

# **Proposed High-Level Test Design for the California Spanish Assessment**

---

**Contract #150012**

*Proposed high-level test design for the California Spanish  
Assessment (CSA) in 2015–2018*

Prepared for the California Department of Education by  
Educational Testing Service

Presented August 4, 2016



*Measuring the Power of Learning.™*



## Table of Contents

<b><u>Section 1 Overview</u></b> .....	<b>1</b>
1A. Introduction.....	1
1B. Document Structure.....	1
1C. Key Assumptions.....	2
1D. Design Team.....	2
1E. Assessment Purpose.....	3
1F. Design Considerations.....	3
1G. Overview of Test Design .....	4
<b><u>Section 2 Design Methodology</u></b> .....	<b>6</b>
2A. Standards and Claims .....	6
<b><u>Section 3 Test Development Process</u></b> .....	<b>8</b>
3A. Step 1—Blueprint Development .....	8
3B. Step 2—Item Development .....	8
3C. Step 3—Forms Development .....	8
<b><u>Section 4 Item Types</u></b> .....	<b>9</b>
<b><u>Section 5 Accessibility Resources</u></b> .....	<b>10</b>
<b><u>Section 6 Test Design Assumptions and Psychometric Considerations</u></b> .....	<b>11</b>
6A. Design Assumptions.....	11
6B. Test Design .....	11
Test Format.....	11
Test Participation Survey .....	11
Test Development Stages .....	12
Pilot Test in Fall 2017.....	12
6C. Psychometric Considerations.....	14
Stand-Alone and Embedded Field Testing.....	14
<b><u>Appendix A. High-Level Test Development Timeline</u></b> .....	<b>17</b>
<b><u>Appendix B. Biographical Summaries: ETS Assessment Design Team</u></b>	
<b><u>Members</u></b> .....	<b>18</b>
Dr. Kenji Hakuta .....	18
Dr. Guadalupe Valdés.....	18
Dr. Patricia Baron.....	19
Dr. Danielle Guzman-Orth.....	19
Dr. Alexis A. López.....	20
Dr. Maurice Cogan Hauck.....	20
Dr. Joyce Wang.....	21
Helen McMahon .....	21
Dr. Ralph Morris .....	21
Jason Gonzalez .....	22
Zulma Torres.....	22
Dr. Rose Payán.....	22
<b><u>Appendix C. References</u></b> .....	<b>24</b>

**List of Tables**

[Table 6.1 2017 Pilot Test Design ..... 13](#)  
[Table 6.2 2018 Field Test Design ..... 14](#)  
[Table 6.3 Psychometric Analyses Plan ..... 15](#)

---

# Section 1 Overview

## 1A. Introduction

California is a state with great linguistic diversity. More than 40 percent of students in California speak a language other than English.<sup>1</sup> Of these students, over 1.2 million speak Spanish.<sup>2</sup> The student population in California includes students who are native speakers of Spanish and students who are learning Spanish as an additional language. California’s educational system includes instruction in Spanish in various forms.

California encourages and values bilingualism and biliteracy; an assessment measuring Spanish reading/language arts skills can communicate acknowledgment of students’ progress toward these goals. Pursuant to California *Education Code (EC)* Section 60640(j), the California Department of Education (CDE) may make available to local educational agencies (LEAs) a primary language assessment aligned to the adopted common core standards in English language arts. In the following pages, Educational Testing Service (ETS) recommends a high-level test design for developing the California Spanish Assessment (CSA). This new computer-based assessment for students in grades three through eight and high school is proposed to measure Spanish skills in reading, writing, and listening. It is part of the California Assessment of Student Performance and Progress (CAASPP) System of assessments.

## 1B. Document Structure

ETS will adopt an evidence-centered assessment design approach for developing and validating the CSA. Suggested high-level claims for the consideration are included in Section 2.

To ensure transparency during the item development process, ETS has outlined the steps the Assessment Development team will take to generate items for the CSA, beginning with the creation of an item based on the standards, and moving through the various stages of content review. The process by which items will be developed is described in Section 3.

ETS will develop machine-scorable item types for online administration by leveraging the most current assessment innovations conducive to assessing reading/language arts skills and practices that can be sustained by California’s assessment delivery system. A description of item types is found in Section 4.

ETS will develop the CSA so that it is accessible to all students, including students with disabilities and English learners. Section 5 provides a description of the research efforts ETS will undertake to identify the accessibility resources for the CSA.

Section 6 provides an overview of the design assumptions and psychometric considerations. 6A summarizes the design assumptions; 6B provides the test format, and describes the participation survey, as well as pilot test and field test plans. Section

---

<sup>1</sup> California Department of Education, DataQuest, Language Census Data for 2012–13

<sup>2</sup> EDFacts/California Consolidated State Performance Report, 2012–13 and 2013–14

6C discusses the psychometric considerations for the CSA. These issues must be kept in mind during the test development process, given the importance of these testing results in demonstrating the level of students' reading/language arts competency through Spanish.

