for raters to keep in mind the specific testing purpose of each type of scoring unit. Some are for the purpose of seeking preservation of specific semantic content, whereas only one type of scoring unit is specifically seeking to measure grammatical correctness (see Section 10.1). When grammatical flaws occur in the other nine types of scoring units (e.g., lack of subject/verb or adjective/noun agreement) they should not generally have any effect on assessing those scoring units as incorrect if the only deficiency is grammatical in nature. Grammar errors in the other categories may result in an incorrect determination when a grammar error itself results in a change in the accuracy of the interpretation of a given scoring unit.
- 11.2 If the semantic content of the rendering is simply wrong and bears no resemblance to a faithful interpretation of a given scoring unit, it is always wrong, regardless of the category of scoring unit in which it has been classified.
- 11.3 There is only one category of scoring units where there is considerable latitude in the degree to which semantic precision is required, and that is Embeddings and Position (I).

This scoring unit tests whether these particular linguistic elements are being left out altogether versus being interpreted in a fully faithful manner.

- 11.4 At times, especially in the simultaneous section of the test, renditions that are close but not precisely on the mark may be considered "slips of the tongue." When this happens, raters should accept such scoring units as correct.

## [Section 12]

## Guidelines for Assessing Scoring Units as Correct or Incorrect

---

- 12.1 The interpretation of a scoring unit shall be considered **correct** if:
- It is listed as acceptable in the scoring dictionary (even if the rater disagrees);
  - The meaning is preserved within the limits of the target language (e.g., the interpreter has used a linguistically appropriate way to preserve the purpose of the scoring unit some way other than what is listed in the scoring dictionary and which the test writers may not have anticipated);
  - Research conducted by the raters establishes that the word or term appears in a published, reputable reference with the intended meaning and the raters have come to consensus that it is correct; or
  - There is something on the recording that is not the fault of the examinee that prevents the rater from being reasonably certain what was said at all (e.g., the quality of the recording makes it impossible to hear and the problem is not caused by the examinee's low voice).
- 12.2 The interpretation of a scoring unit shall be considered **incorrect** if:
- The word or phrase has been designated as unacceptable in the scoring dictionary (even if the rater disagrees);
  - It is omitted when it should not have been omitted (e.g. "the **tall** man" is interpreted merely as "the man") and the candidate did not preserve the meaning and purpose of the scoring unit in some other way;
  - The meaning is changed or otherwise not preserved within the limits of the target language, e.g., "the **tall** man" is interpreted as "the **short** man;" or
  - The rater cannot determine what the examinee actually said because the examinee was mumbling, speaking inarticulately, or was otherwise unintelligible due to the examinee's own performance.
- 12.3 If a candidate's interpretation of a scoring unit is unfamiliar to a rater or raters and is not in the scoring dictionary, raters must check monolingual and bilingual dictionaries, thesauruses, and websites sponsored by official agencies (not private or individual sites

or publications), and official legal publications, whether online or not. If the term appears in an acceptable reputable reference work with the intended meaning, credit should be given. If no decision can be made, the benefit of the doubt shall be given to the candidate and it should be accepted. Raters should document the difficulty with the scoring unit on a *Scoring Unit Suggestion Form* (see Appendix 3) and return it to Consortium staff or the program manager along with the other test materials.

[Section 13]

## Mechanics of Scoring Exams as Correct or Incorrect and Compiling Scores

---

- 13.1 Raters shall write on each test script the identification number of each exam being rated before beginning the rating process.
- 13.2 When more than one section of an exam is to be rated, raters shall rate exams in the following sequence: English-to-Other Language Sight, Other Language-to-English Sight, Consecutive, and Simultaneous.
- 13.3 When doing the preliminary rating individually, each rater shall listen to the entire test straight through one time, marking the scoring units according to the standards below.
- 13.4 While rating exams, raters shall use a uniform marking system. There are certain kinds of marks that all raters must use consistently and there are others that may be used according to each rater's individual preferences. Raters shall refrain from making any other marks, except as provided elsewhere in §13. A sample marked test script to illustrate such markings is provided in Appendix 6.
  - Required uniform markings for all raters to employ:
    - A large “X” shall be written through the bolded, underlined text of each scoring unit determined to be incorrect, regardless of whether the scoring unit was misinterpreted, left out, or deemed to be wrong for any other reason. Only one “X” should be written on the text for any scoring unit. Small marks of any kind over or near the superscript identifiers of the scoring unit (i.e., the number and letter) are not acceptable.
    - Parentheses must be placed around words and phrases that the candidate omitted. Please note that if there is a scoring unit between a set of such parentheses, it must be marked wrong.
    - When any rater has initially marked a large “X” through a scoring unit to indicate that it is incorrect and subsequently determines it is correct, the rater shall write a clear “OK” above that scoring unit and cross out the “X” as

much as possible. This can happen either as a result of further deliberation by an individual rater or the consensus process.

