

ELAAS: A Partnership between the California Department of Education and the Council of Chief School Officers

Table of Contents

Theory of Action..... 2
Assessment Design 7
Assessment Development Plan 18
Research and Evaluation 29
Professional Capacity and Outreach 36
Technology Approach..... 42
Project Management 45
Summary 60

ELAAS: A Partnership between the California Department of Education and the Council of Chief School Officers

Theory of Action

The **California Department of Education (CDE)** is pleased to submit this proposal to develop a new set of English language acquisition assessments to better serve Limited English Proficient (LEP) students. Working closely with the Council of Chief State School Officers (CCSSO), 18 member states, and nationally-recognized experts and organizations from across the nation, the CDE will lead a project entitled English Language Acquisition Assessment System (ELAAS) designed to enhance English language proficiency (ELP) assessment practices and use of assessment results to improve teaching of and performance of LEP students in English language acquisition and, ultimately, in core content areas.

The education community needs a thoughtful and timely effort to address the barriers to educational attainment by the fastest-growing segment of the United States student population, LEP students. If LEP students, including those with disabilities, are to have realistic opportunities to prepare for college and careers, a measured, coordinated effort that integrates their needs into the context of the implementation of the new college- and career-ready standards and emerging, aligned assessment systems must be addressed with new commitment as well as collective state and community action. This proposal is directed at that need. The theory of action driving the ELAAS consortium is a robust and coherent ELP assessment system that:

- Provides understandable, comprehensive data about LEP student performance in a timely manner
- Is accessible to a broad range of LEP students
- Improves teaching and learning for LEP students, not only in the area of ELP, but also in the academic content areas

Assessment results will be used to determine which LEP students are able to (i) meaningfully participate in instruction in the core content areas, (ii) display their understanding of the content, and (iii) comprehend and effectively employ the language features that are necessary in different academic content

areas and for college- and career-readiness. In addition to providing useful information about students' progress, the results of the ELAAS assessments will also drive instruction. The assessments will be provided within an extensive professional development (PD) framework as well as an online community of practice to allow educators to interpret results, use these results to plan for and deliver instruction in language and content to LEP students, and network with each other to share ideas and resources for teaching LEP students.

The ELAAS consortium is unprecedented in its potential to change the way in which LEP students are taught and assessed across the country. To this end, the assessments will be built on a foundation of current research and best practice to ensure meaningful outcomes and address key threats to validity that exist within current ELP assessment frameworks, including: inadequate access to tested content; the use of assessment results for purposes that are not aligned with their intended uses; a disconnect between the assessment and instructional practice; and inadequate support for educators related to the administration of the assessments and use of assessment results to inform instruction.

How the Assessment Results Will Be Used

By systematically addressing key threats to validity and ensuring that ELAAS assessment outcomes are meaningful, the assessment system will yield results that will be used by stakeholders at the state, local, school, classroom, and student levels to:

- Address the particular needs of the different LEP subgroups
- Accurately identify LEP students and each student's level of English language acquisition
- Apply an understanding of ELP levels to curriculum design and instructional practice
- Facilitate student progress monitoring
- Support data-based instructional decisions that addresses students' needs and developing skills
- Support the appropriate, full transition of LEP students into mainstream classrooms without linguistic support so that these students may access core academic content and be on par with their peers

- Inform the development of targeted strategies that address the needs of all LEP students, including recently arrived LEP students, former LEP students, migratory LEP students, and LEP students with disabilities
- Support the teaching and learning of English language acquisition necessary for LEP students to succeed in academic content classes
- Inform the PD efforts across and within states to build the capacity of all teachers, with a special focus on content teachers, to support LEP students as they develop proficiency in language and content
- Report progress by the state educational agency (SEA) and local educational agencies (LEAs) required for purposes of accountability under Titles I and III of the Elementary and Secondary Education Act (ESEA)

While it is important to focus on ELP, we recognize that the end goal of the English language acquisition process is that LEP students be able to demonstrate proficiency in core content areas (English-language arts, mathematics, social studies, and science). Therefore, it is imperative that students' progress toward being able to master the linguistic competencies inform the language instruction that takes place simultaneously within these content areas. We recognize that in order for LEP students to be taught language and content simultaneously, all teachers (including content teachers) must be provided support in order to do so. As such, this collaborative project will place an emphasis on including content area teachers in PD efforts, as these teachers may be unaware or misinformed of how to use ELP assessment results to guide instruction for their LEP students. The ELAAS will pay special attention to building a shared sense of responsibility among all stakeholders (including all teachers and administrators) for teaching LEP students and will focus on providing targeted PD to content teachers who work with LEP students.