Finally, Appendix A contains a proposed high-level test development timeline leading to the field test in 2018. Appendix B documents the design team biographies and Appendix C contains references that were used in developing this document.

## 1C. Key Assumptions

For planning and development purposes, ETS makes the following assumptions about the CSA:

1. The assessment will be developed with a focus on reading, writing, and listening.
2. The target population for the assessment will be:
  - a. students receiving instruction in Spanish in California; and/or
  - b. students seeking a measure that recognizes their Spanish-specific reading, writing, and listening skills. Section 1E covers more information about this diverse population of test takers.
3. The assessment will be an online, linear test (i.e., not adaptive).
4. No paper-pencil versions of the assessment will be developed (though it will support item by item Print on demand, as an accessibility feature).
5. The CSA will be aligned with the *California Common Core State Standards en Español*.
6. The pilot test for the assessment will be administered in the fall of 2017.
7. The field test for the assessment will be administered in the fall of 2018.
8. General performance level descriptors will be developed.
9. A standard-setting process will be designed and implemented after the first operational administration of the assessment. (Note, this is beyond the timeline of the current ETS contract.)
10. Stakeholder input will be a critical component of the development process; there will be a number of stakeholder reviews throughout the test development process.
11. The CSA will offer a pathway that will not require human scoring.

## 1D. Design Team

The CSA Design Team is comprised of ETS assessment development experts, psychometricians, and research scientists experienced in developing assessments for English learners. The team also includes two nationally recognized experts in linguistics: **Dr. Kenji Hakuta**, the Lee L. Jacks Professor of Education at Stanford University, teaches courses on language development, bilingual education, research methods, and statistics; **Dr. Guadalupe Valdés**, the Bonnie Katz Tenenbaum Professor

of Education at Stanford University, works in the area of applied linguistics. Appendix B provides biographies for members of the Design Team.

## **1E. Assessment Purpose**

ETS, in collaboration with nationally renowned experts, Kenji Hakuta and Guadalupe Valdés, collaborated on the development of the purpose statement for the CSA. The purpose statement was guided by *EC* Section 60640(j), which outlines the intent of a primary language assessment.

The purpose of the CSA is to measure a student’s competency in Spanish language arts in grades three through eight and high school for the purpose of:

- providing student-level data in Spanish competency;
- providing aggregate data that may be used for evaluating the implementation of Spanish language arts programs at the local level;
- providing a high school measure suitable to be used, in part, for the State Seal of Biliteracy.

The targeted test-taking population of the CSA consists of:

- students receiving instruction in Spanish in California
- students seeking a measure that recognizes their Spanish-specific reading, writing, and listening skills

The CSA will be designed for computer delivery and will be 100% computer based. This is a departure from the existing Standards-based Tests in Spanish (STS) which is delivered entirely on paper. The CSA will be aligned with the *California Common Core State Standards en Español*.

## **1F. Design Considerations**

When designing the CSA, it is critical to recognize that there may be competing factors that drive the development process. These factors include the target population for the test, their diverse characteristics, and the context in which the test purpose and use are situated. Today’s students in California are facing a shifting landscape of educational assessment as the state develops and adopts more innovative computer-based assessments designed to measure more rigorous content standards. This paradigm holds true for students taking the CSA.

Attention to the target population and their diverse characteristics is necessary for the design and development of the CSA. As with other CAASPP assessments, the CSA will include a diverse group of test takers, which is comprised of multiple subgroups of students. Anticipated subgroups within the target population include:

- recently arrived English learners (ELs)
- late arrivals (i.e., students arriving in middle school or later)
- students with interrupted formal educational (SIFE) experiences (i.e., students with interrupted schooling, such as migrant or refugee students)

New arrivals have the potential to arrive at any time during the school year and at any grade level from kindergarten through grade twelve (e.g., SIFE, late arrivals). Late arrivals include students arriving to the U.S. in middle school or later.

It is important to recognize that some students may enter California schools with limited experience in formal instruction, which often corresponds to limited exposure to formal assessments in general and, in particular, tests on a computer.

Other considerations that must be taken into account when designing the CSA are sociolinguistic in nature, such as the different varieties of Spanish spoken in California (Valdés, Fishman, Chávez, & Pérez, 2006). Language may vary according to the setting, the relative status of the interlocutors, the topic, or the functional purpose (Finegan & Biber, 2001). Finally, some students in California who experience Spanish as the language of instruction may be learning Spanish as an additional language.

Students seeking a measure that recognizes their Spanish-specific reading, writing, and listening skills may be a heterogeneous subgroup of test takers. Students may:

- use Spanish as their primary language or as an additional language
- be considered heritage speakers of Spanish or Spanish language learners
- have received classroom instruction in English or in Spanish as a foreign language
- have been mainly instructed in Spanish

Attention to the variety in both the student characteristics and the context in which Spanish will be used will be important in assessment design, as well as in score interpretation, particularly as students approach college and careers. When analyzing the CSA test-taking population, a given subgroup will likely have higher representation at certain grade levels, given current demographic trends and the programs in which Spanish instruction is available in California. Students may transition from one type of instruction to another as they progress across grade levels. Attention to these unique population characteristics will help guide design considerations for the CSA.