- Discretionary markings that are optional and which raters may or may not choose to use:
  - A checkmark (✓) may be written for scoring units interpreted correctly above, on, or near the text of the scoring unit.
  - A small question mark (“?”) may be placed on or near a scoring unit if the rater is initially not sure how to assess it and a large question mark may be written in either margin of the line on which that scoring unit occurs. The purpose of the small mark is to indicate the specific scoring unit that must be reviewed later and the purpose of the larger question mark is to make it easy to locate such scoring units and ensure that no such scoring units are overlooked. Raters may need to mark scoring units for further review due to a number of reasons. Some are attributable to the candidate (e.g., mumbling or inarticulate speech, a burst of rapid speech, and convoluted renditions). Sometimes it is due to mechanical problems such as a change in volume of the recording or a recording that is not picking up the candidate’s voice clearly. In addition, it is sometimes the rater who has experienced a momentary distraction or lost focus.
  - Once the rater has returned to each such scoring unit and made a determination, the question mark should be struck out to indicate that the scoring unit has been reviewed and a final determination has been reached.
  - As an aid to count the number of incorrect scoring units on a given page of test text, raters may write in either margin the number of incorrect scoring units on lines where there are one or more incorrect scoring units, e.g., -1, -2, -3.

13.5 Raters are encouraged to write what they heard candidates say for certain scoring units or notes about incorrect or doubtful interpretations near the scoring unit or in the margins. New raters may find it difficult to write such notes until they have accumulated a considerable amount of experience. Experienced raters find this to be very useful, especially when going through the consensus process, because it reduces the need to replay segments of the exam to determine what the candidate actually said.

13.6 Once the rating process has been completed, the raters shall complete the scoring grid provided at the end of each test script. Each scoring grid has three components: the number of scoring units in the immediately preceding part or section of the exam (“*Possible Points*”), a box to write the number of scoring units which were assessed by the raters as being incorrect (“(-) # *Incorrect*”), and a box to write the number of scoring units assessed as being interpreted correctly (“*Total Correct*”). This is illustrated by the following grid from the simultaneous section:

Possible Points	75
(-) # Incorrect	
Total Correct	

13.7 Each scoring grid shall be completed as follows:

- Each part of the sight section:
  - Total the number of incorrect scoring units by counting the X'd scoring units on the page and write that number in the right box on row two of the grid.
  - Subtract row two from row one, which results in the number of correct scoring units for that test section, and write that number in the right box on row three.
- Consecutive and Simultaneous sections:
  - Count the number of incorrect scoring units on page one and record that number at the bottom of page one.
  - Carry that number over to the bottom of page two, followed by a slash mark (/).
  - Count the number of incorrect scoring units on page two and write that number to the right of the slash mark.
  - Add the two numbers at the bottom of page two and write the sum at the bottom of page 3, followed by a slash mark.
  - Count the number of incorrect scoring units on page three and write that number to the right of the slash mark.
  - Add the numbers to the left and the right of the slash mark.
  - If there are four or more pages in a section, continue with the same methodology.
  - The result on the last page is the sum of incorrect scoring units for the section.
  - Write that sum in the right box on row two of the grid.
  - Subtract row two from row one and write the total number of correct scoring units in the right box on row three.
- To ensure that the count of incorrect scoring units is accurate, count all of the X'd scoring units in each part or section again. If this count is different from the initial one, a third count is required to be absolutely certain. When a number originally written needs to be rewritten, make sure the final number is written clearly and the original number is either erased or clearly struck through.

13.8 When all scoring grids have been filled in, one of the raters must complete the electronic *Test Results Report Form* in Excel (see Appendices 4 and 5). For each section of the exam, enter the total number of correctly interpreted scoring units after “#

Correct” in the corresponding cells for each test section. The totals and percentage scores will compute automatically.