How the Assessments and Assessment Results Will Be Incorporated Into Coherent Educational Systems

English language acquisition has historically been seen as an isolated process that has relevance to LEP students' success in other content areas but which is structured around language proficiency standards that may only loosely be related to content standards. While ELP standards are meant to be the on-ramp to content area standards, ELP standards are often ignored as their correlation with content standards is poorly understood by teachers. To address this issue, correspondence between the ELP standards, the ELAAS assessments, and the college- and career-readiness standards will be established and inform PD design as well as assessment design and development. The ELAAS will be situated within a solid framework of teacher capacity development to ensure that the system is fully understood and implemented with fidelity in order to improve LEP student achievement in the content areas.

Consistent with this theory of action and in order to ensure that assessment results will be incorporated into and cohere with state educational systems, the ELAAS assessments will:

- Be aligned to a set of quality ELP standards agreed to by the member states
- Correspond to the college- and career-ready standards
- Incorporate aspects of language necessary for students to meaningfully engage with and master content at each grade level across the core academic subjects
- Complement and be compatible with the assessments being created by the Partnership for Assessment of Readiness for College and Careers (PARCC) and the SMARTER Balanced Assessment Consortium (SBAC)

Many states under the current ESEA laws use ELP assessments to determine proficiency levels, placement and annual growth of English proficiency. However, across states there is inconsistency in the proficiency levels, proficiency level descriptors, and definitions. With a very mobile and diverse LEP population we recognize the need to have common definitions and proficiency levels from this work that will apply to LEP students regardless of where they attend school within any given state or district. Also, these measures will correspond to the college- and career-ready standards recently adopted by the members of this consortium in order to provide comparable information across states and districts. While

it is important to focus on ELP, we recognize that the end goal of the ELP process is that students will have amassed proficiency in listening, speaking, reading, and writing that will allow them to demonstrate proficiency in core content areas. Unlike some previous efforts, this connection between a complex system of standards and assessments will directly tie ELP assessments to the overall educational effort in schools, ensuring LEP students form a central part of educational reforms, not an afterthought.

How Those Educational Systems as a Whole Will Improve Student Achievement

Our ultimate goal through creating this system is to improve achievement for LEP students. The ELAAS will do this in the following ways. First, the assessment will be a more accurate representation of LEP students' language abilities, which will provide more accurate data for all stakeholders in LEP students' education, including their parents. Second, because the correspondence to college- and career-readiness standards is a central part of the design, this assessment will ensure successful access to the academic concepts while students progress toward attaining English proficiency. Finally, because we are focused on providing meaningful data to teachers in a timely, understandable format, coupled within extensive PD and a community of practice, this system will positively influence the quality of classroom instruction for LEP students.

While we understand the challenges of creating a new ELP assessment system, we are resolute that the time has come for better measures of English proficiency. The decisions that will be made throughout the design process will be rooted in the focused principle of improving LEP achievement. With that focus, we will produce a new assessment system that will provide better teacher support and increase student achievement to more robust standards. In addition to better training and guidance, all teachers will receive better information on their students' English proficiency and how it relates to accessing instruction. By using student performance outcomes on the ELP assessments in conjunction with other student performance indicators, states, districts, schools, and classrooms will be able to target

professional development for teachers and administrators in order to inform and improve instructional decisions and practice for LEP students.

Assessment Design

The purpose of this project is to develop an ELP assessment system that provides reliable and valid measures of students' level of English proficiency in the four domains of reading, writing, speaking and listening for LEP students in kindergarten through grade 12 for diagnostic and summative purposes. The English Language Acquisition Assessment System (ELAAS) will include two major assessments, a Screener/Diagnostic, and a Summative both of which will be aligned with ELP standards that are common across the participating states and correspond with college- and career-ready standards. In this section of the proposal, we present a summary plan of the system and then discuss the system with respect to the selection criteria set forth in the proposal.