## **1G. Overview of Test Design**

Despite the wide variability in the target population and context surrounding the CSA, the focus of the assessment in grades three through eight and high school, will be students' demonstrated competency in attaining reading, writing, and listening skills through Spanish. The assessment will provide students an annual opportunity to measure their reading/language arts competency through Spanish. In the future these results may become part of the State Seal of Biliteracy.

As reading and listening are key emerging skills that support student success across all subjects, these skills will be more heavily emphasized in the earlier grades. The assessments in grades three through eight will be shorter than the high school assessment. All grade-level assessments will include a reading, writing, and a listening component.

For high-school students, ETS recognizes California's desire for a Spanish reading/language arts assessment that measures a high level of competency

demonstrated by students, newly arrived or otherwise, who are on track for exiting public instruction as biliterate graduates as highlighted in *EC* Section 51460(a) (State Seal of Biliteracy). In the interest of providing ample opportunity to measure the requisite skills, the high school assessment will include lengthier and more complex passages.

---

## Section 2 Design Methodology

### 2A. Standards and Claims

The foundation of any assessment is the content standards on which the test is based. The Spanish version of the Common Core State Standards, *California Common Core Standards en Español*, was developed as a joint effort between the San Diego County Office of Education, Council of Chief State School Officers, and the CDE.

The *California Common Core State Standards en Español* are a translated and linguistically augmented version of the English-language Common Core State Standards (CCSS) for English Language Arts and Literacy (ELA).

The selection of the *California Common Core State Standards en Español* allows ETS to develop items that will measure a student's competency in Spanish language arts in grades three through eight and high school.

The *California Common Core State Standards en Español* are organized into the following domains:

- Reading standards
- Writing standards
- Speaking/Listening standards
- Language standards

The *California Common Core State Standards en Español* guide instruction in a multitude of contexts, including in-class collaborative activities, group reading of antiquated versions of the language, and Question & Answer sessions following a presentation by peers. Consequently, certain standards, while useful benchmarks for teacher-supported classroom learning, are not conducive to large-scale assessments of a student's performance in isolation.

It should also be noted that while the focus of the *California Common Core State Standards en Español* is acquired language arts competency, the domains above are also harmonious with a four-skill language-learning framework (e.g., listening and reading, known as "receptive" skills, and speaking and writing, known as "productive" skills).<sup>3</sup>

Using the domains as a guide for the test design, ETS recommends the following claims for the CSA:

- A. Claim for Grades Three through Eight: Students can demonstrate progress toward a high level of competency in attaining reading/language arts skills and practices through Spanish.

---

<sup>3</sup> The language standards, which focus on vocabulary, can be seen as an integral support of each of the four skills.

- B. Claim for High School: Students can demonstrate a high level of competency in attaining reading/language arts skills and practices through Spanish.

ETS also recommends the following Spanish language arts competency claims for all grade levels:

- A. Reading: Students can read, analyze, and interpret a variety of texts and genres through Spanish.
- B. Writing: Students can write texts for a range of purposes and audiences in order to accurately and convincingly present, describe, and explain ideas through Spanish.

The CSA will offer two models:

- The first model will consist of items in reading, listening and writing mechanics that can be machine-scored. This model will be used in the pilot test.
- The second model will include the machine-scored items in the first model with the inclusion of constructed- response items aligned to the writing domain. This model could be available in the future. (Note that scoring constructed response items is not covered in the current ETS contract.)

- C. Listening: Students can comprehend spoken Spanish in a range of contexts.

---

## **Section 3 Test Development Process**

### **3A. Step 1—Blueprint Development**

The first step in developing fair, valid, and reliable assessments is to develop high-quality test blueprints and specifications targeted to the *California Common Core State Standards en Español* and aligned with general Performance Level Descriptors (PLDs). The ETS design team is developing blueprints that will align with current best practices for reading/language arts assessments. The general PLDs, and blueprints will be presented to the CDE and California State Board of Education (SBE) for approval.

### **3B. Step 2—Item Development**

ETS assessment specialists will begin to develop items for the assessment that are aligned with the *California Common Core State Standards en Español* and consistent with the goals of California’s testing program. Items will be written by ETS assessment developers and trained item writers familiar with assessment development in Spanish and specifically trained for the CSA as well as California educators who have received item writer training. All items will be reviewed by ETS content and editorial staff, the CDE, and a review panel comprised of California educators.

### **3C. Step 3—Forms Development**

ETS assessment specialists and psychometricians will work closely with the CDE to create test forms as specified in the test design.

---

## Section 4 Item Types

ETS will develop machine-scorable item types for online administration by leveraging the most current assessment innovations conducive to assessing reading/language arts skills.

