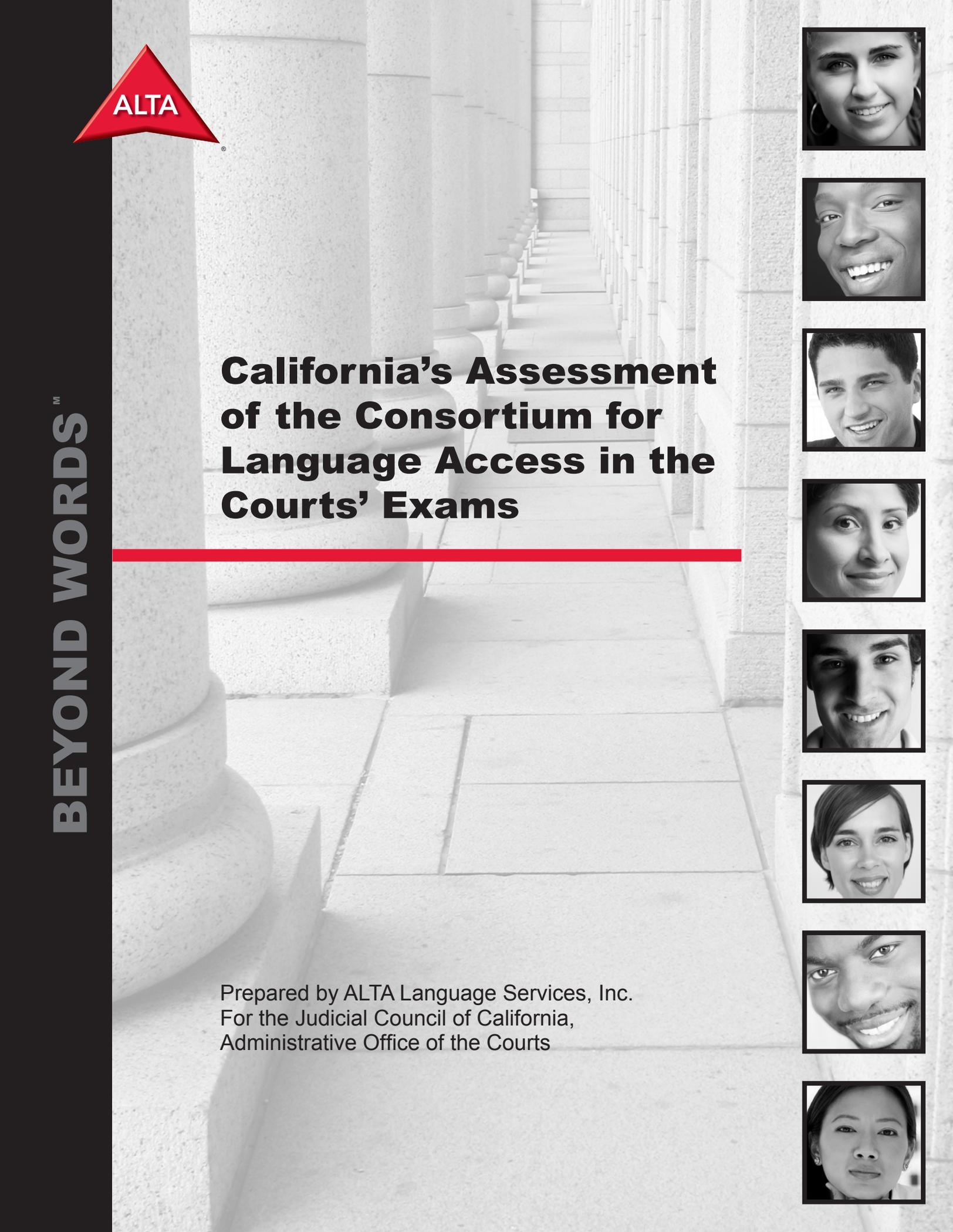




ALTA

BEYOND WORDS™



California's Assessment of the Consortium for Language Access in the Courts' Exams



Prepared by ALTA Language Services, Inc.
For the Judicial Council of California,
Administrative Office of the Courts

This report is produced for the Judicial Council of California, Administrative Office of the Courts,
455 Golden Gate Avenue, San Francisco, CA 94102-3688, www.courtinfo.ca.gov

Copyright© 2010 by Judicial Council of California/Administrative Office of the Courts. All rights reserved.

Prepared for the Judicial Council of California by ALTA Language Services, Inc., 3355 Lenox Road,
Suite 510, Atlanta, GA 30326

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, online, or mechanical, including the use of information storage and retrieval systems, without permission in writing from the copyright holder. Permission is hereby granted to nonprofit institutions to reproduce and distribute this publication for educational purposes if the copies credit the copyright holder.

For additional information, please contact:

Lucy Smallsreed, Manager
Court Interpreters Program,
Executive Office Programs Division,
Administrative Office of the Courts
455 Golden Gate Avenue
San Francisco, CA 94102-3688
courtinterpreters@jud.ca.gov

This report is also available on the California Courts Web site: www.courtinfo.ca.gov/programs/courtinterpreters/

Printed on 100% recycled and recyclable paper.

California's Assessment of the Consortium for Language Access in the Courts' Exams

Prepared by ALTA Language Services, Inc.
For the Judicial Council of California,
Administrative Office of the Courts



FORWARD

In May of 2009, the Judicial Council of California, Administrative Office of the Courts (AOC) contracted ALTA Language Services, Inc. (ALTA) to assess the Consortium for Language Access in the Courts' (CLAC or Consortium) examinations for certifying member state court interpreters.

The purpose of this assessment was to determine the level of functional equivalency between California's court interpreter certification examinations and the Consortium's exams, and to determine how California could use Consortium exams if comparable testing standards were established.

ALTA was charged with analyzing Consortium test content and identifying the overall strengths and weaknesses of the Consortium testing program. Additionally, ALTA examined the degree to which the knowledge, skills, and abilities required of a California court interpreter are covered by the Consortium exams.

With the full cooperation of Consortium staff, ALTA conducted a rigorous and thorough assessment of the Consortium's exams for certifying court interpreters as well as a comprehensive comparative analysis of CLAC's and CCIP's exam programs. ALTA has found that the CCIP and CLAC exams are comparable in structure, content, and level of difficulty, and that both programs certify interpreters at the same high standards. The results of this study provide a strong basis for California to adopt Consortium exams in the future.

Sincerely,



Chris Roosevelt, Ph.D.
Vice-President of Testing and Training
ALTA Language Services, Inc.

ALTA Language Services, Inc.
3355 Lenox Road, Suite 510 Tel 404.920.3800
Atlanta, GA 30326 Fax 404.920.3801
www.altalang.com Email info@altalang.com



TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
INTRODUCTION: PROJECT SCOPE AND DELIVERABLES	7
1 ASSESSMENT OF THE CONSORTIUM FOR LANGUAGE ACCESS IN THE COURTS' TEST DEVELOPMENT PROCESSES.....	11
1.1 Methodological Approach of the Assessment of the Consortium Test Development Processes.....	12
1.2 Findings of the Assessment of the Consortium Test Development Processes	12
1.3 Analysis of Findings of the Assessment of the Consortium Test Development Processes ...	19
1.4 Conclusions Regarding the Assessment of the Consortium Test Development Processes	19
1.5 Recommendations from the Assessment of the Consortium Test Development Processes....	19
2 COMPARATIVE ANALYSIS OF THE CALIFORNIA COURT INTERPRETER PROGRAM EXAMS AND THE CONSORTIUM FOR LANGUAGE ACCESS IN THE COURTS' EXAMS	21
2.1 Methodological Approach of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams	22
2.2 Findings of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams	22
2.3 Analysis of Findings of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams	27
2.4 Conclusions Regarding the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams	29
2.5 Recommendations from the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams	30
APPENDICES	31
Appendix 1 – List of Acronyms.....	32
Appendix 2 – Glossary of Testing Terminology	33
Appendix 3 – Interview Questions	34
Appendix 4 – KSAs Essential for Court Interpretation	43
Appendix 5 – KSAs Measured by the California Written Exam	44
Appendix 6 – KSAs Measured by the Consortium Written Exam	45
Appendix 7 – KSAs Measured by the California Oral Exam.....	46
Appendix 8 – KSAs Measured by the Consortium Oral Exam.....	47
Appendix 9 – Brief History of the Federal Court Interpreter Certification Exam	48
Appendix 10 – Comparison between the Federal Court Interpreter Certification Oral Exams and the Consortium's Oral Exams	50
Appendix 11 – Description of Consortium Scoring Unit Categories.....	52
Appendix 12 – Standard Distribution of Scoring Units for Consortium Oral Exams.....	54
Appendix 13 – California and Consortium Written Certification Exam Comparison	55
Appendix 14 – California and Consortium Oral Certification Exam Comparison	56
Appendix 15 – Description of Recommended Standard-Setting Approaches.....	58
Appendix 16 – Literature Review	59

EXECUTIVE SUMMARY

I. Purpose of Study

The Judicial Council of California, Administrative Office of the Courts (AOC) contracted ALTA Language Services, Inc. (ALTA) to assess the Consortium for Language Access in the Courts' (CLAC or Consortium) examinations for certifying member state court interpreters. This study was conducted to establish the level of functional equivalency between California's court interpreter certification assessments and the Consortium's assessments, as well as to develop recommendations for standard-setting of the Consortium assessments that would ensure the same passing level of competency required of a candidate to pass the California assessments. (For a complete list and explanation of acronyms referenced throughout this Report, see Appendix 1.) Results of analyses from the study also were to establish whether the Consortium's interpreter certification assessments met the standards required of California interpreter certification assessments and, if not, what improvements could be made to achieve those standards.

II. Background

The Consortium for Language Access in the Courts has created a comprehensive interpreter certification testing program that encompasses both an English-only written examination and a bilingual oral performance examination, modeled after the Federal Court Interpreter Certification Examination (FCICE), in more than a dozen languages. The California Court Interpreter Program (CCIP) also consists of a written screening examination and an oral performance examination for certified court interpreters.¹ Both written

“Oral performance examinations for both CCIP and CLAC test the three modes of interpreting required of a court interpreter: the simultaneous mode, the consecutive mode, and sight translation.”

examinations score candidates based upon their ability to correctly answer a variety of multiple-choice questions. Oral performance examinations for both CCIP and CLAC test the three modes of interpreting required of a court interpreter: the simultaneous mode, the consecutive mode, and sight translation.² Additionally, oral test script development for each program is very similar. (For a detailed chart comparing each program's oral exam structure and content, see Appendix 14.) For oral exams, California uses a combined scoring method that includes a holistic evaluation of language and interpreting skills in addition to objective scoring units, while the Consortium strictly uses objective scoring units.³

III. Methodology

Various methodological steps were taken to examine the process used by the Consortium to develop its testing tools, and to present a comparative assessment of the exams used for certifying court interpreters by both CCIP and CLAC. These steps included the following:

- A. Literature and Document Review:** Literature pertaining to industry standards for valid test development was reviewed in addition to documentation provided by Consortium staff and documentation found on the Consortium's website (see Appendix 16).
- B. Qualitative Interviews:** ALTA conducted interviews with Consortium key staff members as well as former and current language consultants and exam writers.

1 Currently, California has certification examinations in 12 designated languages: Arabic, Eastern Armenian, Cantonese, Japanese, Korean, Mandarin, Portuguese, Russian, Spanish, Tagalog, Vietnamese, and Western Armenian. Interpreters of American Sign Language are eligible to become certified California court interpreters if they hold the Specialist Certificate: Legal granted by the Registry of Interpreters for the Deaf (RID). Non-designated languages are referred to as registered languages. Interpreters of registered languages must pass an English proficiency exam (both written and oral).

2 Simultaneous interpretation is the mode in which the interpreter lags slightly behind the source language speaker, interpreting the message into the target language at almost the same time as the original message is being said. In consecutive interpretation, the interpreter listens to a unit of speech (approximately 40-60 words at a time) in the source language and then conveys that message into the target language. Sight translation is the oral interpretation of a written document.

3 According to the Consortium's *Court Interpreter Oral Examination: Test Construction Manual (TCM)*, scoring units are linguistic phenomena that interpreters must be able to render to deliver a complete and accurate interpretation.

- C. **Federal Court Interpreter Test Development Review:** ALTA examined documentation regarding the test development process and structure of the FCICE in comparison to CLAC's test development process and structure. (For a brief history of the development of the Federal Exam, see Appendix 9.)
- D. **Test Content Review:** An in-depth assessment of CLAC's written and oral test content was conducted by a panel of SMEs.
- E. **Comparative Analysis of Test Programs:** Results from the *Study of California's Court Interpreter Certification and Registration Testing*⁴ (*Study 2007*) were compared to the test content review conducted of the Consortium exams; the extent to which the CCIP and CLAC test programs are similar and/or different was assessed.
- F. **Psychometric Audit:** A psychometric evaluation⁵ of the Consortium test content framework and standards was performed by a team of psychometricians and compared to current test industry standards.
- G. **Standard-Setting Session:** A standard-setting session was conducted by a panel of thirteen SMEs to establish the level of functional equivalency between the passing requirements of each program's oral examinations.

IV. Findings

Findings of the research conducted to assess the Consortium's test development process and determine how the Consortium's process compared to test industry standards include the following:

- The Consortium has created a comprehensive interpreter testing program with an impressive body of technical manuals;
- Consortium exam content has been constructed based on solid and consistent test development processes that respect test industry standards with a few areas in need of improvement;
- The Consortium has created and uses a *Court Interpreter Oral Examination: Test Construction Manual (TCM)*, which provides a transparent blueprint for Consortium oral exam development and promotes test equivalency across languages and test versions;
- Many aspects of the Consortium's oral exams were modeled after the Federal interpreter certification exams; and
- The Consortium employs highly qualified subject matter experts at every stage of the process of test development.



4 In June of 2006, the Administrative Office of the Courts of California contracted for an analysis of the examination process and testing instruments for the certification and registration of California state court interpreters. The results and final recommendations of this analysis were concluded in 2007, and the report based upon those results and recommendations is referred to as *Study 2007*.

5 A psychometric evaluation is conducted by testing experts to document and assess a testing program's test blueprints and specifications, as well as any available statistical analyses.