[Section 14]

## Scoring Unit Suggestion Forms and Rater Research

---

- 14.1 When raters are rating examinations, they will hear candidates’ renditions of the various scoring units and make determinations about whether the rendition of each scoring unit was acceptable (and therefore correct) or unacceptable (and therefore incorrect). When a rendition is heard that does not appear in the scoring dictionary for that examination, and, in the opinion of the raters, is a rendition that other candidates may use in the future, rater(s) should mark the scoring unit according to their determination of acceptability or unacceptability, and complete a *Scoring Unit Suggestion Form* (see Appendix 3). The rater(s) must include information about why he/she or both believe the scoring dictionary should be updated or revised (e.g., the rendition appears in a reputable, published dictionary).<sup>7</sup>
- 14.2 A team of two experts designated by Consortium staff pursuant to Section 14.1 of the *Test Construction Manual* will review the *Scoring Unit Suggestion Form(s)* and if the suggestion is obvious (i.e., the suggested rendition appears in a reputable, published dictionary as acceptable or unacceptable), will immediately update the scoring dictionary.
- 14.3 If the experts reviewing the suggested revisions disagree with the suggestion or are uncertain of the accuracy of the suggestion, they may confer with other raters or test writers, update the scoring dictionary, and notify the rater(s) who made the suggestion.
- 14.4 When a rater is working alone, he/she should conduct research, as described below, for any rendition of a scoring unit about which the rater is unsure. When two raters are working together, they should identify and discuss those scoring units on which they disagreed during the initial scoring process, conduct the research necessary to conclude that the interpretation was acceptable or unacceptable, and report one consensus score.
- 14.5 “Research” in this context means the following:

---

<sup>7</sup> When in doubt as to whether a rendering that is not in the scoring dictionary should be considered for inclusion in the dictionary as acceptable or unacceptable, the basic guiding principle is whether **a recent arrival from any part of the world where the language being tested is from** would be likely to understand it. If a recent arrival would understand it, then it should be considered for addition to the renderings that are deemed acceptable. If a recent arrival would not understand it, then it should be considered for addition to the unacceptable renderings.

- The rater(s) will first check the scoring dictionary.
- If the candidate’s rendering is there as an acceptable rendering, the rater(s) shall accept it as correct. If the candidate’s rendering is listed as an unacceptable rendering, it must be accepted as incorrect and the rater(s) shall mark the script as prescribed above. If the rater(s) disagree with the scoring dictionary, the rater(s) should check monolingual and bilingual dictionaries, thesauruses, and web sites sponsored by official (not private or individual sites or publications) agencies, and official legal publications, whether online or not. Then the rater(s) should send a completed *Scoring Unit Suggestion Form* to Consortium staff, providing supporting documentation from the reference source(s) supporting a change in the scoring dictionary.
- If the candidate’s rendering is not in the scoring dictionary as either acceptable or unacceptable, the rater(s) should check monolingual and bilingual dictionaries, thesauruses, and web sites sponsored by official (not private or individual sites or publications) agencies, and official legal publications, whether online or not. If the rendering is supported by one or more of these sources, the rater (s) shall accept the rendering is an accurate interpretation. If the rendering is not supported by one or more of these sources, the rater(s) shall mark it as incorrect. In either case, the rater(s) shall complete and submit a *Scoring Unit Suggestion Form*.
- If the research does not resolve how to assess a given rendition and there is still uncertainty or disagreement, the rater(s) should document the uncertainty or disagreement on a *Scoring Unit Suggestion Form* and return it to Consortium staff or the program manager along with the other test materials. Raters should assess the scoring unit as correct unless giving credit for the scoring unit will move the score from a fail status to a pass or a pass to a fail status. If that is the case, the rater(s) should notify Consortium staff and staff will attempt to quickly ascertain, with help from other experts, how the rendition should be scored.

14.6 When raters discover that they have a difference of opinion with other experts who contributed to the scoring dictionary, the raters should first ask themselves what reasons their colleagues might have for taking the position they did. It has been found that raters’ opinions vary from one context to another and taking a wider view often makes a difference in arriving at a consensus.

14.7 When raters consistently experience disagreements about how to score renderings of a specific scoring unit, it probably means that the unit should be replaced. The discovery of hidden problems with units only occurs through test use. These units should be changed as soon as is practicable, and it is very important that this information be communicated using the *Scoring Unit Suggestion Form*.