There are both stand-alone items and passage-based items; all items may contain a stimulus (e.g., a passage, video, or image). Many of the items have technology-enhanced interactions. These interactions include having a student respond by typing an answer, completing a graph, dragging a response to a designated area, using drop-down box selection, or selecting multiple areas in a graphic (also known as a “hot spot”). The assessment industry does not currently offer artificial intelligence scoring of Spanish written responses; items developed for the CSA will be machine-scored.

---

## **Section 5 Accessibility Resources**

The CSA may offer the accessibility resources commonly used in computer-based assessments, where applicable for the construct.

ETS will work with experts to review the field of accessibility resources available for CSA and determine the appropriate supports for this assessment and targeted test taking population. As the CSA will be delivered entirely in Spanish, careful consideration of the quality of supports available in Spanish will be necessary.

Language-based supports, such as glossaries, dictionaries, or even the opportunity to translate test directions (into English), are all being considered for possible inclusion in the assessment. As the development of the CSA ensues, accessibility resources offered in the different phases of the assessment (i.e., pilot test, field test, etc.) will continually be refined.

---

## Section 6 Test Design Assumptions and Psychometric Considerations

### 6A. Design Assumptions

When developing a test, a number of key assumptions have to be made regarding the assessment, such as the purpose of the test, the target population, and the test format. As the assumptions solidify and the requirements of the test become known, elements in the test design will change to support those requirements. Although the scope of this contract is to develop and execute a pilot test and a field test, it is important that these tests are designed to support the provisional operational test design. The following is a list of key assumptions about the CSA:

- It is a language arts assessment that measures a student’s competency in attaining reading, writing, and listening skills and practices through Spanish.
- It targets the following students:
  - Students receiving instruction in Spanish in California; and/or
  - Students seeking a measure that recognizes their Spanish-specific reading, writing, and listening skills.
- It is aligned with the *California Common Core State Standards en Español*.
- It is a linear test delivered online under untimed testing conditions.

The approved draft test blueprint will be available prior to item development for the large-scale field test.

### 6B. Test Design

#### Test Format

The CSA will be a linear test delivered online under untimed testing conditions. It will be untimed in order to allow students to complete the test. Testing time guidelines will be developed and provided to LEAs for scheduling purposes. The developmental maturity and attention span of students will be taken into consideration. It is also important to note, in view of California’s desire for a Spanish language arts assessment that measures a high level of competency demonstrated by students who are on track for exiting public instruction as biliterate graduates, that the high school test will be based on lengthier, more complex passages. Finally, we recognize that the variety of Spanish-language programs and curricula available to students may pose some additional challenges for score interpretation.

#### Test Participation Survey

The CSA targets two distinct voluntary populations: students receiving instruction in Spanish in California and students seeking a measure that recognizes their Spanish-specific reading, writing, and listening skills.

ETS plans to collect information on test participation from LEAs and/or schools through a survey in 2016. Questions on the survey may include, but are not limited to, the following:

- What is the projected test volume for each grade?
- What are the purposes for administering the CSA at each grade?

Results of the participation survey will help refine the sampling plan for the pilot and field test planning for the CSA.

### **Test Development Stages**

To support the 2018–19 operational test design, there will be two test development stages:

- Stage 1: Pilot test development in 2017
- Stage 2: Field test development in 2018

### **Pilot Test in Fall 2017**

The objective of the pilot test is to try out new items and/or new item types on the computer-delivery platform. ETS wants to ensure that taking tests on a computer will not prohibit students from demonstrating what they know and can do. Therefore, the main question ETS wants to address through the pilot test is how students perform in an online environment using the functionalities of online items. The focus is not on the content of the items except as it interacts with the online testing mode. A stratified sample, including all target populations, will be recruited to participate in the pilot test. ETS plans to recruit approximately 300 students per grade band. The results from the test participation survey will help refine the sampling plan for the pilot test.

One pilot test form will be available for each of the following grade bands:

- Upper elementary (grades three–five);
- Middle school (grades six–eight); and
- High school (grades nine–twelve).

The pilot test experience, as outlined in Table 6.1, will be used to refine item-writing guidelines, determine item types that could be administered operationally, observe students' interactions with the *California Common Core State Standards en Español*, and provide other ancillary information.

**Table 6.1 2017 Pilot Test Design**

<b>Item Type</b>	<b># of Pilot Test Items per Form</b>	<b>Estimated Testing Time</b>
Selected-response items (machine-scored)	20–25 items	Grades 3–8: 25–30 minutes High School: 35–45 minutes
Technology-enhanced items (machine-scored)	5–7 items	Grades 3–8: 10–15 minutes High School: 20–30 minutes
Estimated Total	25–32 items	Grades 3–8: 35–45 minutes High School: 55–75 minutes

**Field Test in Fall 2018**

Since the field test forms will reflect the operational test blueprint, the main objective of the field test is to evaluate item and form performance, in addition to providing students an opportunity to familiarize themselves with this new test. The field test will provide item-level information as well as the group performance data required for statistical analyses. For students not participating in the pilot test, the field test will be the first opportunity for them to experience this new online assessment. Results from the field test will allow ETS and the CDE to review the draft test blueprint. For the field test, forms will be administered for grades three, four, five, six, seven, eight, and high school. Table 6.2 shows a tentative field test design. ETS is committed to the development of the number of items shown in Table 6.2. However, the voluntary nature of this test may impact the number of items that are field tested due to potentially low student counts.