Findings from the comparative analysis of CCIP and CLAC exams conducted include:

- While the constructs of the written exams for each program are somewhat different, they both serve the same function at very similar levels of difficulty: eliminating unqualified candidates from moving on to the oral exam phase of the certification process, while allowing qualified candidates to proceed in the process;
- The structure, content, and level of difficulty of the oral exams for CCIP and CLAC are comparable and adhere to equivalent standards;
- One essential difference between CCIP and CLAC exams is the method of scoring used by each program for its oral examinations: CLAC uses only objective scoring units, while CCIP uses a combination of holistic and objective scoring methods;
- A high degree of overlap exists between CCIP and CLAC's written and oral exam KSAs; and
- All of the KSAs required in the area of Interpreting Skills are tested by both CCIP and CLAC.

V. Analysis

An analysis of findings determined that most areas of the Consortium interpreter certification program meet test industry standards, while a few are in need of improvement to meet test industry standards. Modeling the Consortium oral exams after the Federal oral exams, which have legally been shown to be valid (Arjona, 1985), provided the Consortium with a strong foundation for oral test development. The Consortium's development and use of the *TCM* and an objective scoring method has further facilitated the development of a consistent and transparent method of achieving test validity during each stage of test development for every language for which the Consortium develops an interpreter certification test.

“CLAC's key areas of strength lie in the qualifications of personnel involved in test development, its objective method of oral exam scoring, and its use of the *TCM*.”

CLAC's key areas of strength lie in the qualifications of personnel involved in test development, its objective method of oral exam scoring, and its use of the *TCM*. CLAC's key areas for improvement include the need for further collection of documentation and oral exam statistical data.

The comparative analysis of CCIP and CLAC examinations determined that CCIP and CLAC written and oral exams test candidates at comparable standards of difficulty and at similar standards according to test validity practices. The one KSA (the ability to preserve accuracy) tested by CCIP's written

exam that is not tested by CLAC's is tested in CLAC's oral examination. KSAs tested by CCIP but not tested by CLAC's oral exams were identified as speaking skills that are not tested due to CLAC's objective method of scoring. The use of objective scoring units does not allow for an evaluation of subjective components, such as the testing of a candidate's pronunciation or accent in the foreign language. In fact, the Consortium considers scoring elements such as foreign language accent difficult to test in an objective manner and therefore specifically avoids testing KSAs that cannot be measured through objective scoring. Poor speaking skills will, however, indirectly impact a candidate's score if raters cannot hear or understand a response given on the oral exam.

VI. Conclusions

ALTA's conclusions of the findings and analyses are that, since its inception in 1995, CLAC has created a comprehensive interpreter certification testing program with an impressive body of technical manuals and candidate resources. CCIP and CLAC have many comparable testing elements and both programs have created exams at similar standards and levels of equivalency. In particular, the oral exams of each program contain many common elements, including types of scripts used and the testing of grammatical structures and legal terminology. A high overlap of KSAs tested also exists between the CCIP and CLAC written and oral exams; the difference in KSAs is attributable to the distinct methods of scoring used by each program, but all interpreting KSAs adopted by California are tested in CLAC's exams.

While all testing programs require the continual need for review and maintenance and contain areas that could benefit from improvement, CLAC has created a testing program that is equivalent to California's in terms of validity and content. Results of the standard-setting session conducted for the CLAC oral exam provided further support that Consortium exams certify candidates at a standard comparable to that of California.

VII. Recommendations

It is ALTA's recommendation that California could use CLAC's written exam as a screening device to qualify certified and registered court interpreter candidates. This recommendation is based on CLAC's adherence to test industry standards when developing its written exam content and cut-score.

It is also recommended that since oral certification exams for both California and CLAC have been found to be comparable in structure, content, and level of difficulty, CLAC oral exams could also be used by CCIP to certify California court interpreters. Additional recommendations have been made in the full final report of this study to enhance CLAC's testing program, but are not considered prerequisites to California's use of the Consortium exams.

INTRODUCTION: PROJECT SCOPE AND DELIVERABLES

In May of 2009, the Judicial Council of California, Administrative Office of the Courts (AOC) contracted ALTA Language Services, Inc. (ALTA) to assess the examination development process for the Consortium for Language Access in the Courts' (CLAC or Consortium) certification of member state court interpreters. ALTA was charged with assessing the historical and current examination development process of the Consortium, so as to develop recommendations regarding efforts to establish functional equivalence between California's court interpreter certification assessments and the Consortium's assessments. Additionally, ALTA examined which knowledge, skills, and abilities (KSAs) are tested through CLAC's examination process, as well as the educational level needed to successfully complete the exam. A psychometric evaluation of the Consortium's test content and standards was also performed to ensure that testing industry standards were being met through the Consortium's test development process. In an attempt to further establish equivalence between CCIP (California Court Interpreter Program) and CLAC's testing programs, ALTA also addressed and developed recommendations pertaining to standard-setting



of the Consortium assessments that would ensure the same passing level of competency required of a candidate to pass the California assessments. (For a complete list and explanation of acronyms referenced throughout this Report, see Appendix 1.)

To produce a comprehensive report presenting specific recommendations concerning the assessment of the Consortium for Language Access in the Courts' state court certification examination process and functional equivalence between CLAC exams and CCIP exams, the following deliverable items were included in the scope of services:

- A. Assessment of the Consortium's Test Development Processes:** An assessment was conducted to determine the historical as well as the current test development practices of the Consortium for Language Access in the Courts. From this research, the processes used for test writing, item selection, pilot testing, psychometric evaluations, qualifications, training, and recruitment processes for exam writers and developers, adaptation of foreign language test components and the adaptation of non-Consortium exams to meet Consortium guidelines were analyzed and compared to current Consortium test practices.
- B. Comparative Analysis of CCIP Exams and Consortium Exams:** Written and oral Consortium exam structure and content was analyzed, and the knowledge, skills, and abilities measured by the exams were identified. The extent to which the Consortium exams assess the 26 measurable California KSAs, as determined through the *Study of California's Court Interpreter Certification and Registration Testing (Study 2007)*, was also established. A psychometric evaluation of the blueprint/template of the Consortium's test content and standards was also performed during this stage, as well as a comparative analysis of test industry standards with the Consortium's documented test purpose, test development procedures, and testing analysis. Additionally, recommendations were developed regarding standard-setting to ensure functional equivalence between the Consortium and California's testing programs.
- C. Final Report and Recommendations:** A final comprehensive report has been created that addresses each of the Deliverables. An executive summary, methodology, findings, analyses, conclusions, and recommendations have been included.

In this final Deliverable, ALTA was asked to present a comprehensive report that addressed the major findings, analysis, conclusions, and recommendations from the Assessment of the Consortium's Test Development Processes and the Comparative Analysis of CCIP Exams and CLAC Exams.

Project Methodology

Various methodological steps were taken to present a final report on the test development processes used by the Consortium, including the conducting of a psychometric evaluation of current Consortium test practices. Steps were also taken to present an assessment of the exams used for certifying court interpreters by CCIP and CLAC, including a standard-setting recommendation to ensure equivalency between the two programs. These steps consisted of the following:

- A. Literature and Document Review:** Throughout the project, literature and documents pertaining to industry standards for valid test development, test adaptation⁶ practices, and standard-setting practices were reviewed (see Appendix 16). ALTA also reviewed all publicly available information found on the Consortium’s website, in addition to confidential documentation provided by Consortium staff. This review contributed to an understanding of the Consortium’s test development processes, as well as to an understanding of the steps needed to establish functional equivalence between CCIP and CLAC exams.
- B. Qualitative Interviews:** To gather detailed qualitative information regarding the test development processes used for both the written and oral Consortium exams, ALTA conducted interviews with both former and current Consortium key staff members, including subject matter experts⁷ (SMEs) involved in the original test development process. (See Appendix 3 for interview questions.) By examining historical as well as current test development processes, ALTA was able to identify how the various aspects of the Consortium’s testing program have been accomplished, and how those aspects align with testing industry standards.

Former and current language consultants and exam writers were also interviewed to provide further information regarding training, creating test content, and the process of adapting foreign language components of the oral exams, as well as the process of choosing scoring units, and other areas of interest pertaining to Consortium test development practices

- C. Federal Court Interpreter Test Development Review:** ALTA examined documentation regarding the test development process and structure of the Federal Court Interpreter Certification Examination. (For a brief history of the development of the Federal Exam, see Appendix 9.) This review was conducted based on information provided during the qualitative interview process that several aspects of the Consortium’s test development program were modeled after the Federal court interpreter exam program.
- D. Test Content Review:** An in-depth assessment of CLAC’s written and oral test content was conducted. A panel of SMEs was identified to perform the review of Consortium test content and asked to determine the extent to which CLAC’s current testing instruments aligned with the 26 measurable KSAs needed for court interpreting as identified in *Study 2007*. (See Appendix 4 for a list of the KSAs from *Study 2007*.)

ALTA staff also reviewed the scoring unit distributions of each section of the Consortium oral exams as outlined in the *Court Interpreter Oral Examination: Test Construction Manual*⁸ (Consortium for Language Access in the Courts, *TCM*, 2010). The outlined scoring unit category distributions were compared to distributions found in actual test content and analyzed for any discrepancies that might have occurred during the development process. (See Appendix 12 for the scoring unit distribution chart.) The distribution of scoring unit categories also served to enhance results of SME test content reviews.

⁶ Test adaptation is the process used to adapt an exam from one language to another; this involves the consideration of the cultural nuances and grammatical structures of each language.

⁷ Staff members and SME contractors were required to sign confidentiality agreements precluding them from taking a Consortium exam for a minimum of five years, as well as from disseminating Consortium test content.

⁸ The *TCM* serves as a blueprint for oral exams developed by the Consortium, outlining the qualifications of the test development committee (TDC), the test development process, the creation and maintenance of scoring units, and administrative guidelines for pilot testing.

- E. Comparative Analysis of Test Programs:** Results from *Study 2007* formed the basis for understanding California's test development process and content. These results were compared to the test development process review and test content review conducted of the Consortium exams, and the extent to which the two programs are similar and/or different was assessed. This analysis contributed to the determination of the functional equivalence between the two programs, as well as to a determination of how California could best utilize Consortium test content for its own court interpreter certification program.
- F. Psychometric Audit:** A psychometric evaluation of the Consortium test content framework and standards was performed by a team of psychometricians. The psychometric team reviewed all Consortium documents including collected statistical analyses, as well as interview transcripts from the qualitative interviews conducted with Consortium staff and consultants.

Information collected during the psychometric evaluation was then compared to test industry standards and how these standards applied to the Consortium's testing program, as well as the degree to which the Consortium currently conforms to those standards.

- G. Standard-Setting Session:** A standard-setting session was conducted by a panel of thirteen SMEs to establish the level of functional equivalency between each program's oral examinations. SMEs reviewed Consortium test content and determined whether the standard required for a candidate to pass Consortium oral exams was equivalent to that of California.



ASSESSMENT OF THE CONSORTIUM FOR LANGUAGE ACCESS IN THE COURTS' TEST DEVELOPMENT PROCESSES



An assessment was conducted of the test development processes used by CLAC including a description of its test writing process, test item selection process, pilot-testing process, and any psychometric evaluations performed. An analysis was performed of the adaptation of foreign language exam components and donated non-Consortium exams, and a description was written concerning the qualifications, training, and recruitment processes for exam writers and developers.

Through this comprehensive assessment, ALTA was able to assess the Consortium's historical and current test development processes and determine how these processes compared to current test industry standards as defined by the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999).⁹

1.1 Methodological Approach of the Assessment of the Consortium Test Development Processes

To assess the test development processes used by the Consortium, a multi-step analysis was employed which included qualitative interviews with Consortium key personnel and language consultants, a review of the Federal court interpreter test development process, a psychometric evaluation, and an extensive review of documentation and literature pertaining to historic and current Consortium test development processes. (For further explanation of each methodological approach taken, see the Project Methodology on pages 9 – 10 of the Introductory section of this Report.)

1.2 Findings of the Assessment of the Consortium Test Development Processes

The Consortium for Language Access in the Courts (formerly known as the Consortium for State Court Interpreter Certification) was founded in 1995 by four original member states — Minnesota, Oregon, Washington, and New Jersey — to address the need for a national testing body in the field of court interpretation. These four states pooled their resources to develop a court interpreting certification program closely modeled after research conducted on the FCICE, CCIP, the state of New Jersey's court interpreter program, and the state of Washington's court interpreter program (Hewitt, 1995). Since 1995, the Consortium has grown into a comprehensive interpreter testing program with 40 member states.

1.2.1 The Consortium for Language Access in the Courts' Written Exam Development Processes

CLAC's written exam was donated by a member state in the late 1990's and adapted by an appointed Consortium task force in early 2000, and several practicing interpreters, a psychometrician, judges, and attorneys reviewed the written exam content during the adaptation process. The Consortium written exam is an English-only exam with a total of 135 multiple-choice questions divided over ten sections, consisting of:

1. General Vocabulary - Sentence Completion
2. General Vocabulary – Synonyms in Context
3. General Vocabulary – Synonyms
4. General Vocabulary – Antonyms
5. Idioms
6. Sentence Completion
7. Court-Related Questions
8. Sequence
9. Professional Responsibility and Ethics
10. Scenarios

⁹ The *Standards for Educational and Psychological Testing* is widely considered the primary source for establishing testing industry standards and test validity for test publishers, test-takers, and test-users (Hambleton, Merenda, & Spielberger, 2006). The *Standards for Educational and Psychological Testing* was originally developed in 1985 by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education, and was revised in 1999. It provides a comprehensive set of testing guidelines reflecting U.S. Federal law and measurement trends in test validity.

The total time given to complete the test is 2 hours and 15 minutes. Candidates must pass the written exam with an 80% correct score. In 2004, a panel of 53 subject matter experts consisting of judges, interpreters, and statewide program managers determined this percentage/cut-score¹⁰ using an accepted method of standard-setting known as the Angoff Method.¹¹

Several pilots of CLAC's written exam have been performed to cull poorly-performing test items and replace certain items with new ones. Written exam pilot analysis includes examining item-level statistical sets containing data with regard to standard deviation, standard error of measurement, and alpha reliability. (For a glossary of testing terminology, see Appendix 2.)