- 15.1 While the primary function of raters is to rate exams, they have an important secondary function in terms of providing ongoing feedback on how well the exams are functioning. The *Test Construction Manual* states, “As long as any test form is being administered, raters who grade the test will be engaging in ongoing review and evaluation of that test. For example, raters sometimes discover that scoring units do not perform well over time or identify shortcomings in the selection and/or classification of scoring units...” (see Section 13.2).
- 15.2 Accordingly, while rating exams, raters shall watch for problems with each test form being rated and provide feedback when appropriate. Any time a rater or rating team has any questions, concerns, or suggestions for improving any aspect of any exam, they are urged to write a description of the nature of the problem or concern, offer one or more solutions to each such problem or concern if they can, provide any other information or perspectives that may be useful, and submit a report to Consortium staff directly (if rating pursuant to a member contract with Consortium staff) or indirectly (through a program manager). All such feedback will be reviewed carefully by the test writers and Consortium staff and, when deemed appropriate, will prompt revisions in test scripts. In the event the rater or rating team was also the test writing team, Consortium staff will work directly with them to reach an appropriate disposition of the report.
- 15.3 Consortium staff will provide feedback to each rater who provides such reports on the ultimate resolution of their suggestions and keep the Technical Committee apprised of all such transactions.
- 15.4 Raters also have the function of monitoring performance of individuals who proctor the exams. Any time raters discover that proctor error has created a problem that interferes with the fairness of the test-taking experience or creates impediments for the rating process, raters shall write a report in which each problem is described objectively and the corresponding impact on the integrity of the test administration and rating process is clearly set forth. These reports shall be submitted directly to Consortium staff, who will review such reports and share the results of their analysis of the program manager of the member jurisdiction, including any suggestions for remedial action that may be warranted.
- 15.5 When there are mechanical or proctor problems that render all or part of a test difficult or impossible to rate, the rater or rating team shall immediately contact Consortium staff for guidance on how to proceed if the Rating Supervisor is not available to provide adequate guidance.

- 16.1 Persons who have written Consortium exams and wish to become raters must complete the rater training program established by the Technical Committee and coordinated by Consortium staff *before* they may rate any exams, except as provided for members or test writing teams in §4.3 of the Test Construction Manual.
- 16.2 All other persons who wish to become raters must complete the rater training program established by the Technical Committee and coordinated by Consortium staff *before* they may rate any exams.
- 16.3 In addition to the initial rater training, raters shall participate in periodic “refresher” training events and any other training the Technical Committee may determine from time to time. If possible, the mandatory periodic events will be a remote, interactive training module that can be used for all raters in all languages. No substitutions for this training are allowed without approval from Consortium staff.

- 17.1 Raters’ performance will be supervised by Consortium staff, rating supervisors, and, to a lesser degree, program managers, to ensure that raters perform their duties in a manner consistent with the provisions of this Manual. This essential aspect of quality control includes but is not limited to the following:
- Test scripts will occasionally be reviewed to ensure that rating mechanics are being followed, scores are being accurately tabulated, and forms are being filled out correctly.
  - Performance during the rating process will be monitored by rating supervisors who will report any concerns to Consortium staff.
  - Timeliness of completing rating assignments and working collegially with other raters will be monitored by Consortium staff and program managers.
  - A record of all such concerns will be maintained by Consortium staff, who will also address such concerns in a forthright, professional manner directly with each rater for whom such matters arise.