**Table 6.2 2018 Field Test Design**

<b>Item Type</b>	<b># of Items Developed per Grade</b>	<b># of Items Taken by Each Student</b>	<b>Estimated Testing Time</b>
Selected-response items (machine-scored)	115–120 items	40–45 items	Grades 3–8: 50–55 minutes High School 60–75 minutes
Technology-enhanced items (machine-scored)	25–30 items	5–10 items	Grades 3–8: 15–20 minutes High School 20–35 minutes
Estimated Total	140–150 items	45–55 items	Grades 3–8: Up to 75 minutes High School: Up to 110 minutes

Based on the estimated testing time of the field test, the approximate testing time for the operational administration of the CSA would be up to 75 minutes.

## **6C. Psychometric Considerations**

Psychometric components of the assessments discussed in this section are based on the test design assumptions described in Section 6A.

### **Stand-Alone and Embedded Field Testing**

In the stand-alone field test and ongoing embedded field test, forms can be spiraled at the student level to achieve random samples for test items. Pilot sampling strategies will be based on an agreement between the CDE and ETS such that conclusions can be drawn for all three target populations of the CSA. An advantage of online testing is the ability to embed field test items dynamically to achieve the desired field test sample size and better meet the requirements of psychometric models.

The embedded field testing of items can occur on an ongoing basis for each operational test administration to support a mutually agreed upon refresh rate. Anticipated testing volumes at various grade levels will be taken into account when considering the frequency of the item refresh rate. Based on low volume counts, a yearly item refresh may not be necessary.

This was the case with the previous STS; however, a key challenge in embedding field test items during operational administrations at the various grade levels is small

population size.<sup>4</sup> If encountered, this restricts the number of items that can be field tested in each operational test administration. There is a greater opportunity to field test more items and refresh test content more rapidly at the lower grade levels, where there is a larger population of eligible students. At the middle and upper grades, enrollment in Spanish instruction has attenuated, but test takers seeking to demonstrate Spanish-specific reading, writing, and listening skills for other purposes may not yet be included. ETS recognizes this as an ongoing challenge; the variety of Spanish-language programs and curricula available to students will require attention when considering score reporting and score interpretation.

### Psychometric Analyses Plan

The psychometric tasks and/or analyses that will be performed after each test administration are outlined in Table 6.3. These analyses are described in more detail below.

**Table 6.3 Psychometric Analyses Plan**

<b>Year</b>	<b>Test Development Stage</b>	<b>Psychometric Task</b>
2015–16	Test Design	Test design, model selection, sampling plans, identification of possible psychometric issues for special studies
2016–17	Pilot Test	Item analysis, differential item functioning (DIF) analysis, and possible item response theory (IRT) item calibration
2017–18	Field Test	Item analysis, DIF analysis, IRT item calibration, and standard setting for achievement levels

### Item Analyses

After each test administration, item analyses and DIF analyses will be conducted. Field-test items that do not meet certain statistical criteria will be excluded from item calibration and later-stage testing.

### Item Parameter Estimation and Scoring

Item calibration will be conducted based on the Rasch model (Rasch, 1960/1980) for multiple-choice items and the Rasch partial credit model (Masters, 1982) for polytomous items. A Rasch model is assumed based on small sample sizes expected at higher middle grade levels; it also supports number correct to scale score tables for operational scoring, which means student scores will be dependent on the number of items

<sup>4</sup> The STS target population consisted of Spanish-speaking ELs who were receiving instruction in Spanish or who had attended school in the U.S. for fewer than 12 months. The target population size ranged from 6,259 in grade three to 669 in grade eleven in 2013; after 2014, the test is no longer mandatory.

answered correctly rather than item patterns. Each grade level test will be on an independent continuous scale so that scores within each grade level can be compared.

### **Equating**

Although operational test equating is beyond the scope of the current work, the discussion and concerns are related to the field test design.

After the field test, base scale and achievement levels will be established and will form the basis for operational score reporting. For the first operational test administration, post-equating can be implemented. Post-equating provides more accurate item parameter estimates, which can be obtained from operational student samples; additionally, it will support standard setting.

After the first operational test administration, it may be necessary to recalibrate items to update item statistics in the item bank.