1.2.2 The Consortium for Language Access in the Courts' Oral Exam Development Processes

The *TCM* was developed from 1995 to 1996 by several recognized SMEs in the field of court interpreting, and was independently reviewed by a psychometrician at least once during its developmental stages.¹² Since 1996, the *TCM* has undergone several revisions to provide an in-depth, language-neutral test blueprint that describes each section of Consortium oral exam development as well as the percentage and category types of scoring units to be included in each section. According to the *TCM*, scoring units are linguistic phenomena that interpreters must be able to render to deliver a complete and accurate interpretation. The *TCM* ensures that each oral exam has the same distribution of scoring units, plus or minus 10 percent. The weighting of scoring units in the *TCM* was heavily determined by scoring unit distributions found in the Federal exam, as well as the expert judgments of the SMEs involved in original test construction. (For a detailed description of the Consortium oral exam as compared to the Federal model, see Appendix 10.) By using the *TCM* during each stage of test development for every language developed, the Consortium has created a consistent and transparent method of pre-establishing test equivalency across languages of the oral exam.

“ The *TCM* ensures that each oral exam has the same distribution of scoring units, plus or minus 10 percent. ”

Original and current oral examinations developed for the Consortium are high-stakes criterion-referenced assessments¹³, testing the three modes of interpreting required of an interpreter in the courtroom with a total of 215¹⁴ scoring units:

- Simultaneous Interpreting: English to Foreign Language
- Consecutive Interpreting: English to Foreign Language and Foreign Language to English
- Sight Translation: Foreign Language to English and English to Foreign Language

The number of required scoring units was closely modeled after the Federal exam's 220 scoring units. (For a detailed description of the CLAC's current oral exam, see Appendix 14.) According to Consortium staff interviewed, every oral exam, regardless of language, has the same number of scoring units per section. The distribution of scoring unit categories included might vary slightly across languages, but exams for each language must stay within the recommended weight and category of units per section as identified in the *TCM*.

Currently, the Consortium has full oral interpreter exams in 16 languages.¹⁵ Original English-based scripts for the oral exams were chosen by Federally certified court interpreters who searched through authentic court documents and as a group chose the best base documents for test development.

¹⁰ A cut-score is a specified point on a score scale, such that scores at or above that point are interpreted or acted upon differently from scores below that point (AERA, APA, & NCME, 1999).

¹¹ The Angoff Method relies upon the judgment of subject matter experts (panelists) who examine the content of each test item/task and predict the proportion of minimally-qualified candidates who will answer the item correctly. The average of the panelists' predictions for a test item becomes the predicted difficulty of the test item.

¹² Currently, a Consortium staff member is responsible for psychometric evaluations of Consortium test data.

¹³ Criterion-referenced tests are used to make score interpretations in relation to a functional performance level (AERA, APA, NCME, 1999).

¹⁴ Based on the current total number of scoring units; some older versions of the consecutive section contain 75 units as opposed to the updated number of 90 units.

¹⁵ The languages are Modern Standard Arabic (simultaneous and sight translation sections only) and Egyptian Colloquial Arabic (consecutive section only), Cantonese, French, Haitian Creole, Hmong, Ilocano, Korean, Laotian, Mandarin, Polish, Portuguese, Russian, Somali, Spanish, and Vietnamese. In addition to the full oral exam administered by the Consortium, an abbreviated oral exam is also offered for certain languages, including Bosnian, Chuukese, Marshallese, and Turkish. The abbreviated exam consists of a simultaneous interpreting section and a spoken English proficiency test.

The Consortium employs highly qualified personnel as part of its test development team. Typically a Consortium test development team consists of one practicing interpreter and one linguist, ideally with the highest credentials possible in their respective fields. Each team of language consultants/exam writers is provided with a test writing handbook and the *Court Interpreter Oral Examination: Test Construction Manual*. A test development manager (TDM) is responsible for the creation and implementation of each new oral exam and for the oversight of language consultants.

The team of experts assigned to a particular language first works on choosing an English to foreign language base sight translation that is culturally appropriate. The team then identifies the scoring units by following the test construction standards found in the *TCM* for scoring unit selection and proper distribution.

Next, the consecutive section is developed. The language team reviews an English base script to ensure that the content is culturally sensitive and does not raise any unnecessary linguistic challenges for that particular language. Developers then adapt portions of the English base script into the foreign language. The process of adaptation requires that developers not merely translate those sections of the script, but that they consider the nuances of the language both culturally and grammatically when transferring the information from one language to another. It is the job of the test development manager to “explain to test writers that they should produce a document that is a true dynamic equivalent of the English text and does not sound like a translation” (Consortium for Language Access in the Courts, *TCM*, 2010). Once the

adaptation process has taken place, the development team reviews the script to check for fluidity of the language and assigns scoring units.

“ It is the job of the test development manager to “explain to test writers that they should produce a document that is a true dynamic equivalent of the English text and does not sound like a translation.” ”

**Consortium for Language Access
in the Courts, *TCM*, 2010**

The simultaneous section is created much like the English sight translation in that developers work with an English base script, determine if any of the material is culturally inappropriate, and choose the appropriate scoring units for that section.

The last section developed is the foreign language to English sight translation. Test developers locate a document which might be encountered in a courtroom setting, and which is written in the target language. If a document of that nature is not found, they revert to searching for English-language documents and begin the process of translation/adaptation for that document. If a language consultant adapts a script or document

for use in a particular language, a professional translator, ideally with American Translators Association (ATA) credentials, reviews that script or document in order to confirm the accuracy and consistency of the adaptation.

After a review of oral exam scripts (including judicial and technical committee reviews) has been conducted, the exam is piloted. Due to the small size of the pilot candidate pool, most pilots are conducted from live versions of the exams. After a sufficient number of candidates have taken the exam to establish a representative analysis, the Consortium looks at classical statistics such as the mean score and standard deviation to ensure that a test is functional, valid, and reliable. Ideally, the Consortium considers 200 candidates to be a representative sample for pilot-testing. However, due to the size of the candidate pool for certain less commonly tested languages, this is not always possible. For less commonly tested languages, the Consortium will run statistical analyses with 30 - 40 sets of exam data.

To ensure that the oral exam scripts reflect national courtroom perspectives, the Consortium Technical Committee¹⁶ developed a glossary of legal terms. Using member state feedback, this glossary was used to identify terms that could be used as scoring units, and to provide potential candidates with a useful study guide of these terms. Items included in the glossary were required to be items accepted across all judicial systems, and not state-specific. The Consortium also limits the use of regionalisms in all test content and attempts to keep test content neutral and universal. Idiomatic expressions, slang, and vulgarities included in oral tests for the Consortium should be content that is recognized in all countries that speak the target language. However, a candidate is given credit if he/she uses a regionalism during the test, as long as that regionalism carries the same meaning as the word or scoring unit being rendered.

According to interviewees, the register¹⁷ of the language found in Consortium oral exams can contain a range of phrases from legalese to profanity in either language; therefore, a candidate needs to possess both the lowest and highest levels of vocabulary and grammar in both languages to accommodate for the range in complexity of the test content.

To pass a Consortium oral exam, a cut-score of 70% is needed. This cut-score was determined by the Consortium Technical Committee based on the Federal court interpreter oral exam cut-score and research conducted by the state of New Jersey during its original program implementation. Specific data considered when evaluating the performance of an oral exam consists solely of an objective assessment and the cut-score percentage assigned based on the scoring units rendered correctly by a candidate.¹⁸

Additionally, for each oral exam developed, a scoring unit dictionary is created to facilitate the rating or scoring process of that exam. A scoring unit dictionary is a dictionary of acceptable and unacceptable terms initially identified by language consultants and updated by exam raters in an ongoing process. Once the script and scoring units for an exam are selected, all of the scoring units are put into a table that lists the scoring unit number, its category/definition, the actual scoring unit, and two columns denoting the acceptable and unacceptable interpretations of the scoring unit (see Table 1).

Table 1: Sample Scoring Guide

NO.	SCORING UNIT CATEGORY	SCORING UNITS	ACCEPTABLE RESPONSE(S)	UNACCEPTABLE RESPONSE(S)
1	D: Legal Terms/ Phrases	state's attorney	Acceptable interpretations of the various scoring units would be entered in the spaces below.	Unacceptable interpretations of the various scoring units would be entered in the spaces below.
2	C: General Vocabulary	county		
3	D: Legal Terms/ Phrases	jury		
4	I: Embeddings/ Position	hours		

Taken from Appendix 2, page 21 of the *Court Interpreter Oral Examination: Test Construction Manual*

¹⁶ The Consortium's Technical Committee is a standing committee of the Consortium, responsible for oral test instruments, standards of test administration, and the development and modification of Consortium manuals. The Technical Committee is comprised of five members from member states. At least three of the members should have substantial expertise in as many of the following as possible: interpreter test construction, rating, or professional interpreting. A fourth member should be an individual who is a practicing court interpreter who has passed the Consortium test or equivalent, or who is Federally certified and has expertise in test development and test rating. The fifth member should be an individual who will work toward developing the kind of interpreting/test construction and rating expertise required of the first three members.

¹⁷ *Register* refers to the ability to reflect the level and tone of the language being used, including formal or informal.

¹⁸ Previously, the Consortium also used a holistic score to judge a candidate's performance based on overall performance, but this holistic rating method has since been discontinued. The holistic score is currently being phased out of the oral exam evaluation process because it was deemed no longer necessary due to the fact that a situation in which a candidate renders the scoring units correctly but does not have the requisite language skills has never occurred throughout the history of oral exam administration.

1.2.3 Process Used for Adapting Donated Non-Consortium Tests to Meet Consortium Guidelines

The *Agreements for Consortium Organization and Operation* states that tests donated to the Consortium by member states are referred to as “Consortium-modified tests” and have “undergone modifications of [their] text and scoring units through substitution, additions or deletions, followed by Technical Committee-approved professional review or pilot-testing.” The *Agreements* additionally explain that “all Consortium tests, whether new or modified, will reflect standardized testing objectives related to the general professional responsibilities of interpreters and the common needs of state courts, as defined by the *Court Interpreter Oral Examination: Test Construction Manual*.”



The adaptation of non-Consortium exams occurs infrequently, but, as noted during the interview process, if the Consortium were to adapt donated test material in the future, its procedure would be to compare the donated exam against the *TCM*.

1.2.4 Qualifications, Training, and Recruitment Processes for Exam Writers and Developers

Ideally, language consultant teams consist of at least one “practicing professional interpreter who possesses the highest credentials in the field” and who has experience in applied linguistics, and one “theoretical, scholarly linguist who has the most formal academic training possible in the linguistics of the language” (Consortium for Language Access in the Courts, *TCM*, 2010). Language consultants are led through the test development process by the TDM, who is a Consortium staff member with prior test development experience, and who is assigned to train and manage the development team for each particular language.

The initial orientation process and training for exam writers and developers is always conducted face-to-face. Consultants receive at least one day of development process training that consists of learning the process of adapting a test for use in a foreign

language, learning what scoring units are and how to choose them, and a discussion of the categories assigned to each scoring unit. Additional training occurs to reacquaint an exam writer and/or developer with the testing process if a test is revised or newly created.

In recruiting exam writers and developers, the Consortium relies heavily on professional recommendations.

1.2.5 Psychometric Audit Findings

After performing an initial review of the Consortium’s written and oral test development processes, ALTA engaged a team of psychometricians to conduct a psychometric audit of the Consortium’s court interpreter certification testing program. The relevant audit standards for the program were categorized according to the suggested outline for a technical manual for a testing program, as determined by the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999) and the *Handbook of Test Development* (Downing & Haladyna, 2006).

The audit identified strengths of the Consortium’s processes and areas for improvement. The key sections evaluated were:

1. Organizational Characteristics
2. Purpose of the Testing Program
3. Test Development
4. Psychometric Properties
5. Administration
6. Security

1.2.5.1 Organizational Characteristics

The purpose of this section was to review the Consortium's staff qualifications and experience along with the organizational structure and relationships to any training or preparation programs.

The Consortium employs a staff with extensive experience in courtroom procedures, including interpretation in the courts. In addition, the staff works with external consultants who represent expertise in test development, and more specifically, language testing. The Technical Committee of the Consortium currently consists of five representatives from the states of Colorado, California, Texas, and Ohio, and one working professional from the court interpreting community. The Technical Committee of the Consortium also retains consultation services with an external psychometrician and language testing specialists. These specialists provide technical guidance on test development and psychometric requirements and analyses.

1.2.5.2 Purpose of Testing Program

CLAC's test purpose is explicitly stated on the first page of the *TCM*:

"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 intended test population is also referred to in the test purpose: bilingual candidates who would like to become court interpreters.

In addition to the test purpose, a technical manual that includes specific studies focused on the design of the exams and how the exams measure the intended knowledge, skills, and abilities required for court interpretation would enhance the validity evidence needed to support the documented intended uses of the exam results.

“ *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.* ”

1.2.5.3 Test Development

The Consortium completed an effort to determine the full set of knowledge, skills, and abilities (KSAs) required for court interpreting; this was done through the collaboration of prominent SMEs to define the content covered by the oral exams through the creation of the *TCM* between the years 1995 and 1996. This included input from judges and courtroom interpreters for both the written and oral exams. In addition, each member state had the opportunity to contribute to the legal glossary from which the scoring units for the oral exams were derived. The content sections for each exam are clearly defined and justified in terms of the importance for credential-worthy performance in the court interpreting occupation and profession.

The *TCM* provides a structured process by which the oral exams must be created and submitted to the Technical Committee for review. This includes the specifications for the scoring units and the criteria that must be met within each exam form.

However, an internal process document (e.g., a test construction manual) for the written exam that is similar to the one currently available for the oral exam would provide systematic development, review, and revision processes for subsequent written exam versions. Additionally, for both the written and oral exams, a formal process for reviewing pilot test results, making changes to items when necessary, and re-piloting items to gather new data for evaluation prior to operational use would further contribute to the validity of the Consortium's test program.

1.2.5.4 Psychometric Properties

For both written and oral CLAC exams, score reliability information was provided across several versions and/or languages. For many of the examinations, information in the form of classical item analysis and test analysis statistics collected suggested support for the internal consistency and reliability values for exam versions.

As mentioned previously, a modified Angoff (Angoff, 1971) standard-setting procedure was conducted for CLAC's written exam in 2004 involving 53 subject matter experts (SMEs) including attorneys, judges, interpreters, and statewide program managers. The overall modified Angoff standard-setting resulted in a standard cut-score of 80% for the written exam versions. While a formal standard-setting process was not conducted to determine the 70% cut-score for the oral exam, the cut-score was determined based on the Federal court interpreter model's cut-score and research conducted by the state of New Jersey during its original program implementation.