- 18.1 Any program manager who has evidence that a rater has not abided by the provisions of this Manual may file a written complaint specifying the details of the alleged misconduct to Consortium staff. Raters may file complaints about raters with whom they have rated and shall file such complaints with the office that arranged for the rating process (i.e., the program manager or Consortium staff, depending on the circumstances).
- 18.2 Complaints may be filed for the following types of misconduct:
- Willful or negligent compromising of the security or confidentiality of any test materials.
  - Use of any test content in a training event or the promise of an increase in exam performance as a result of any training.
  - Discussion of the test performance of any examinee.
  - Refusal to follow and abide by scoring dictionaries and scoring candidates on the basis of one's own preferences, or declining to do the necessary research and documentation to support deviations from the scoring dictionaries.
  - Intentional scoring of an examinee's performance unfairly or differently from other examinees for any reason.
  - Failure to complete required training.
  - Performance of duties in an untimely manner or in a manner inconsistent with the provisions of this Manual (e.g., working ineffectively with another rater or declining to accept guidance from Approved Lead Raters or others guiding the rating process), failing to use appropriate scoring mechanics (including marking scripts), completing *Test Results Report Forms* incorrectly, and submitting *Scoring Unit Suggestion Forms* without sufficient research or failing to submit such forms when appropriate.
- 18.3 Consortium staff, under the oversight of the Technical Committee, will review and process all complaints filed against raters. Every effort will be made by Consortium staff and, when applicable, the Technical Committee, to provide relevant and objective feedback and guidance when any kind of problem arises with raters, when raters show any reluctance to follow Consortium policies and procedures, or when raters engage in conduct that compromises the security and confidentiality of exam contents, privacy of candidates, or the validity and reliability of the rating process.
- 18.4 A range of disciplinary actions may be imposed by Consortium staff upon approval by the Technical Committee, including either temporary (and, if temporary, for what specific period of time) or permanent removal from the Rater Contact List,

- 18.5 All information pertaining to specific complaints against raters shall be confidential and available only to Consortium staff, members of the Technical Committee, and, when deemed to be appropriate, program managers.

# Appendices

---



The Consortium for Language Access in the Courts discourages member states from using interpreter training providers who also serve as Consortium test writers or raters unless such activities are fully disclosed. In its *Court Interpreter Oral Examination: Test Rating Manual*, the Consortium states:

*“There are occasions when a test rater also serves as an interpreter trainer, e.g., delivering training at institutions of higher education, orientation programs, skill-building seminars, or professional conferences. Although serving as a trainer does not automatically disqualify an individual from serving as a rater, such information must be disclosed to avoid the appearance or perception of a conflict of interest.” (Section 8.3)*

\_\_\_\_\_, hereinafter referred to as Training Provider, agrees to provide skill, ethical, or other training to court interpreters in the State of \_\_\_\_\_, a member of the Consortium for Language Access in the Courts, and (check one):

- Training Provider hereby discloses that he/she also serves as a test writer or test rater of Consortium examinations and agrees to sign a disclosure agreement.
- Training Provider hereby certifies that he/she does not serve as a test writer or test rater of Consortium examinations.

\_\_\_\_\_  
Signature of Training Provider

\_\_\_\_\_  
Printed Name of Training Provider

\_\_\_\_\_  
Date

(This signed form should be turned in to the program manager in the state where the training will be provided.)

I, \_\_\_\_\_, Training Provider, hereby disclose that I serve as a test writer or test rater for the Consortium for Language Access in the Courts and also serve as a trainer of court interpreters.

I agree that I will not use or disclose test content in any way in my training materials or presentation content. I further agree that I will not market my training session(s) as Consortium oral examination preparation training nor advertise an expected or guaranteed increase in the likelihood of passing a Consortium oral examination.

\_\_\_\_\_  
Signature of Training Provider

\_\_\_\_\_  
Printed Name of Training Provider

\_\_\_\_\_  
Date

---

(This signed form should be turned in to the program manager in the state where the training will be provided.)

TEST NUMBER \_\_\_\_\_

RATER NUMBER \_\_\_\_\_

**Please note that suggestions will not be considered if there is no supporting evidence or resource provided.**

<input type="checkbox"/> English Sight <input type="checkbox"/> Foreign Sight <input type="checkbox"/> Consecutive <input type="checkbox"/> Simultaneous  <input type="checkbox"/> Add to acceptables <input type="checkbox"/> Add to unacceptables <input type="checkbox"/> Other comment / problem <b>This box must be completed</b>	UNIT #:	UNIT TEXT:
	Supporting evidence or resource _____	

<input type="checkbox"/> English Sight <input type="checkbox"/> Foreign Sight <input type="checkbox"/> Consecutive <input type="checkbox"/> Simultaneous  <input type="checkbox"/> Add to acceptables <input type="checkbox"/> Add to unacceptables <input type="checkbox"/> Other comment / problem <b>This box must be completed</b>	UNIT #:	UNIT TEXT:
	Supporting evidence or resource _____	

<input type="checkbox"/> English Sight <input type="checkbox"/> Foreign Sight <input type="checkbox"/> Consecutive <input type="checkbox"/> Simultaneous  <input type="checkbox"/> Add to acceptables <input type="checkbox"/> Add to unacceptables <input type="checkbox"/> Other comment / problem <b>This box must be completed</b>	UNIT #:	UNIT TEXT:
	Supporting evidence or resource _____	