---

## Appendix A. High-Level Test Development Timeline

Activity	Date
SBE action on the Proposed High-Level Test Design for the California Spanish Assessment	September 2016
SBE action on the proposed test blueprints and general PLDs	March 2017
Administration of the pilot test	Fall 2017
Administration of the field test	Fall 2018
Operational test	Spring 2019
SBE action on the performance-level threshold scores and PLDs	Fall 2019

---

## Appendix B. Biographical Summaries: ETS Assessment Design Team Members

### Dr. Kenji Hakuta

Dr. Kenji Hakuta is the Lee L. Jacks Professor of Education at Stanford University, where he teaches courses on language development, bilingual education, research methods, and statistics. He received his PhD in Experimental Psychology from Harvard University in 1979, has held faculty positions at Yale University and the University of California at Santa Cruz, and served as the founding dean of the University of California, Merced. He currently serves as the co-chair of the Understanding Language Initiative that addresses the challenges and opportunities of the CCSS for English learners (ELs). Hakuta is a member of the National Academy of Education and the American Educational Research Association and a fellow of the American Association for the Advancement of Science. Hakuta's research is in the areas of psycholinguistics, bilingualism, language shift, and the acquisition of English in immigrant students. He is the author and editor of many articles and books, including *Mirror of Language: The Debate on Bilingualism* (1986) and *In Other Words: The Science and Psychology of Second Language Acquisition* (1994), both considered classics in the field. Besides research, Hakuta is professionally active in the areas of language policy, education of language-minority students, affirmative action in higher education, and improvement of quality in educational research. He has served on the boards of the Spencer Foundation and the ETS, and he chaired the National Educational Research Policy and Priorities Board of the U. S. Department of Education. He currently serves on the boards of the National Academy of Education and California Education Partners. Hakuta is actively involved in supporting the work of school districts and states around the country, and he leads several professional learning communities, including at school districts in rural central California and a learning community of state leaders organized by the Council of Chief State Schools Officers.

### Dr. Guadalupe Valdés

Dr. Guadalupe Valdés is the Bonnie Katz Tenenbaum Professor of Education at Stanford University. Working in the area of applied linguistics, much of her work has focused on the English-Spanish bilingualism of Latinos in the United States and on discovering and describing how two languages are developed, used, and maintained by individuals who become bilingual in immigrant communities. Dr. Valdés has investigated Latino students in elementary, middle school, high school, and college, leading to 6 books and more than 70 articles. In the last several years, her work includes a number of articles, including "Toward an ecological vision of languages for all: The case of heritage languages" in A. Heining-Boynton's *Realizing Our Vision of Languages for All* (2006) and "Bilingualism, heritage learners and SLA research: Opportunities lost or seized" in the *Modern Language Journal* (2005). Valdés is also the coauthor of a best-selling Spanish language textbook that focuses on the teaching of Spanish to Hispanic bilinguals. *Español Escrito* (first published by Scribner in 1978 and now published by Prentice Hall) is now in its sixth edition. She was awarded the Joshua Fishman Award for Outstanding Contributions and Leadership in the Heritage Language Field from the National Heritage Language Resource Center at the University of California, Los Angeles in 2010. Valdés is a member of the American Academy of Education, a fellow of the American Educational Research Association, and a

member of the Board of Trustees of ETS. She serves on the editorial boards of a number of journals, including the *Review of Educational Research*, *Bilingual Review*, *Written Communication*, *Modern Language Journal*, and *Hispanic Journal of the Behavioral Sciences*. In May 2000, Valdés received an honorary doctorate from the University of Arizona for her work on the use of Spanish in the United States.

### **Dr. Patricia Baron**

Dr. Patricia Baron, Lead Research Project Manager, has served as the standard-setting director, researcher, and lead facilitator in ETS's Center for Validity Research. In this role, she directs standard setting for ETS's K–12 testing programs. This experience includes providing consultation for the California Standardized Testing and Reporting program, the grades 2–11 Spanish-based assessments of reading and mathematics, the grades 3–11 and Educational Opportunity Center (EOC)–modified assessments, and the California Alternate Performance Assessment, in addition to directing standard setting and performance level descriptor development for the Tennessee EOC, the Proficiency Assessments for Wyoming Students, and the Wyoming Student Assessment of Writing Skills. Dr. Baron has also designed and conducted validation studies and evaluated the alignment of innovative item types to the CCSS. For the past seven years, she has focused on research in factors contributing to the variability in standard setting, the development of mixed methods in curriculum and standards validity studies for state assessments, and on the assessment of young ELs in the international context. Significantly, she completed design and implementation of a standard-setting tool for the Bookmark method, which provides a mechanism for expedited analysis and reporting with high quality assurance standards. Before transitioning into her current position, she worked as the director of Government Relations and Assessment Services, and she was a senior psychometrician in the Research and Development division, conducting hundreds of equating and scaling studies for sixteen years. During her time at ETS, she has been the lead psychometrician on high-stakes undergraduate and graduate admissions tests, outcome assessments for college and higher level programs, and a national assessment for Qatar. She has led development of the vertical scale and test design and helped plan standard setting for Qatar in Arabic and English. She earned her EdD and M.Ed. in Educational Psychology with a specialization in Educational Statistics and Measurement from Rutgers University, where she also earned her BA in Psychology.