1.2.5.5 Test Administration

The Consortium has test administration and scoring processes that are standardized, and are supported by policy documents.



The Consortium provides overview and exam section manuals for the written and oral exams. These materials summarize the test purpose, specify the procedures for test administration, and define the intended populations of test takers. Additionally, the Consortium provides training for test administration supervisors and test proctors for the exams to ensure a standardized administration and scoring process.

1.2.5.6 Security

The Consortium provides acceptable security measures for test items, test materials, score keys, and administration procedures. Confidentiality and

nondisclosure forms are signed by all test development personnel, test administration personnel, and examinees; access to test scores is confidential to the examinee and to trained and qualified member state program directors and Consortium staff.

One potential security issue is the small number of test versions for the written and oral exams by language. This could lead to unintended overexposure. All testing programs are potentially affected by this security issue, with language testing programs affected in particular due to the difficulty and expense of developing language exams. However, it is important to remain aware of the potential of overexposure and the need for revision in those exams that have only one version.

1.3 Analysis of Findings of the Assessment of the Consortium Test Development Processes

The Consortium for Language Access in the Courts has created a comprehensive interpreter certification testing program based on solid and consistent test development processes that encompass both an English-only written exam and a bilingual oral performance exam in more than a dozen different languages. Based on these findings, ALTA was able to gain an in-depth understanding of the Consortium's test development processes and its adherence to current test industry standards.

The Federal court interpreter exam, which has been legally shown to be valid (Arjona, 1985) provided the Consortium with a strong model for test development. CLAC's key strengths lie in personnel involved in test development, its use of the *TCM*, its objective scoring method for oral exams, the standard-setting process used to determine the cut-score for the written exam, and the test analysis statistics collected for the written exam. Specifically, the *TCM* provides test developers with detailed guidelines necessary to establish test validity.

CLAC's key areas for improvement identified during the psychometric audit include the need for further documentation of its development processes to maintain a database of evidence that CLAC exams are meeting their intended test purpose and design, and the collection of oral exam statistical data to enhance analysis of the performance of test items.

1.4 Conclusions Regarding the Assessment of the Consortium Test Development Processes

Since its inception in 1995, the Consortium has developed a valid interpreter certification program that meets test industry standards with a few exceptions. Considering the limited number of established court interpreter testing programs in 1995, the Consortium engaged several of the best subject matter experts in the country to design and implement its original testing program.

The creation and evolution of the *Court Interpreter Oral Examination: Test Construction Manual* has played a pivotal role in establishing the validity of the Consortium's test development procedures. The *TCM* is a key element in establishing test equivalency across versions and languages by ensuring that each oral exam has the same distribution of scoring units, plus or minus 10 percent. The transparency and consistency as evidenced by the use of the *TCM* during every phase of Consortium test development contributes to a documented method of establishing test validity and equality.

While relevant testing standards for the Consortium's court interpreter certification program were reviewed and suggested improvements were made during the psychometric audit process, it is important to note that all testing programs require continual upgrades and maintenance and contain areas that could benefit from improvement.

1.5 Recommendations from the Assessment of the Consortium Test Development Processes

Based on the Consortium's adherence to test industry standards when developing its written exam, specifically the designation of the written exam cut-score and the SMEs involved in the exam's construction, it is ALTA's recommendation that the test development process of the CLAC exams is sufficient so that California could use CLAC's written exam as a screening device for certified and registered court interpreter candidates.

It is also recommended, based on the rigorous standards inherent in the test development process employed during the construction of CLAC's oral exams, specifically the use of the *TCM* and the SMEs involved in oral exam construction, that the CLAC's oral exams could also be used by CCIP to certify court interpreters in California.

Based on results of the psychometric evaluation, improvements to the Consortium's testing process were suggested to enhance the validity of the Consortium's test program but are not considered prerequisites to California choosing to adopt Consortium exam content and working with the Consortium in the future. These suggested improvements include:

- An expert in testing and measurement processes should be a regular participant of the Consortium Technical Committee; and/or, staff members at the Consortium should be trained to have an understanding of the test development and validation processes.
- All test development steps should be documented and stored in a secure location for future reference and added validity measures.
- While the Consortium does perform pilot testing, a more rigorous pilot testing process¹⁹ is recommended to provide further evidence that test items/scoring units are functioning correctly.
- In advance of any development activities, the Consortium should develop a policy manual for the written exam using the same structure that exists for the oral exam. Any subsequent development activities should then be documented (including who was involved, what process was followed, what the results were, and what actions were taken based on the results) to demonstrate adherence to the Consortium's development policy.

19 Formal pilot-testing occurs with 30 – 100 candidates within a target test population prior to a test becoming operational. Decisions about good vs. bad items included in a test can be made with as few as 30 people per test version as long as they are representative of the test population (Downing and Haladyna, 2006).



**COMPARATIVE ANALYSIS
OF THE CALIFORNIA
COURT INTERPRETER
PROGRAM EXAMS AND
THE CONSORTIUM FOR
LANGUAGE ACCESS IN
THE COURTS' EXAMS**



After performing an assessment of CLAC's test development process, ALTA conducted a review and analysis of CLAC's current written and oral examinations in comparison to CCIP's written and oral examinations. This analysis included an identification of the 26 measurable KSAs (as determined in *Study 2007*) measured by the CLAC's exams, as well as a mapping of the results of analysis of CCIP exams produced through *Study 2007* to the results of analysis of the Consortium's exams produced by the current study. (For a list of the KSAs produced by *Study 2007*, see Appendix 4.) Standard-setting recommendations were also made to ensure equivalency between the passing scores required for each program.

2.1 Methodological Approach of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams

To compare the exams used for certifying court interpreters by CCIP and CLAC, the following methodological approaches were included: a test content review conducted by SMEs of the Consortium's written and oral exams, a comparative analysis of CCIP and CLAC test programs, a standard-setting session, and a continued literature and document review. (For further explanation of each methodological approach taken, see the Project Methodology on pages 9 – 10 of the Introductory section of this Report.)

2.2 Findings of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams

2.2.1 California Court Interpreters Program and Consortium for Language Access in the Courts' Written Exam Comparison

The CCIP and CLAC written exams are both screening exams in which test candidates are scored upon their ability to correctly answer a variety of multiple-choice questions. The Consortium uses an 80% cut-score determined by an Angoff Panel while California uses a 70% cut-score that is scaled to ensure comparability of exams across the 12 distinct language combinations tested by its written certification exams. However, for all non-designated languages²⁰ in California, the written exam consists of three English-only portions: English Vocabulary, English Grammar and Word Usage, and Reading Comprehension – English. In addition to these three sections, the written exam for registered languages also contains a section on interpreting principles, practices, and procedures. For purposes of the current Consortium study, only the English-only portions of the California written exam were compared to the Consortium exam since the Consortium's written exam is an English-only exam. (For a detailed comparison of each testing program's written exam structure, see Appendix 13.)

A review of the test content of both programs revealed several similarities. Both written exams test an extensive understanding of English vocabulary, word usage, and idioms. The Consortium tests an understanding of English legal terminology as well. The Consortium written exam tests reading comprehension skills and knowledge of interpreter ethics, while the CCIP written exam for registered languages also tests reading comprehension skills and knowledge of interpreter ethics. Both programs test only a passive knowledge of English since they each rely on the proper selection of answers from a limited range of responses and do not test productive speaking skills.

Study 2007 indicated that the educational level of California's written test content ranges between grade level 13 (first-year collegiate) and grade level 14 (second-year collegiate). The test content review performed of the Consortium's written test content determined the same educational level range, between grades 13 and 14, with additional courtroom experience and/or training in legal terminology. Educational levels for both California and the Consortium's written test content were determined through the average educational level assigned by SMEs to each section of the written exam from reviews of test items and overall content for each program conducted during *Study 2007* and the current study of the Consortium's exams.

²⁰ See page 2, footnote 1.

2.2.2 California Court Interpreters Program and Consortium for Language Access in the Courts' Oral Exam Comparison

The oral exam test purpose is similar for both CCIP and CLAC testing programs. Each program's website lists the intended use of the tests as a means of measuring the abilities needed to accurately interpret a broad range of information from one language into another in a courtroom setting.

Prior to initiating new test development, CCIP and the Consortium both conduct item-writer training sessions to familiarize SMEs involved in the test development process with the expectations and process of each program. Once test development for a new language is complete, both organizations also conduct a review process of newly created items that includes review by SMEs.

The CCIP and Consortium oral exams both test the three main modes of interpreting required of a court interpreter: simultaneous interpreting, consecutive interpreting, and sight translation. The types of passages used for oral test script development for each program are very similar. (For a detailed chart comparing each program's oral exam structure and content, see Appendix 14.)

Candidates for the sight translation section of the oral exams of both programs are tested from English to the foreign language and from the foreign language to English. The types of documents used for the sight translations are comparable across programs, consisting of police or investigative reports and correspondence or affidavits. Six minutes are allotted to complete each sight translation, and each document is assigned 25 scoring units for both programs (see Table 2 for a comparison of the number of scoring units in each program's oral exam sections). Scoring units across both programs test a variety of different categories including grammatical structures, legal terminology, general vocabulary, and some dates.

Table 2: CCIP and CLAC Scoring Unit Comparison

Oral Exam Component	CCIP # of Scoring Units	Consortium # of Scoring Units
Simultaneous	50	75
Consecutive	50	90
Sight: FL to English	25	25
Sight: English to FL	25	25
Total	150 Scoring Units	215 Scoring Units

The consecutive section for both programs is approximately 20 minutes long and consists of a dialogue in the form of witness testimony. The Consortium consecutive section tests a significantly higher number of scoring units and may therefore encompass a broader range of categories. Test content review for both programs' consecutive sections revealed that scoring units measured knowledge of legal terminology, general vocabulary, idiomatic expressions or colloquialisms, numbers, and grammatical structures. Additionally, the Consortium's consecutive section also measures slang, false cognates, register, markers/intensifiers, and position. (For a complete description of the Consortium's scoring unit categories, see Appendix 11.)

The simultaneous section for both CCIP and the Consortium is an opening/closing argument monologue delivered at similar speed ranges.²¹ California's simultaneous section is recorded at 120 – 140 words per minute, while the Consortium's is recorded at 110 – 130 words per minute with an average of 120 wpm. Both monologues are delivered in English and test a candidate's ability to render an English monologue into the foreign language. Test content review for both programs indicated that the scoring units for each section were measuring a candidate's knowledge of grammatical structures, legal terminology, idiomatic expressions and colloquialisms, and numbers.

²¹ While the majority of the Consortium's simultaneous section is in monologue form, a brief section of dialogue by several speakers may be included in some language scripts in addition to the monologue.

“ Test content review for both programs indicated that the scoring units for each section were measuring a candidate’s knowledge of grammatical structures, legal terminology, idiomatic expressions and colloquialisms, and numbers. ”

Additionally, the Consortium’s simultaneous section contains 75 units in comparison to California’s 50 scoring units. The higher number of scoring units found in the Consortium simultaneous section might allow for a broader range of scoring unit categories within this particular section.

The main area of difference between CCIP and the Consortium falls under the practice of scoring the oral examinations. California uses a combined scoring method that includes a holistic evaluation of language and interpreting skills in addition to objective scoring units. Overall scores for California are reported on a scale of 1 – 5 with a passing candidate scoring a 4 or above in

language proficiency and interpreting skills, or “a 4 in three of the four tested components and score of 3+ in any one remaining component except the consecutive section which must remain at 4 or above” (Prometric Services, *California Certified Interpreter Oral Performance Evaluation Guide*, 2006). In contrast, the Consortium’s oral exams are scored using only scoring units; a candidate must render a minimum of 70% of the scoring units correct for each section of the exam in order to pass.

Study 2007 indicated that the educational level of California’s oral test content is set at a level ranging between grades 13 and 14 (entry-level collegiate to second-year collegiate). The test content review²² performed of the Consortium’s oral test content also determined the educational level to be at a range between grade levels 13 – 14, with additional courtroom experience and/or training recommended.

2.2.3 Knowledge, Skills, and Abilities Tested by Each Program

The *Study of California’s Court Interpreter Certification and Registration Testing (Study 2007)* found a total of 32 KSAs required for the function of court interpretation. (For a complete list of the KSAs identified by *Study 2007*, see Appendix 4.) Of these 32 KSAs, 26 were determined measurable through examination; these 26 KSAs were adopted by California as the set of KSAs essential for court interpreter certification in California. For the purposes of the current study, the extent to which CLAC exams align with the 26 measurable KSAs identified by *Study 2007* was also determined.

2.2.4 Knowledge, Skills, and Abilities Tested by the Written Exams

Analysis of the written exam versions across both programs illustrated that the same KSAs were being measured by each program with one exception: The California exam measures the ability to preserve accuracy (KSA 5I), while the Consortium’s exam does not. This difference is due to the fact that California’s written certification exam is a bilingual exam that includes a Foreign Language Sentence Translation Section and is not relevant when comparing the English-only sections of each exam program (See Appendices 5 and 6 for the distribution of KSAs by written exam section for California and the Consortium.)

Both the CCIP Written Exam and the Consortium Written Exam test the following KSAs:

2.2.4.1 Linguistic Skills

- Knowledge and use of a broad range of vocabulary, including legal terminology, subject-specific terminology, and slang.

2.2.4.2 Reading Comprehension Skills

- Ability to read and recognize various written contexts, including formal and informal text, subject-specific vocabulary, idiomatic expressions, and colloquialisms.

²² It should be noted that SMEs determined the educational level of the oral exams without reviewing the foreign language sections of the exams. Instead, SMEs performed a review focused on English sections of the exams and oral examination technical manuals found on the Consortium website.

2.2.4.3 Interpreting Skills

- Ability to concentrate and focus;
- Ability to think analytically;
- Ability to select appropriate equivalent for vocabulary or phrases; and
- Ability to conserve intent, tone, style, and utterances of all messages.

Both exam programs test a passive knowledge of English with a focus on vocabulary selection skills. Speaking skills, listening comprehension skills, and the majority of interpreting skills are not tested by either program's written examinations.