**COURT INTERPRETER ORAL EXAMINATION  
TEST RESULTS REPORT FORM**

Consortium for Language Access in the Courts

**Candidate Information**

<b>Test Date:</b>	<b>Language:</b>
<b>NCSC ID:</b>	
<b>State ID:</b>	

**Oral Examination Test Results**

<b>Evaluator #1:</b>	<b>Evaluator #2:</b>
----------------------	----------------------

<b>SIGHT TRANSLATION</b>		<b>(25 units possible for each part)</b>	
English - Foreign Language	# Correct		0%
Foreign Language - English	# Correct		0%
Total Number Correct	<b>0</b>		<b>0%</b>
<b>CONSECUTIVE</b>		<b>(75 units possible)</b>	
Total Number Correct			<b>0%</b>
<b>SIMULTANEOUS</b>		<b>(75 units possible)</b>	
Total Number Correct			<b>0%</b>

Feedback about common exam performance deficiencies can be accessed at:  
[http://www.ncsc.org/Web%20Documents/Consortium\\_CommonErrors.pdf](http://www.ncsc.org/Web%20Documents/Consortium_CommonErrors.pdf)

**COURT INTERPRETER ORAL EXAMINATION  
TEST RESULTS REPORT FORM**

Consortium for Language Access in the Courts

**Candidate Information**

<b>Test Date:</b>	<b>Language:</b>
<b>NCSC ID:</b>	
<b>State ID:</b>	

**Oral Examination Test Results**

<b>Evaluator #1:</b>	<b>Evaluator #2:</b>
----------------------	----------------------

<b>SIGHT TRANSLATION</b>		<b>(25 units possible for each part)</b>	
English - Foreign Language	# Correct		0%
Foreign Language - English	# Correct		0%
Total Number Correct	<b>0</b>		<b>0%</b>
<b>CONSECUTIVE</b>		<b>(90 units possible)</b>	
Total Number Correct			<b>0%</b>
<b>SIMULTANEOUS</b>		<b>(75 units possible)</b>	
Total Number Correct			<b>0%</b>

Feedback about common exam performance deficiencies can be accessed at:  
[http://www.ncsc.org/Web%20Documents/Consortium\\_CommonErrors.pdf](http://www.ncsc.org/Web%20Documents/Consortium_CommonErrors.pdf)

## Practice Examination #1

## Consecutive

## Corporal Punishment

Q. Mr. Rodriguez, my name is Barbara Davies. I spoke with you on the telephone on April 16<sup>th</sup>G. Do you remember that conversation?

A. O sea<sup>21</sup> el día que me llamó al trabajo... sí, creo que sí.

Q. We talked about your work history and your home and about your children (Pedrito and Celia, and Paula<sup>3G</sup>, their mother? (Do you recall that? )

A. Sí, hablamos un buen rato. Recuerdo especialmente que hablamos acerca del problema de Pedro, mi hijo.

Q. (You mentioned) you've worked (a long time<sup>4B</sup>) for a hotel, I think it's the Paradise Island Hotel, right?

A. Antes se llamaba así. Le cambiaron el nombre hace dos, dos años y medio<sup>5G</sup>. Ahora se llama Hotel Crystal Palace, pero es el mismo, (correcto.)

Q. Now you work as a manager for that property, but you have held<sup>6A</sup> several different positions<sup>7B</sup> in the hotel and have a very strict work ethic<sup>8C</sup>, do you not? You always arrive at work on time and you're devoted<sup>9A</sup> to your work?

A. Así es, me han tratado muy bien. Yo empecé lavando platos, pero siempre he sido muy emprendedor<sup>10C</sup> y gracias a mi esfuerzo llegué a la gerencia<sup>11C</sup>. Yo soy leal<sup>12C</sup> a mis patrones y por eso que me han dado varios<sup>13I</sup> ascensos.

Q. So now, after you worked your way up the ranks to where you are today, you make very good money<sup>14E</sup>, right?

A. Yo no diría<sup>15A</sup> que gano mucho dinero, pero me va bien<sup>16H</sup>.

Q. And, correct me if I'm wrong, you live in a house located at 2491 Bayfront<sup>17G</sup> in one of the swankiest<sup>18J</sup> neighborhoods on the island?