### **Dr. Danielle Guzman-Orth**

Dr. Danielle Guzman-Orth, Research Scientist, specializes in monolingual and bilingual assessments, with particular focus on accessibility and accommodations for ELs, including ELs at risk and ELs with disabilities. Along with her involvement in the California Assessment of Student Performance and Progress (CAASPP) primary language stakeholder meetings, Dr. Guzman-Orth has led and consulted on research studies for state and consortia contracts, such as the English Language Proficiency Assessments for California (ELPAC), Smarter Balanced, Partnership for the Assessment of Readiness for College and Careers, and English Language Proficiency Assessment for the 21<sup>st</sup> Century (ELPA21). Her current research projects focus on improving assessment practices for young dual-language learners and ELs with disabilities. Before coming to ETS, she gained valuable classroom experience with ELs, ELs at risk, and ELs with disabilities in P-20 settings. She trained tutors to implement instructional interventions with ELs and students

with disabilities, provided English language development instruction to ELs in first through sixth grade, and taught reading intervention for first grade ELs. She holds an MA and PhD in Education with a specialization in Special Education, Disabilities, and Risk Studies from University of California, Santa Barbara and a BA in Psychology and English from California State University, Stanislaus.

### **Dr. Alexis A. López**

Dr. Alexis A. López, Research Scientist, is focusing on the assessment of language proficiency and the assessment of content knowledge for K–12 ELs in the Center for English Language Learning and Assessment at ETS. For the past four years, Dr. López has conducted research on the use of translanguaging in content assessments, dual-language assessments, and technology-enhanced assessments. He has also led or coled research studies for state and consortia contracts, such as the ELPAC, the ELPA21, and the California English Language Development Test (CELDT) item alignment to the 2012 English Language Development Standards. He previously worked as an associate professor at Universidad de los Andes in Bogotá, Colombia, and as a test development specialist at Second Language Testing, Inc. in Washington, DC. He has participated in all facets of the test development process, including developing test specifications, item writing, field testing, standard setting, and conducting validation and alignment studies. He earned both his PhD in Education and MA in Teaching English as a Second Language from the University of Illinois at Urbana-Champaign, and his BS in English and Spanish from the Universidad Pedagógica Nacional in Bogotá.

### **Dr. Maurice Cogan Hauck**

Dr. Maurice Cogan Hauck, Assessment Development Strategic Advisor, is responsible for all aspects of ETS assessment development work on K–12 English Language Learning assessments for use in the United States, including ETS's work on the ELPAC and the CELDT. In 2014, Dr. Hauck led ETS's work on the design and development of a pool of over 2,500 test items for the ELPA21 consortium, including a large proportion of innovative, technology-enabled task types. He has also led or coled conceptualization, design, and development efforts for ETS assessment programs, including the Test of English as a Foreign Language (TOEFL) Junior Comprehensive, TOEFL Junior Standard, TOEFL Primary, and ELTeach™. Previously, he held a series of senior management positions in the ETS Assessment Division in which he was responsible for ETS's content development of assessment programs, including SAT Reasoning and the Graduate Record Examinations® General test. Before that, he spent four years managing the ETS K–12 English Language Learning group, for which he led the development of several new tests of English language proficiency, including the Comprehensive Learning English Assessment (CELLA). In addition to his work at ETS, he has 10 years of experience as a teacher of English as a Second Language and academic writing and is the coauthor of three textbooks. He earned his PhD in Language, Literacy, and Society and his MA in Applied Linguistics, both from Columbia University. He holds a BA in English Literature from the University of California at Berkeley.

### **Dr. Joyce Wang**

Dr. Joyce Wang, Senior Psychometrician, will provide oversight for the technical and psychometric tasks and issues that relate to pilot test and field test forms, sampling design, item analyses, standard setting, research studies, and other technical analyses for the California Spanish Assessment (CSA). Dr. Wang has more than seven years of experience in psychometrics, and for the past three years, she has worked as senior psychometrician in Research and Development for ETS. She is responsible for designing and implementing complex designs for the scaling and linking of K–12 EL products and large-scale state contracts. She directs and supervises research and statistical analysis activities, including item- and test-level statistical analyses and research studies on technical issues, as well as scaling and equating results. Prior to joining ETS in 2011, she was a research scientist at CTB/McGraw Hill. Before that, she worked as a psychometrician for ETS from 2002–2006. She earned her PhD in Education with an emphasis on Research Methodology from the University of California, Santa Barbara. She earned her MEng in Nuclear Engineering with an emphasis on Health Physics from the University of Florida, and she earned her BA in Nuclear Engineering from the National Tsing Hua University in Taiwan.