2.2.5 Knowledge, Skills, and Abilities Tested by the Oral Exams

Similar to the written exam, a high degree of overlap exists between CCIP and CLAC's oral exam KSAs. The chart below illustrates the KSAs that California and the Consortium have in common and of the four sections of the test (simultaneous interpretation section, consecutive interpretation section, English to foreign language sight translation section, and foreign language to English sight translation section), the percentage of the overall test that contains each KSA. (See Appendices 7 and 8 for the distribution of KSAs by oral exam section for California and the Consortium.)

Study 2007 KSAs		% Tested	
		CCIP	CLAC
Linguistic Skills			
1A	Native-like proficiency in all working languages;	100%	100%
1B	Ability to think and react communicatively in all working languages	100%	100%
1C	Knowledge and use of a broad range of vocabulary, including legal terminology, subject-specific terminology, and slang; and	100%	100%
1D	Knowledge and use of cultural nuances, regional variations, idiomatic expressions, and colloquialisms in all working languages.	100%	100%
Speaking Skills			
2A	Ability to speak with proper pronunciation, diction, and intonation in all working languages;	100%	0%
2B	Ability to speak with a neutralized accent in all working languages; and	100%	0%
2C	Ability to project and/or speak softly.	100%	0%
Listening Comprehension Skills			
3A	Ability to listen to and comprehend different rates of speech in all working languages;	50%	50%
3B	Ability to listen to and comprehend various regional accents and/or dialectical differences in all working languages; and	25%	0%
3C	Ability to ignore auditory distractions and focus on source speaker.	0%	0%

Reading Comprehension Skills			
4A	Ability to read and comprehend overall meaning and specific details of written text in all working languages;	50%	50%
4B	Ability to read and recognize various written contexts, including formal and informal text, subject-specific vocabulary, idiomatic expressions, and colloquialisms; and	50%	50%
4C	Ability to read quickly and with little preparation	50%	50%
Interpreting Skills			
5A	Ability to concentrate and focus;	100%	100%
5B	Ability to process linguistic information quickly;	100%	100%
5C	Ability to make quick linguistic decisions regarding word choice or terminology selection;	100%	100%
5D	Ability to apply short-term memory skills in retaining small units of information;	50%	50%
5E	Ability to think analytically;	100%	100%
5F	Ability to utilize predictive thinking skills to anticipate incoming messages;	75%	75%
5G	Ability to convey meaning;	100%	100%
5H	Ability to provide transference from one language to another;	100%	100%
5I	Ability to preserve accuracy;	100%	100%
5J	Ability to select appropriate equivalent for vocabulary or phrases;	100%	100%
5K	Ability to conserve intent, tone, style, and utterances of all messages;	100%	100%
5L	Ability to reflect register; and	100%	100%
5M	Ability to self-monitor and self-correct.	100%	100%

Of the 26 measurable KSAs listed, CCIP and the Consortium have 21 in common. Neither testing program tests the ability to “ignore auditory distractions and focus on the source speaker.” (See Appendices 7 and 8 for individual charts of the KSAs tested by each program’s oral examinations.)

The main skill set of the KSAs being tested by California that is not tested by the Consortium is Speaking Skills. Both programs test all of the KSAs required in the category of Interpreting Skills.

2.3 Analysis of Findings of the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams

An analysis of findings determined that CCIP and CLAC written and oral exams test candidates at comparable standards of difficulty and at similar standards according to test validity practices. To provide further validation of the comparability found between the CCIP and CLAC oral exam structure and KSAs, a standard-setting session was conducted to analyze the equivalent passing level required between CCIP's oral exams and CLAC's oral exams.

2.3.1 Analysis of the California Court Interpreter's Program and the Consortium for Language Access in the Courts Written Exam KSAs

While the constructs of the written exams for each program may be somewhat different, they both serve the same function at very similar levels of difficulty: eliminating unqualified candidates from moving on to the oral exam phase of the certification process, while allowing qualified candidates to proceed in the process.

The only KSA not tested by the Consortium's written exam that is tested by California is KSA 5I, "the ability to preserve accuracy." In California's written exam, this KSA is only tested in the Foreign Language Sentence Translation Section and is not relevant when comparing the English-only sections of each exam program. Additionally, this KSA is fully tested in the interpreting portions of both the California and the Consortium oral exams.

2.3.2 Analysis of the California Court Interpreter's Program and the Consortium for Language Access in the Courts Oral Exam KSAs

The structure, content, and level of difficulty of the oral exams for CCIP and CLAC were found to be comparable and to adhere to equivalent standards.

The testing of speaking KSAs by CCIP but not the Consortium is due to the different scoring methods each program currently employs. The Consortium's method of scoring by solely using objective scoring units does not allow for the testing of "proper pronunciation, diction, and intonation," (KSA 2A) the "ability to speak with a neutralized accent in all working languages," (KSA 2B) or the "ability to project and/or speak softly" (KSA 2C). The Consortium considers scoring elements such as proper pronunciation difficult to test in an objective manner and therefore specifically avoids testing KSAs that cannot be measured through objective scoring. However, poor speaking skills are indirectly tested through the rating process used by the Consortium. If a test rater cannot hear or understand a candidate's response on a CLAC oral exam, the candidate would not receive credit for the response.

The remaining California KSA not tested by the Consortium oral exams is the "ability to listen to and comprehend various regional accents and/or dialectical differences in all working languages" (KSA 3B). The TCM emphasizes the use of a standardized or neutral accent for the foreign language recordings within the consecutive section of the exam because comprehensive regionalisms and accents are difficult to test in an objective manner. Therefore, the Consortium also specifically avoids testing this KSA.

Of the 26 measurable KSAs identified in *Study 2007*, all of the KSAs required in the area of Interpreting Skills are tested by both the Consortium and CCIP exams.

2.3.3 Standard-Setting Analysis of California's and the Consortium's Oral Exams

In addition to performing a comparative analysis of CCIP exams and CLAC exams, recommendations regarding standard-setting methods were made to ensure functional equivalence between the scoring practices employed by each program.

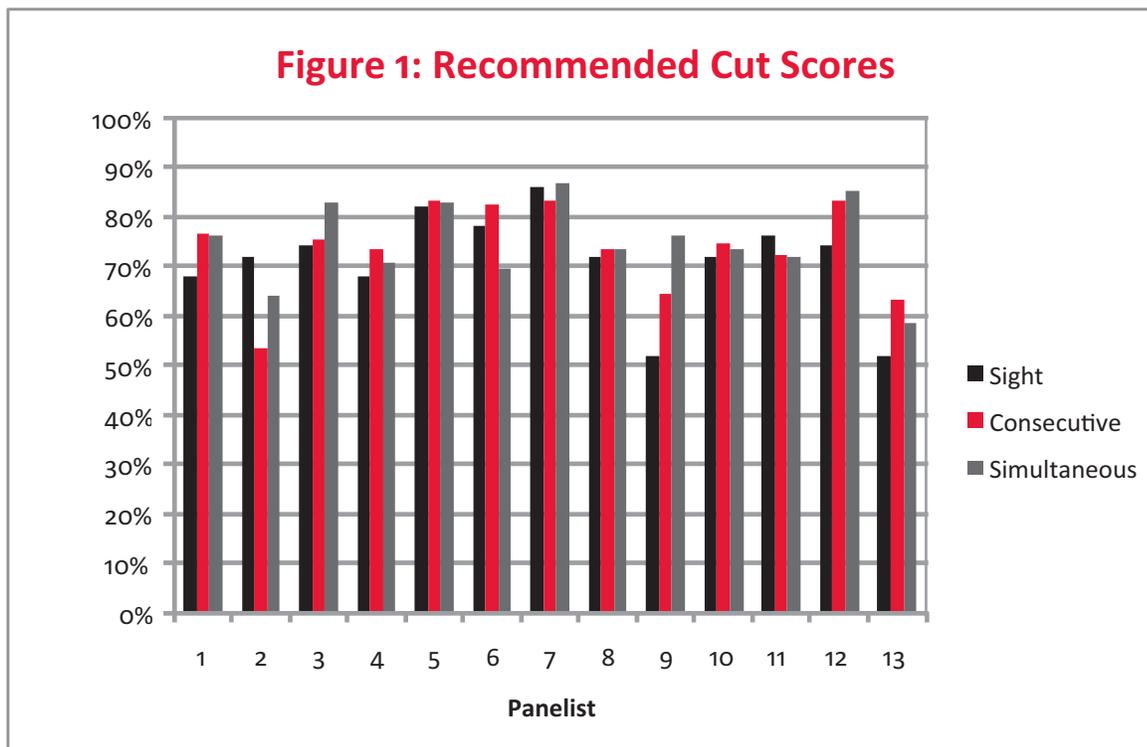
During the research and analysis conducted of standard-setting methods, it was determined that to establish functional equivalence between the passing standards required for the CCIP and CLAC oral exams, a standard-setting session should be conducted in the form of an Angoff Method. (For a detailed description of what occurs during an Angoff Method, see Appendix 15.)

While the test content and level of difficulty are very similar for both CCIP and CLAC oral exams, to ensure equivalent passing standards for both testing programs, an Extended Angoff Method²³ was recommended and conducted to determine the equivalent level of passing scores needed for the highest impact group: Spanish.

To determine the cut-score that would establish equivalency between Consortium test content and California test content, an Extended Angoff Method standard-setting workshop was conducted for the Consortium Spanish oral exam under the supervision of a psychometrician and ALTA staff. Thirteen SMEs from varying regions of the U.S. participated; their qualifications ranged from Federally certified interpreters to university linguists.

SMEs reviewed the list of KSAs required of a California court interpreter as identified in *Study 2007* as well as the necessary requirements needed to pass the California court interpreter exams as listed in the *California Certified Interpreter Oral Performance Evaluation Guide* (Prometric Services, *California Certified Interpreter Oral Performance Evaluation Guide*, 2006), and were asked to determine the percentage of minimally-qualified candidates that would accurately interpret each scoring unit correctly based on the California standards for a court interpreter. Figure 1 shows the recommended cut-scores from each SME by exam section.

23 As with a traditional Angoff Method, in an Extended Angoff Method participants also estimate the number of scale points that they believe borderline candidates would obtain on each performance-based item. However, during an Extended Angoff Method study, panelists are also provided with several different candidate performances to illustrate the range of candidate performances and potential cut-scores.



A range of percentages was identified, with the sight translation section having the broadest range as indicated by the standard error, and the consecutive and simultaneous sections having a narrower range. (For a glossary of testing terminology, see Appendix 2.)

Sight Translation: mean 71.2 %, standard error 2.75%

Consecutive Interpretation: mean 73.8 %, standard error 2.50%

Simultaneous Interpretation: mean 74.7 %, standard error 2.50%

The average mean determined by the SMEs across all three sections of the Consortium Spanish oral exam was 73.2% and the average standard error was 2.6%.

2.4 Conclusions Regarding the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams

The Consortium has done an excellent job of structuring exam content and distributing the use of scoring units across languages and versions. Information gathered during the comparative analysis process indicates that the Consortium and CCIP testing programs share many similarities in the area of court interpreter certification standards and test content. Both programs have also achieved similar standards of difficulty within their test content.

There is a high overlap of the KSAs tested by both CCIP and CLAC written and oral exams. The main difference in KSAs was found in the category of speaking skills and is attributable to the distinct methods of scoring used by each program. These speaking KSAs are indirectly tested by CLAC oral exams during the rating process, since if a rater cannot hear or understand a candidate's utterance of a scoring unit, the candidate would not receive credit for that particular unit. It is also important to note that *Study 2007* recommended that CCIP adopt a method of scoring that emphasizes the use of objective scoring units to standardize the oral examination rating process.

The cut-score for California's use of the Consortium oral Spanish test depends on several considerations. The starting point is the mean cut-score assigned by the total group of SMEs on the entire test. This was calculated to be 73.2% based on equal weighting of the three sections. Next the range of cut-scores for the

entire test using plus and minus two standard errors was calculated to give a score range of 68% to 78.4%. The final cut-score should be within this range. The preferred cut-score is 70% since this is the score that the Consortium currently uses based on its standard setting approach. Moreover, the number of additional test items that need to be answered correctly to get from 70% to 73.2% is very small (approximately six scoring units).

2.5 Recommendations from the Comparative Analysis of the California Court Interpreter Program Exams and the Consortium for Language Access in the Courts' Exams

It is ALTA's recommendation that California could use the CLAC written and oral exams to certify California court interpreters. This recommendation is based on results from the comparative analysis of CCIP exams and CLAC exams, including the structural overlap of both programs' exams, the high overlap of KSAs for both programs, and the results of the standard-setting session.

While the constructs or make-up of the written exams for both CCIP and CLAC are somewhat different, they both serve the same function at very similar levels of difficulty: eliminating unqualified candidates from moving on to the oral exam phase of the certification process, and allowing those candidates who do pass to move on in the process. Due to the difference in test construction, an Extended Angoff Method and/or concordance study of the California and Consortium written exams content would verify that candidates are being screened through the Consortium's written exam at the equivalent standard currently in use in California. (For a detailed description of what occurs during an Angoff Method and concordance study, see Appendix 15.) However, performing an Extended Angoff Method and/or concordance study are not considered prerequisites to California's use of the Consortium written exam based on the nature of the exams as a screening device and the adherence to test development procedures employed during the Consortium's written exam development.

A few CLAC exam sections reviewed were not in full compliance with the standards outlined in the Court Interpreter Oral Examination: Test Construction Manual; all exams or exam sections should be brought into full compliance with those standards to eliminate the potential for bias.

The use of an oral proficiency interview²⁴ (OPI) is one option recommended to ensure that the speaking KSAS that are tested by CCIP but not CLAC are being measured in the screening phase of the testing program.

Given that the results from the Extended Angoff Method for the Consortium Spanish oral exam were within such a close range of CLAC's existing oral exam cut-score, ALTA recommends that California use the 70% cut-score as already established by the Consortium for its oral examinations should California decide to adopt Consortium Spanish oral test content. Furthermore, California could use the 70% cut-score on all oral certification exams in the additional languages shared between California and the Consortium since they have been found to be comparable in structure, content, and level of difficulty, although future standard-setting sessions for CLAC exams in languages other than Spanish are also recommended to provide further evidence of cut-score validity.

²⁴ An OPI is used to holistically assess a candidate's functional speaking ability in a particular language.

APPENDICES

APPENDIX 1 – LIST OF ACRONYMS

ALTA	ALTA Language Services, Inc.
AOC	Administrative Office of the Courts, Judicial Council of California
ATA	American Translators Association
CCIP	California Court Interpreter Program
CIAP	Court Interpreters Advisory Panel
CLAC	Consortium for Language Access in the Courts
FCICE	Federal Court Interpreter Certification Exam
ITC	International Testing Commission
KSA s	Knowledge, Skills, and Abilities
MQC	Minimally Qualified Candidate
NCSC	The National Center for State Courts
OPI	Oral Proficiency Interview
SME	Subject Matter Expert
TCM	Court Interpreter Oral Examination: Test Construction Manual
TDM	Test Development Manager

APPENDIX 2 – GLOSSARY OF TESTING TERMINOLOGY

The following is a list of testing related terminology that can be found throughout the report:

Alpha Reliability	Developed by Lee Cronbach, alpha reliability is a test to determine how reliable a multi-item scale may be for a given population.
Angoff Method	A widely accepted standard-setting approach, in test development, used to establish the cut-score for a test.
Angoff Panel	The subject matter experts who are responsible for determining the cut-score during an Angoff Method.
Equipercentile Equating	An equation method that makes scores on two distinct test forms interchangeable.
Extended Angoff Method	As with a traditional Angoff Method, in an Extended Angoff Method participants also estimate the number of scale points that they believe minimally qualified candidates would obtain on each performance-based item. However, during an Extended Angoff Method study, panelists are also provided with several different candidate performances to illustrate the range of candidate performances and potential cut-scores.
Internal Process Document	Test construction manual
Mean	Average; approximating the statistical norm or average or expected value.
Mean Distribution	The mean of the test population from which scores were sampled.
Point-Biserial Correlations	The correlation between an item score and the total score on a test.
Psychometrics	The branch of psychology that deals with the design, administration, and interpretation of quantitative tests for the measurement of psychological variables such as intelligence, aptitude, and personality traits.
P-values	The probability of achieving a test statistic consistent with or more extreme than the original test results.
Score Reliability	Consistency of the scoring method used by a test.
Scoring Unit	Specific linguistic phenomena that interpreters must be able to render to deliver a complete and accurate interpretation.
Standard Deviation	Standard deviation is a widely used measure of the variability of test scores. It may be thought of as the average difference of the scores from the mean of distribution.
Standard Error	The standard error of a method of measurement or estimation is the standard deviation of the sampling distribution associated with the estimation method.
Test Blueprint	The test blueprint provides the outline for the test forms that will be developed.

APPENDIX 3 – INTERVIEW QUESTIONS



ALTA Language Services, Inc.
Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

Interview for Analysis of Consortium Test Development Process for State Court Interpreter Certification

This is the form for documenting information from an interview. The purpose of the interview is to identify the process used in developing exams for the certification of state court interpreters for the Consortium of State Court Interpreters.

Answers given will be confidential. All answers will be linked to an identification number, not to an individual's name.

Test Development—General

1. Do I have permission to record our interview?

2. Where are the Consortium exams developed? Does the development team for a test work together at Consortium headquarters or is it done remotely?

3. Who is responsible for overall test development procedures and what are his/her qualifications?

4. When hiring exam writers, developers, and language consultants, what are the minimum qualifications they should have? What is the process for recruiting qualified individuals?

5. What type of training do exam writers, developers, and language consultants receive before they begin working on test development? On-going training?

6. How many exam writers, developers, and language consultants are used per test? Is there a standard procedure/number of experts involved?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

7. What types of consultants were used throughout the original test development process? What types of consultants are currently used?

8. Does a test blueprint and/or test specifications exist for both the written and oral exams?

9. What type of review process were/are tests put through? How many times are they reviewed before they are piloted?

10. How is the review process documented?

11. How are the review changes implemented?

12. In what manner does the Consortium proceed in order to meet the *Standards for Educational and Psychological Testing* during exam development?

13. What is the procedure for deciding whether or not to adopt an exam donated by another organization for Consortium use?

14. How does the Consortium adapt, integrate, and equate non-Consortium tests to meet its guidelines?

15. How are psychometric evaluations conducted during the development process?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

16. Who is responsible for conducting the psychometric evaluations?

17. What qualifications must psychometricians have in order to be hired?

18. What is the Consortium's pilot process once an exam has been newly developed or updated?

19. How are pilot candidates identified? What type of feedback, if any, is received from pilot candidates? How is feedback recorded and implemented?

20. How are statistics from the pilot process used?

21. How does the Consortium establish and/or monitor equivalency across test versions and languages?

22. How does the Consortium monitor revisions to an exam and check performance indicators?

23. What are the determining factors for retiring a test version?

24. What statistical data does the Consortium collect and for what purpose?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

- 25. What steps are taken to protect the security of the exams?

- 26. How are the exams protected from overexposure?

- 27. Throughout the Consortium's history, what improvements have been made to its test development process?

Written Exam

- 28. Please describe the process used to create the written exam. How was the content chosen for the original test items?

- 29. How has the written exam been modified and updated since its original development phase? How often is the written exam reviewed and updated?

- 30. How many versions of the written exam are currently in use?

- 31. How was it determined to use an English-only written exam?

- 32. How was the number of questions determined for each section of the written test, as well as the overall test, determined?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

33. How was the allotted time for each section of the written test, as well as the time allotted for the overall test, determined?

34. Please describe how the categories that comprise the ten sections contained in parts I and II of the exam were determined.

35. What educational level is the content found in the written test supposed to reflect? Why was this level chosen, and how was it determined?

36. Describe the decision that led to the development of the written exam as a screener. What validation studies, if any, has the Consortium done to determine the effectiveness of the written exam as a screener?

37. Explain how the ten-sentence translation component was developed and what led to the decision to make it optional.

Oral Exam

38. How did the Consortium identify the first set of languages for oral test development?

39. How does the Consortium currently determine which languages to add to its oral test bank?

40. Please describe the process used to create the oral exams.



ALTA Language Services, Inc.
Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

41. Please describe the process used to develop the oral test construction manual. How has it been modified over the years?

42. How were the scripts that were used for the original oral test development identified? How have they been modified/updated over the years?

43. How was the length and type of scripts used for the oral exam sections determined? How was each script matched to the interpreting mode being evaluated?

44. How are the scripts adapted? Are they created by translating a base script from English, or are they created directly in the target language?

45. If scripts are translated, what methods are used to establish the accuracy of translation? What qualifications do the translators have?

46. How does the Consortium ensure that the oral exam content reflects national courtroom perspectives (legal language, laws, etc.)?

47. What educational level is the content found in the oral test supposed to reflect? Why was this level chosen, and how was it determined?

48. How was the allotted time for each section of the oral test, as well as the time allotted for the overall test, determined?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

49. What is the rate of words per minute for each script? How was this rate decided upon?

50. How does the Consortium ensure standardized recordings (rates of speech, audibility, etc.) across languages?

51. What are examples of language and culture-specific issues that may arise during the development process and how are these issues resolved?

52. What is the process for producing the oral exam recordings? When a language has more than one distinct regional variety in terms of pronunciation, vocabulary, and/or grammar, what is the procedure for choosing the variety to be used on the oral exam? How are non-selected varieties accommodated?

53. How did the Consortium arrive at the decision to offer abbreviated oral exams? What is the development and validation process for these exams?

Test Scoring Practices

54. How are scoring unit dictionaries developed for each oral exam? How have they been modified over time?

55. Describe the process used to accept and add newly introduced renditions for a scoring unit to the scoring dictionary.

56. How was the percentage each section of the oral exam is worth determined?



ALTA Language Services, Inc.
 Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

57. What is the process used for choosing scoring units?

58. How was the number of scoring units for each section of the oral test, as well as the overall test, determined?

59. Does the number of scoring units vary according to version and/or language? Explain.

60. How was the number of repetitions allowed during the consecutive section of the oral exam determined?

61. What specific data is considered when evaluating the performance of an oral exam?

62. How was the current cut-score determined? Written? Oral?

63. Is an Angoff procedure performed for each test version in every language?

64. How does the Consortium keep the cut-score the same across languages? Is it scaled according to language or is it through the initial weighting process during development?

65. How does the Consortium determine the validity of the cut-score?



ALTA Language Services, Inc.
Tel: 404-920-3800 Fax: 404-920-3801 www.altalang.com

66. How was the holistic score developed? What is the basis for the three scales of measurement (unacceptable, borderline, and acceptable)?

67. Is the holistic scoring procedure still in use? If not, please explain why not.

68. What are the equating strategies for scoring multiple test versions of the same language? Explain.

69. How did the Consortium arrive at the different units of measurement in the weighting process for the oral exam (grammar, false cognates, technical vocabulary, etc.) of scoring categories?

Closing Questions & Comments:

70. What improvements are planned for future development processes?

Additional Comments:

APPENDIX 4 – KSAS ESSENTIAL FOR COURT INTERPRETATION

The following knowledge, skills, and abilities (KSAs) essential for court interpretation were defined in *Study 2007*:

Linguistic Skills	
1A	Native-like proficiency in all working languages;
1B	Ability to think and react communicatively in all working languages
1C	Knowledge and use of a broad range of vocabulary, including legal terminology, subject-specific terminology, and slang; and
1D	Knowledge and use of cultural nuances, regional variations, idiomatic expressions, and colloquialisms in all working languages.
Speaking Skills	
2A	Ability to speak with proper pronunciation, diction, and intonation in all working languages;
2B	Ability to speak with a neutralized accent in all working languages; and
2C	Ability to project and/or speak softly.
Listening Comprehension Skills	
3A	Ability to listen to and comprehend different rates of speech in all working languages;
3B	Ability to listen to and comprehend various regional accents and/or dialectical differences in all working languages; and
3C	Ability to ignore auditory distractions and focus on source speaker.
Reading Comprehension Skills	
4A	Ability to read and comprehend overall meaning and specific details of written text in all working languages;
4B	Ability to read and recognize various written contexts, including formal and informal text, subject-specific vocabulary, idiomatic expressions, and colloquialisms; and
4C	Ability to read quickly and with little preparation
Interpreting Skills	
5A	Ability to concentrate and focus;
5B	Ability to process linguistic information quickly;
5C	Ability to make quick linguistic decisions regarding word choice or terminology selection;
5D	Ability to apply short-term memory skills in retaining small units of information;
5E	Ability to think analytically;
5F	Ability to utilize predictive thinking skills to anticipate incoming messages;
5G	Ability to convey meaning;
5H	Ability to provide transference from one language to another;
5I	Ability to preserve accuracy;
5J	Ability to select appropriate equivalent for vocabulary or phrases;
5K	Ability to conserve intent, tone, style, and utterances of all messages;
5L	Ability to reflect register; and
5M	Ability to self-monitor and self-correct.

APPENDIX 5 – KSAS MEASURED BY THE CALIFORNIA WRITTEN EXAM

The following chart illustrates the KSAs as defined in *Study 2007* that are tested by the bilingual written examination sections of the California exams. See Appendix 4 for a description of each KSA.

Section:	1A	1B	1C	1D	2A	2B	2C	3A	3B	3C	4A	4B	4C	5A	5B	5C	5D	5E	5F	5G	5H	5I	5J	5K	5L	5M
1 English Vocabulary			✓									✓		✓				✓					✓	✓		
2 Foreign Language Vocabulary			✓	✓								✓		✓				✓					✓	✓		
3 English Grammar and Word Usage			✓	✓							✓	✓		✓				✓								
4 Foreign Language Grammar and Word Usage			✓	✓							✓	✓		✓				✓								
5 Reading Comp. English			✓	✓							✓	✓		✓				✓								
6 Reading Comp. Foreign Language			✓	✓							✓	✓		✓				✓								
7 English to Foreign Language Vocabulary			✓									✓		✓				✓					✓	✓		
8 Foreign Language to English Vocabulary			✓									✓		✓				✓					✓	✓		
9 Foreign Language Sentence Translation			✓	✓							✓	✓		✓				✓				✓	✓	✓		

APPENDIX 6 – KSAS MEASURED BY THE CONSORTIUM WRITTEN EXAM

The following chart illustrates the KSAs as defined in *Study 2007* that are tested by the written examination sections of the Consortium exams. See Appendix 4 for a description of each KSA.

Section:	1A ²⁵	1B	1C	1D	2A	2B	2C	3A	3B	3C	4A	4B	4C	5A	5B	5C	5D	5E	5F	5G	5H	5I	5J	5K	5L	5M
1 General Vocabulary Sentence Completion			✓									✓		✓				✓					✓	✓		
2 General Vocabulary Synonyms in Context			✓									✓		✓				✓					✓	✓		
3 General Vocabulary Synonyms			✓									✓		✓				✓					✓	✓		
4 General Vocabulary Antonyms												✓		✓				✓					✓	✓		
5 Idioms			✓	✓							✓	✓		✓				✓					✓	✓		
6 Sentence Completion			✓	✓							✓	✓		✓				✓					✓			
7 Court-Related Questions			✓	✓							✓	✓		✓				✓					✓			
8 Sequence			✓								✓	✓		✓				✓								
9 Ethics			✓								✓	✓		✓				✓								
10 Scenarios			✓								✓	✓		✓				✓								

²⁵ KSA 1A: This is an English-only written exam; therefore other languages are not considered “working languages” for native-like proficiency.

APPENDIX 7 – KSAS MEASURED BY THE CALIFORNIA ORAL EXAM

The following chart illustrates the KSAs as defined in *Study 2007* that are tested by the oral examination sections of the California exams. See Appendix 4 for a description of each KSA.

Section:	1A	1B	1C	1D	2A	2B	2C	3A	3B	3C	4A	4B	4C	5A	5B	5C	5D	5E	5F	5G	5H	5I	5J	5K	5L	5M	
1 Simultaneous Section	✓	✓	✓	✓	✓	✓	✓	✓						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2 Consecutive Section	✓	✓	✓	✓	✓	✓	✓	✓	✓						✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
3 Sight Translation: Foreign Lang	✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
4 Sight Translation: English	✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

APPENDIX 8 – KSAS MEASURED BY THE CONSORTIUM ORAL EXAM

The following chart illustrates the KSAs as defined in *Study 2007* that are tested by the oral examination sections of the Consortium exams. See Appendix 4 for a description of each KSA.