### **Helen McMahon**

Helen McMahon, Senior Director, K–12 Assessment, is the senior K–12 staff and content manager at ETS responsible for the overall quality of the work on K–12 large-scale assessment contracts. In this role, she oversees work done by the assessment development teams on the design of new assessments as well as the subsequent work associated with the development of test content. She has served in a variety of capacities since coming to ETS in 2004. She has worked as an assessment developer and assessment director in the science group, as well as a senior process specialist. In her current role as senior director, McMahon supervises the team of assessment directors who manage the assessment development work in the mathematics, English language arts/literacy (ELA), social studies, and science groups. She ensures that staff with the appropriate knowledge and experience are assigned to programs and is responsible for the quality of the assessments produced by the K–12 division. Before joining ETS, she taught for 14 years at the intermediate and middle school levels. She taught in a variety of configurations, including multisubject upper elementary, as well as mathematics/science and ELA/social studies team groupings. She specialized in teaching ELs and students with learning disabilities. She earned an undergraduate degree in Agriculture Science with additional studies in premed/preveterinary science at Texas State University. Most recently she has completed 33 hours toward a master's degree in Business Administration, also at Texas State University.

### **Dr. Ralph Morris**

Dr. Ralph Morris, Assessment Director, is responsible for the supervision of ETS language arts staff, as well as for the development of language arts assessments in English and Spanish for CAASPP. His experience in assessment and content development will establish comprehensive and reliable subject material. Dr. Morris has worked in language arts assessment development at ETS for ten years in roles of increasing responsibility, including as an assessment specialist, content lead, and assessment lead. Currently, he works as the language arts assessment director, where he is responsible for organizing, implementing, and distributing content assignments for contracts; managing the content

area within projects while enforcing proper processes; and serving as the point of contact regarding product quality and personnel issues. Before joining ETS, he was a middle- and high-school world languages teacher (English as a Second Language, Spanish, French, and German). He earned his PhD and MA from the University of Wisconsin–Madison with a major emphasis in Germanic Philology and Linguistics and a minor emphasis in both Romance Philology and Linguistics and German Literature; he earned a BA in Modern Foreign Languages from Lee University.

### **Jason Gonzalez**

Jason Gonzalez, Test Development Team Lead (TDTL), is responsible for overseeing the overall development schedule and process for the CSAs. Jason has over eleven years of experience working on numerous California state assessments, most notably as the TDTL on the Standards-based Tests in Spanish and the CELDT. Most recently, he helped schedule and organize the primary language, Digital Library, and fine arts stakeholder meetings. He joined ETS in 2002 and prior to the TDTL role, he was the lead editor on several large-scale assessments, including the California High School Exit Exam. His experience with California testing programs, especially with Spanish language assessments, gives him a unique insight into the intricacies of establishing a new statewide assessment for California.

### **Zulma Torres**

Zulma Torres, Program Director, is responsible for overseeing the development and implementation of the CSAs. She joined ETS in 2003, and her effective project management skills have been instrumental in the highly successful management of many programs, including CELLA, Pruebas Puertorriquenas de Aprovechamiento, and the Miami-Dade Interim Assessments. Most recently, she led the team that worked on the ELPA21 item development project. In 2007, she earned ETS's highest honor for employees for her work on the Miami-Dade project. She is fluent in Spanish and earned her bachelor's degree from Rider University.

### **Dr. Rose Payán**

Dr. Rose Payán, Strategic Advisor, Business Development and English Learner Assessment, worked with the Primary Language design team to provide guidance and information on state and national trends regarding the assessment and instruction of ELs. During her 36-year tenure with ETS, Dr. Payán has worked as a researcher on Hispanic higher education projects, and she has worked closely with state organizations in teacher licensing exams and as a government relations manager for the western region of the United States. Currently at ETS, she is the director of business development and outreach for student and teacher assessments. In this position, she is responsible for ETS's business development in several states, including Arizona, New Mexico, Oregon, and Washington. Her expertise in the assessment of ELs particularly comes in to play when she works with the assessment development and research team in developing twenty-first century assessments for ELs. Prior to her position in K–12 assessments, she worked at the Policy Research and Evaluation Center, where she focused on issues surrounding EL instruction and assessment and Latino education.

Dr. Payán's background includes speaking and organizing numerous conferences relating to Hispanic education. She was instrumental in ETS's hosting of a national conference on Science, Technology, Engineering, and Mathematics education and careers for Latinos, as well as in the ETS Achievement Gap Conference on English Learners. Before coming to ETS, she worked as a speech and language therapist, a school district bilingual program director and multicultural coordinator, and as a bilingual kindergarten teacher. She earned her PhD in Curriculum and Instruction from the University of Colorado Boulder, where she studied measurement and evaluation, curriculum and instruction, and bilingual special education. She earned her MA in Elementary Education from Claremont Graduate School, and her BA in Speech and Language Pathology from the University of Texas at El Paso.

---

## Appendix C. References

- Finegan, E., & Biber, D. (2001). Register variation and social dialect variation: The register axiom. In P. Eckert & J. R. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 235–268). Cambridge: Cambridge University Press.
- Valdés, G., Fishman, J., Chávez, R., & Pérez, W. (2006). *Developing minority language resources: The case of Spanish in California*. Clevedon, UK: Multilingual Matters.