Section:	1A	1B	1C	1D	2A	2B	2C	3A	3B	3C	4A	4B	4C	5A	5B	5C	5D	5E	5F	5G	5H	5I	5J	5K	5L	5M
1 Simultaneous Section	✓	✓	✓	✓				✓						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
2 Consecutive Section	✓	✓	✓	✓				✓						✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓
3 Sight Translation: Foreign Lang	✓	✓	✓	✓							✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
4 Sight Translation: English	✓	✓	✓	✓							✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓

APPENDIX 9 – BRIEF HISTORY OF THE FEDERAL COURT INTERPRETER CERTIFICATION EXAM

In 1978, Public Law 95 – 539 (also referred to as the Court Interpreters Act) was passed by the 95th Congress to provide “more effectively for the use of interpreters in courts of the U.S. and for other purposes” (Arjona, 1985). Passage of the Court Interpreters Act charged the director of the administrative office of the U.S. courts with the task of implementing a certification program for the use of Federal court interpreters and designing an exam that would adequately test the skills needed for court interpreting. A project director, Jon A. Leeth, was also appointed to determine the necessary procedures to implement such a program and test, and make recommendations according to research conducted regarding those procedures.

To determine what would be required of a court interpreter certification test, Mr. Leeth first conducted a nationwide job audit that resulted in a list of the knowledge, skills, and abilities “required for both general and court interpretation.” After determining the necessary KSAs, the director appointed a task force of consultants consisting of court interpreters, bilingual judges, linguists, and interpreters directly involved in other testing and training programs to work together to develop the certification exam for Spanish-English, the most widely used language pair for interpreting in the courts during that time. The task force first designed a written screening exam to test “linguistic competence in English and Spanish.” The written examination was created with a GRE-level of difficulty in mind and developed in both English and Spanish (Arjona, 1985).

The written test was reviewed by Federal judges and piloted with a group of 24 Spanish interpreters. Modifications were made according to data collected from the pilot, and the final test product was completed and administered in 1980. The final written examination was a multiple-choice exam that contained antonyms, synonyms, reading comprehension, and usage sections.²⁶

A second task force was assigned to create the final portion of the certification exam, a performance exam that would test the oral interpreting skills of a candidate once that candidate had successfully completed the written examination. The task force of consultants for the oral exam consisted of the same consultants from Task Force I, as well as an interpreter from the United Nations, and a psychologist with research experience in aptitude testing and simultaneous interpreting. Federal judges also participated in the development of the oral exam and reviewed it once it was completed.

Original oral performance test construction consisted of a general interview section, a simultaneous interpretation of an opening statement, a consecutive interpretation of a cross-examination, one sight translation from English to Spanish and one from Spanish to English, and finally a recall section that required a candidate to provide a summary interpretation.²⁷ To ensure the validity of the test construction process, “actual courtroom proceedings were used as primary sources” for all base script development. Texts for the consecutive and sight translation sections were also “specifically written to follow test design specifications” (Arjona, 1985).

The scoring for the oral examination was determined through the use of 220 objective scoring units that included grammar, vocabulary, and conservation, as well as a holistic evaluation of candidate performance that included his/her language skills. The 80% cut score assigned to the Federal oral exam was independently determined by a panel of SMEs who went through a process of “identifying the highest number of errors that could be committed before the level of interpretation becomes so unacceptable that it affects the fairness of the hearing” (Gonzales, Vasquez, and Mikkelson, 1991). The holistic evaluation is only used to “promote a candidate with an objective score that is below but very near the pass point into the ‘pass’ category” and has “never been used to lower a candidate’s score or to demote a candidate from pass to fail status” (National Center for State Courts, *FCICE for Spanish/English Examinee Handbook*, 2009).

²⁶ Currently the Federal written exam consists of five parts: reading comprehension, usage, error detection, synonyms, and best translation of a word or phrase. (National Center for State Courts, *FCICE for Spanish/English Examinee Handbook*, 2009).

²⁷ The recall section was used for experimental purposes only and was not included in the final score a candidate received.

The Federal oral exam has been revised several times since its first administration in 1980. In 1988, two additional oral exam languages, Navajo and Haitian-Creole, were added to the Federal testing program as well. However, currently the Federal Court Interpreter Certification Examination is only offered in Spanish.

In 1980, the validity of the Federal written certification examination was challenged in the U.S. District Court of New York in a case referred to as *Seltzer v. Foley*. The two plaintiffs involved in the case believed that the written examination “tests applicants on language proficiency unrelated to anything encountered normally in a courtroom by bilingual interpreters” and “uses inaccurate and invalid criteria” (*Seltzer v. Foley*, 1980). During the trial, the development process for both the written and oral Federal examinations was discussed at length including the procedural steps taken to ensure the validity of the exams. The outcome of the trial upheld the validity of the Federal exams, for it was found that “the examinations and the parts thereof in the content and manner in which they were given for certification of interpreters for the United States court were fairly, reasonably and comprehensively developed and prepared under outstanding expert guidance, and properly, fairly and reasonably administered” (*Seltzer v. Foley*, 1980).

APPENDIX 10 – COMPARISON BETWEEN THE FEDERAL COURT INTERPRETER CERTIFICATION ORAL EXAMS AND THE CONSORTIUM’S ORAL EXAMS

SIGHT TRANSLATION - ENGLISH TO FOREIGN LANGUAGE

Test Segment	FCICE	Consortium
Time Allowed	5 minutes	6 minutes
Length of Passage	230 words	200-225 words
Number of Scoring Units/ Elements	22	25
Percent of Total Test	10%	11.6%
Description of Passage	Police or investigative reports	Police or investigative reports

SIGHT TRANSLATION – FOREIGN LANGUAGE TO ENGLISH

Test Segment	FCICE	Consortium
Time Allowed	5 minutes	6 minutes
Length of Passage	230 words	200-225 words
Number of Scoring Units/ Elements	22	25
Percent of Total Test	10%	11.6%
Description of Passage	Correspondence, affidavits (formal language)	Correspondence, affidavits (relatively formal language)

SIMULTANEOUS INTERPRETATION – MONOLOGUE

Test Segment	FCICE	Consortium
Time Required	7 minutes	7 minutes
Length of Passage	840 words	800-850 words
Rate of Speech	120 wpm	120 wpm
Number of Scoring Units/ Elements	65	75
Percent of Total Test	29.5%	34.9%
Description of Passage	Opening/Closing Argument	Opening/Closing Argument

SIMULTANEOUS INTERPRETATION – WITNESS TESTIMONY

Test Segment	FCICE	Consortium
Time Required	4 minutes	N/A
Length of Passage	600 words	
Rate of Speech	Varies – up to 160 words per minute, with pauses between Q&A	
Number of Scoring Units/ Elements	35	
Percent of Total Test	16%	
Description of Passage	Witness Testimony (in English)	

CONSECUTIVE INTERPRETATION

Test Segment	FCICE	Consortium
Time Allowed	15 minutes	22 minutes
Length of Passage	850-900 words	850-950 words
Length of Utterances	1-50 words	1-50 words
Number of Scoring Units/ Elements	30 (English to foreign language) 46 (foreign language to English)	40* (English to foreign language) 50* (foreign language to English)
		* Some early versions of the Consortium's exams included 75 scoring units in the consecutive section. Subsequent exams include 90 scoring units.
Percent of Total Test	34.5%	41.9%
Description of Passage	Witness Testimony (direct or cross – Federal Court)	Witness Testimony (direct or cross – State Court)

PASS REQUIREMENTS

Test Segment	FCICE	Consortium
Oral Cut-Score	220 scoring units comprised of: grammar and usage (false cognates, use of verbs, interference, etc.), general lexical range (general vocab., legal terms and phrases, idioms/ sayings), and conservation (register, slang, numbers/names, modifiers, embeddings, etc.). Must obtain 80% of the items correct. A holistic evaluation is also completed to assess candidate's overall strengths and weaknesses.	215 scoring units comprised of: grammar, false cognates, general vocab., technical vocab., idioms, numbers, names, dates, modifiers for emphasis, register and style, special position of words, and slang/colloquial language. Must score 70% on each section of the exam.

APPENDIX 11 - DESCRIPTION OF CONSORTIUM SCORING UNIT CATEGORIES

The following chart was created from information found in the Consortium for Language Access in the Courts' *Court Interpreter Oral Examination: Test Construction Manual*, pages 8 - 9:

Scoring Unit Category	Description	Testing Goals
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." ²⁸	Ensure that candidates recognize and, within the limits of the source and target languages, satisfactorily handle 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 candidate's 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 the candidate's 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.
E Idioms and Sayings	An <i>idiom</i> 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." ²⁹ <i>Sayings</i> are short expressions such as aphorisms and proverbs that are often repeated and familiar, setting forth wisdom and truth.	Determine the candidate's 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.

²⁸ DiYanni, Robert, and Pat C. Hoy II. *The Scribner Handbook for Writers*. Boston: Allyn and Bacon, 1995, p. 221.

²⁹ *The American Heritage College Dictionary*, Third Edition, Boston: Houghton Mifflin Co., 1997, p. 674.

Scoring Unit Category	Description	Testing Goals
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 the candidate's ability to preserve the level of language so that others' impression of the speaker is not raised or lowered by the interpreter and assess the candidate's ability to interpret offensive terminology.
G Numbers and Names	Any number, measurement, or proper name.	Measure the candidate's 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 metric to English), 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 or middle of a long sentence, second in a string of adjectives or adverbs) or function (tag questions).	Ensure that candidates preserve all elements of the source language, especially those that they 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 test texts only when they have passed into widespread usage across the United States.	Measure the candidate's range of knowledge of nonstandard, informal forms of speech and their ability to interpret the meaning of such words and phrases without being bound to preserve their low register.

APPENDIX 12 - STANDARD DISTRIBUTION OF SCORING UNITS FOR CONSORTIUM ORAL EXAMS

The following chart was taken from page 12 of the Consortium for Language Access in the Courts' *Court Interpreter Oral Examination: Test Construction Manual*:

SCORING UNIT CATEGORY	SIGHT E - FL	SIGHT FL - E	CON	SIM	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, Emphases 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
Total	25	25	90*	75	215	100

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

APPENDIX 13 – CALIFORNIA AND CONSORTIUM WRITTEN CERTIFICATION EXAM COMPARISON³⁰

OVERALL WRITTEN TEST INFORMATION

Test Segment	Consortium Written Exam	California Written Exam
Written Format	<p>135 multiple-choice items in English³¹:</p> <p>Sections in Part I: General Language Proficiency: Sentence Completion, Synonyms in Context, Synonyms, Antonyms, Idioms.</p> <p>Sections in Part II: Court-Related Terms and Usage, and Ethics & Professional Conduct: Sentence Completion, Court-Related Questions, Sequence, Professional Conduct Questions, Scenarios.</p>	<p>155 multiple-choice items in English and Foreign language:</p> <p>Nine sections total, including English vocabulary, Foreign language vocabulary, English grammar and word usage, Foreign language grammar and word usage, Reading Comprehension-English, Reading Comprehension-Foreign Language, English to Foreign Language Vocabulary, Foreign Language to English Vocabulary, Foreign Language Sentence Translation.</p>
Written Cut-score	Must pass with 80%.	Must pass with a scaled score of 70% in both English and the Foreign Language.
Written Administration	<p>Total time allowed:</p> <p>Two hours and 15 minutes.</p> <p>Retake policies vary by state.</p>	<p>Total time allowed:</p> <p>Four hours and 15 minutes; there is no limit on how many times a candidate can take the written test.</p>

³⁰ For the purposes of the current study, ALTA examined the English-only sections of the California written examination.

³¹ Also includes a language assessment consisting of ten sentences for translation from English to Foreign Language (this assessment is not endorsed by the Consortium and is not used by every member state).

APPENDIX 14 – CALIFORNIA AND CONSORTIUM ORAL CERTIFICATION EXAM COMPARISON

SIGHT TRANSLATION – ENGLISH TO FOREIGN LANGUAGE

Test Segment	Consortium Oral Exam	California Oral Exam
Time Allowed	6 minutes (includes prep time)	6 minutes (includes prep time)
Length of Passage	200-225 words	290 words (avg.)
Number of Scoring Units/ Elements	25	25
Percent of Total Test	11.60%	22%
Description of Passage	Police or investigative reports	Legal vocabulary

SIGHT TRANSLATION – FOREIGN LANGUAGE TO ENGLISH

Test Segment	Consortium Oral Exam	California Oral Exam
Time Allowed	6 minutes (includes prep time)	6 minutes (includes prep time)
Length of Passage	200-225 words	280 words (avg.)
Number of Scoring Units/ Elements	25	25
Percent of Total Test	11.60%	22%
Description of Passage	Correspondence, affidavits (relatively formal language)	Correspondence, affidavits (relatively formal language)

CONSECUTIVE INTERPRETATION

Test Segment	Consortium Oral Exam	California Oral Exam
Time Allowed	22 minutes	20 minutes (approx)
Length of Passage	850-950 words	900-1,000 words
Length of Utterances	1-50 words	1-40
Number of Scoring Units/ Elements	40 (English to Foreign Language) 50 (Foreign Language to English)	50
Percent of Total Test	41.90%	33%
Description of Passage	Witness Testimony (direct or cross – State Court) Candidate is permitted two repetitions.	Witness Testimony (direct or cross – State Court) Candidate is permitted six repetitions.

SIMULTANEOUS INTERPRETATION – MONOLOGUE

Test Segment	Consortium Oral Exam	California Oral Exam
Time Required	7 minutes	3 ½ minutes
Length of Passage	800-850 words	470 words (avg.)
Rate of Speech	120 wpm	120-140 wpm
Number of Scoring Units/ Elements	75	50
Percent of Total Test	34.90%	22%
Description of Passage	Opening/Closing Argument	Opening/Closing Argument

OVERALL ORAL TEST INFORMATION

Test Segment	Consortium Oral Exam	California Oral Exam
Oral Format	Sight translation (English to Foreign Language/ Foreign Language to English); Simultaneous (English to Foreign Language-120 wpm) Consecutive (English to Foreign Language/ Foreign Language to English)	Consecutive (English to Foreign language/Foreign language to English); Sight translations (English to Foreign language/ Foreign language to English); Simultaneous (Foreign language to English, 120-140 wpm).
Oral Cut-off Score	215 scoring units comprised of: grammar, false cognates, general vocabulary, technical vocabulary, idioms, numbers, names, dates, modifiers for emphasis, register and style, special position of words, and slang/colloquial language. Must score 70% on each section.	General rating of 1-5 for both language proficiency and interpreting skills. Candidate must score 4 or above in all language proficiency component ratings. Candidates must score 4 or above in all interpreting skill components or 4 or above in three of the four components and 3+ in any one remaining component except the consecutive section.
Oral Administration	Uses pre-recorded simulations of courtroom activity. Retake policies vary by state.	Uses pre-recorded simulations of courtroom activity. Offered several times throughout the year. Once a candidate passes the written exam he/she is given 4 attempts to pass the oral exam (there is no time limit to complete the 4 attempts). Candidates must retake all parts of the oral test is one part is failed. If a candidate fails to take and/or pass the oral exam after 4 attempts, he/ she must restart the certification process.

APPENDIX 15 – DESCRIPTION OF RECOMMENDED STANDARD-SETTING APPROACHES

Angoff Method

An Angoff Method is used in determining the cut-score for each section of an exam. This approach relies on the judgments of a panel of SMEs as to the likelihood that a minimally qualified candidate (MQC) would perform successfully on each item. In this method, each member of the judging panel independently determines whether a candidate who meets the definition of the minimally qualified candidate would respond to each item correctly or incorrectly. The MQC is defined as someone who meets the minimum-level criteria necessary for the skill level. Each judge independently assigns a 0 or 1 to each test item based on his or her determination of the likelihood that the MQC would respond correctly. (A score of 0 signifies the candidate would not answer the item correctly; a score of 1 signifies that a candidate would answer the item correctly.) The percentage of these determinations across the set of judges determines the recommended cut-score for the test. In other words, each judge's final determination is averaged together with all of the determinations of the other judges to arrive at the final cut-score for the overall exam. This cut-score can be further analyzed by exam section if needed. In conducting the analysis, a panel of SMEs is recruited, trained on the methodology, and trained on the profile for the MQC. This profile would be inclusive of the cognitive, content, and performance demands of California as identified through the 26 measurable KSAs. Panel members would be required to assign judgments according to this profile. The result would be a score on each Consortium exam that the minimally qualified candidate would be required to obtain in order to meet California's standards.

Concordance Study

A concordance study is designed to map a cut-score from one program to another. It relies on the assumption that the test population is randomly equivalent and compares the impact data of both programs to establish which cut-score on Exam A is equivalent to the cut-score on Exam B. This method assumes that two randomly equivalent samples of candidates have taken each exam. The process applies the cut score from Exam A to Exam B to find the impact (what percentage of candidates passed using that cut score). By analyzing how candidates performed on Exam A, one could find the cut score that would result in the same percentage of candidates passing Exam A and Exam B.

To have confidence in this process, two major assumptions need to be addressed. The first is that the content on the exams is roughly parallel. This is necessary for it to be logical to assume that the exams are of equal difficulty and should have the same pass rate. The second assumption is that the test populations taking each exam are randomly equivalent. This means the groups should have the same composition in terms of knowledge, skills, and abilities as well as demographic characteristics. Part of evaluating the random equivalency of the groups is considering the pre-requisites required for taking the exam and whether both programs have the same requirements. With these assumptions, minimum sample sizes of 200 for each test would be necessary to have a measure of confidence establishing concordance between the exam programs' pass/fail decisions.

APPENDIX 16 – LITERATURE REVIEW

Bibliography

- Abel, L. (2009). *Language Access in State Courts*. Brennan Center for Justice at New York University School of Law.
- AERA, APA, & NCME. (1999). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- Alderson, J. C. (2005). *Diagnosing Foreign Language Proficiency: The Interface Between Learning and Assessment*. Continuum.
- American National Standards Institute (ANSI), International Standards Organization (ISO) & International Electrotechnical Commission (IEC). (2003). *Conformity Assessment – General Requirement for Bodies Operating Certification Systems of Persons*. Washington, D.C.: ANSI and ASTM International.
- Angoff, W.H. (1971). Scales, Norms and Equivalent Scores. In R.L. Thorndike (Ed.), *Educational Measurement* (2nd ed, pp. 508-60). Washington, D.C.: American Council on Education.
- Arjona, E. (1985). The Court Interpreters Certification Test Design. In L. Elias-Olviarez, E.A. Leone, R. Cisneros, & J.R. Gutierrez (Eds.), *Spanish Language Use and Public Life in the United States* (pp. 181-200). Berlin: Mouton de Gruyter.
- Bachman, L.F., & Palmer, A.S. (1996). *Language Testing in Practice*. Oxford University Press.
- Bachman, L.F., Davidson, F., Ryan, K., & Choi, I. (1993). *Studies in Language Testing (1): An Investigation into the Comparability of Two Tests of English as a Foreign Language*. Cambridge: Cambridge University Press.
- Berk-Seligson, S. (2002). *The Bilingual Courtroom: Court Interpreters in the Judicial Process*. Chicago: The University of Chicago Press.
- Bock, R.D., Muraki, E., & Pfeifferberger, W. (1988). Item Pool Maintenance in the Presence of Item Parameter Drift. *Journal of Educational Measurement*, 25, 275-285.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Brennan, R.L. (2001). *Generalizability Theory*. New York: Springer-Verlag.
- Buckendahl, C.W., & Plake, B.S. (2006). Evaluating Tests. In S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Test Development* (pp. 725–738). Mahwah, NJ: Lawrence Erlbaum Associates.
- Cizek, G.J., & Bunch, M.B. (2007). *Standard-setting: A Guide to Establishing and Evaluating Performance Standards on Tests*. Oaks, CA: Sage Publications Inc.
- Consortium for Language Access in the Courts (CLAC). (2010). *Court Interpreter Oral Examination: Test Construction Manual*.
- . (2009). *Oral Test Writing Handbook Draft*.
- . (2008). *Agreements for Consortium Organization and Operation*.
- . (2008). *Becoming an Arabic Court Interpreter*.

- . (2005). *Overview of the Oral Examination for Prospective Court Interpreters*.
- . (2005). *Overview of the Written Examination for Candidates*.
- . (2001). *Test Rating Standards and Resource Materials for Rater Training: Court Interpreting Oral Proficiency Examination*.
- . (2001). *Abbreviated Test Models*.
- . (2000). *Test Construction Manual*.
- . (1996). *State Court Interpreter Certification Meeting Minutes*.
- Council of Europe Language Policy Division. (2002). *Language Examining and Test Development*. Prepared by M. Milanovic.
- . (2009). *Relating Language Examinations to the Common European Framework of Reference Languages: Learning, Teaching, Assessment (CEFR)*.
- Cronbach, L.J., Rajaratnam, N., & Gleser, G.C. (1963). Theory of Generalizability: A Liberalization of Reliability Theory. *British Journal of Statistical Psychology*, 16, 137–163.
- Cronbach, L.J., Rajaratnam, N., & Gleser, G.C. (1972). *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. New York, NY: Wiley.
- Cziko, G. (1982). Improving the Psychometric, Criterion-Referenced, and Practical Qualities of Integrative Language Tests. *TESOL Quarterly*, 16(3).
- Davidson, F., & Lynch, B.K. (2002). *Testcraft*. New Haven: Yale University Press.
- Davis, B.W. (1999). *Court Interpreter Technical Report*.
- Downing, S.M., & Haladyna, T.M. (1997). Test Item Development: Validity Evidence from Quality Assurance Procedures. *Applied Measurement in Education*, 10(1).
- Downing, S.M., & Haladyna, T.M. (2006). *Handbook of Test Development*. London: Routledge.
- Gonzalez, R.D., Vasquez, V.F., & Mikkelson, H. (1991). *Fundamentals of Court Interpretation: Theory, Policy and Practice*. Durham: Carolina Academic Press.
- Haertel, E.H. (2006). Reliability. In R.L. Brennan (Ed.). *Educational Measurement* (4th ed, pp.65-110). Westport, CT: American Council on Education and Praeger Publishers.
- Hambleton, R.K. (1993). *Translation Achievement Tests for Use in Cross-National Studies*. International Association for the Evaluation of Educational Achievement.
- Hambleton, R.K., & Kanjee, A. (1993). *Enhancing the Validity of Cross-Cultural Studies: Improvements in Instrument Translation Methods*. National Center for Education Statistics.
- Hambleton, R.K., & Pitoniak, M.J. (2006). Setting Performance Standards. In R.L. Brennan (Ed.), *Educational Measurement* (4th ed., pp.433-470). Westport, CD: American Council on Education and Praeger Publishers.
- Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2006). *Adapting Educational and Psychological Tests for Cross-Cultural Assessment*. New York: Lawrence Erlbaum Associates, Inc.
- Hewitt, W.E. (1995). *Court Interpretation: Model Guides for Policy and Practice in the State Courts*. Williamsburg, VA: National Center for State Courts.

- Hewitt, W.E. (2007). *Expert Panel on Community Interpreter Testing and Certification, Final Report*. Retrieved from <http://www.umtia.org/ExpertPanel/Bill%20Hewitt%20-%20Expert%20Panel.pdf>.
- Hoekje, B., & Linnell, K. (1994). "Authenticity" in Language Testing: Evaluating Spoken Language Tests for International Teaching Assistants. *Teachers of English to Speakers of Other Languages, Inc.* (TESOL), 28(1).
- International Testing Commission. *Guidelines*. Retrieved July 20, 2009 from <http://www.intestcom.org/>
- Ingham, K. (2008). The Cambridge ESOL Approach to Item Writer Training: The Case of ICFE Listening. *Research Notes*, 32(2).
- Jones, P., Smith, R.W., & Talley, D. (2006). Developing Test Forms for Small-Scale Achievement Testing Systems. In S.M. Downing & T.M. Haldyna (Eds.), *Handbook of Test Development* (pp. 487-525). Mahwah, NJ: Lawrence Erlbaum Associates.
- Judicial Council of California/Administrative Office of the Courts. (2007). *Study of California's Court Interpreter Certification and Registration Testing*. Prepared by ALTA Language Services, Inc.
- Kolen, M.J., & Brennan, R.L. (2004). *Testing, Equating, Scaling and Linking: Methods and Practices* (2nd ed.). New York, NY: Springer Science + Business Media, Inc.
- Kunnan, J.A. (1998). *Validation in Language Assessment*. New Jersey: Lawrence Erlbaum Associates.
- Lane, S., & Stone, C.A. (2006). Performance Assessment. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., 387-431). Westport, CT: American Council on Education and Praeger Publishers.
- Levine, M., & Rubin, D.B. (1979). Measuring the Appropriateness of Multiple Choice Test Scores. *Journal of Educational Statistics*, 4(4).
- Linacare, J.M. (2009a). FACETS: *Rasch-Model Computer Program*. [Computer Software] Chicago, IL: Mesa Press.
- Linacare, J.M. (2009b). A User's Guide to FACETS: *Rasch-Model Computer Programs*. Chicago, IL: Mesa Press.
- Linacare, J.M., & Wright, B.D. (2000). WINSTEPS: *Multiple-Choice, Rating Scale, and Partial Credit Rasch Analysis*. [Computer Software]. Chicago, IL: MESA Press.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Millman, J., & Greene, J. (1989). The Specification and Development of Tests of Achievement and Ability. In R.L. Linn (Ed.), *Educational Measurement* (3rd ed., pp.335-366). New York, NY: American Council on Education and Macmillan Publishing Company.
- National Center for State Courts. (2009). *Federal Court Interpreter Certification Examination for Spanish/English Examinee Handbook*.
- National Organization for Competency Assurance (NOCA) (2004). *National Commission for Certifying Agencies, Standards for the Accreditation of Certification Programs* Washington D.C.: National Organization for Competency Assurance.
- Pommerich, M., et al. (2004). Issues in Conducting Linkages between Distinct Tests. *Applied Psychological Measurement*, 28(4)
- Prometric Services. *California Certified Interpreter Oral Performance Evaluation Guide*. 2006.

- Raymond, M., & Neustel, S. (2006). Determining the Content of Credentialing Examinations. In S.M. Downing & T.M. Haladyna (Eds.), *Handbook of Test Development* (pp 181-223). Mahwah, NJ: Lawrence Erlbaum Associates.
- Robin, F., Sireci, S., & Hambleton, R. (2003). Evaluating the Equivalence of Different Language Versions of a Credentialing Exam. *International Journal of Testing*, 3(1).
- Romberger, W. (2008). Language Access Centers: A Win-Win Idea. *Future Trends in State Courts*.
- . (2009). Scoring the Consortium for State Court's Oral Proficiency Exam for Interpreters. *The ATA Chronicle*, January 2009.
- Rosenbluth, A.R. (1976). The Feasibility of Test Translation between Unrelated Languages English to Navajo. *TESOL Quarterly*, 10(1).
- Schmeiser, C.B., & Welch, C.J. (2006). Test Development. In R.L. Brennan (Ed.), *Educational Measurement* (4th ed, pp.307-353). Westport, CT: American Council on Education and Praeger Publishers.
- Seltzer v. Foley* abstract, 502 F. Supp. 600 (1980)
- Sireci, S.G., & Allalouf, A. (2003). Appraising Item Equivalence Across Multiple Languages and Cultures. *Language Testing*. 20(2).
- Stansfield, C.W., & Hewitt, W.E. (2005). Examining the Predictive Validity of a Screening Test for Court Interpreters. *Language Testing*, 22(4).
- Tannenbaum, R., & Wylie, E.C. (2005). Mapping English Language Proficiency Test Scores Onto the Common European Framework. *Research Reports*, RR-80.
- Thomson, B.S. (2004). Qualitative Research: Grounded Theory-Sample Size and Validity. Retrieved January 4, 2007, from <http://www.buseco.monash.edu.au/research/studentdocs/mgt.pdf>
- Wang, J., Niemi, D., & Wang H. (2007). Impact of Different Performance Assessment Cut-scores on Student Promotion. National Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles. Report 719.
- Webb, N.L. (2006). Identifying Content for Student Achievement Tests. In S.M. Downing & T.M. Haladyna (Eds.). *Handbook of Test Development* (pp. 155-180). Mahwah, NJ: Lawrence Erlbaum Associates.
- Weir, C., & Wu, J.R.W. (2006). Establishing Test Form and Individual Task Comparability: A Case Study of a Semi-Direct Speaking Test. *Language Testing*, 23(2).
- Yin, P., et al. (2004). Concordance Between ACT and ITED Scores From Different Populations. *Applied Psychological Measurement*, 28(4).
- Zumbo, B.D. (2003). *Re-Visiting Exploratory Methods for Construct Comparability: Is There Something to be Gained From the Ways of Old?* Amherst: University of Massachusetts.