System Description

This system will be developed in accordance with the project's theory of action and by a team of experts with many years of experience in ELP test development and assessment of ELP. The system is heavily based on current research on the development, validation, and implementation of ELP assessments in the nation and around the world. As indicated in the project's theory of action, the ELAAS will be a robust and commonly shared system of ELP assessments that is based on the concept of universal design and provides comprehensive data about student performance in a timely manner that will improve teaching and learning. The system will generate a wealth of data that can be shared by the participating states for assessment, accountability, and research, and provide important data for informing curriculum planning and instruction in order to meet LEP students' academic needs.

The mandate of ESEA ELP assessments prompted researchers to conduct studies and discuss issues on the assessment and accommodations of LEP students and to examine content and psychometric characteristics of these assessments. This project which aims the advancement of ELP assessments that

correspond to the college- and career-ready standards will greatly benefit from the outcomes of these studies and will incorporate research-based recommendations into the test development process.

The system moves beyond existing ELP assessments in several respects: (1) the ELAAS is a comprehensive system with components that will provide assessment outcomes, diagnostic information, and information to improve instruction and curriculum planning; (2) It includes a comprehensive set of items in different types and formats including performance assessment items to ensure LEP students can demonstrate what they know and are able to do; (3) It is a computer-based system that allows the use of innovative technologies in test item presentation, provision of accommodations and scoring; (4) It provides a comprehensive data system for use by the participating states for research and test development purposes; and (5) It generates comprehensive sets of user friendly reports that can help teachers and parents to understand students' strengths and weaknesses in areas that are essential for student academic success.

The initial stage of the assessment design process includes development of test blueprints based on the ELP standards that correspond to college- and career-ready standards. In the pilot testing phase, the process of *think aloud* and *cognitive labs* will help identify test items that may not clearly measure the intended construct. The pilot test phase also includes bias and sensitivity reviews and alignment of the test content with the ELP standards. The field test phase includes careful examinations of the content and psychometric characteristics of the items and the tests. The field test will also include examination of items that may be biased against or in favor of certain groups of LEP students through the application of differential item functioning (DIF) methodology. Operational test forms will be constructed for each of the two assessments (screener/diagnostic and summative forms) in each of the five grade clusters (i.e., kindergarten, grades 1 and 2, 3 to 5, 6 to 8 and 9 to 12). Standard setting procedure will be conducted to establish separate sets of cut scores for each of the 13 grades (K through 12).

The use of grade clusters rather than specific grade assessments provides greater efficiency in item development, professional development for teachers and administrators, and for test administration.

The system will be developed under the supervision of the projects director and in consultation with CCSSO national experts working on specific tasks and the project technical advisory committee in the ELAAS consortium. The collaboration provides opportunity to share resources for conducting analyses that can inform future development and improvements in the system and will also bring considerable cost savings. For example, major studies on the impact of students' background variables can be conducted by using data from multiple locations provided by the participating states in the consortium. In addition, data from the multiple sites can be disaggregated by years in a language instruction educational program, by disability status, and by student's native language, among other categories.

While the system can be presented in a paper-and-pencil mode or in a computer-based mode, the computer-based mode is preferable because it will:

1. Standardize administration of the assessment across multiple states
2. Provide unique opportunities to present accommodations to LEP students with disabilities that are recommended in their Individualized Education Program (IEP) and Section 504 plans
3. Provide computer scoring capabilities that would be difficult to apply in the paper-and-pencil mode
4. Be consistent with the test administration of the future state assessment system (e.g., PARCC and SBAC)
5. Provide rater training that can be facilitated through computer-based practice and feedback
6. Allow the use of innovative item types
7. Assess ELP Standards that are difficult to measure using paper-and-pencil mode
8. Allow adaptive or staged-adaptive delivery of ELP assessments

The main concern with the computer-based system of delivery is the "digital divide" (i.e., students and parents at different socioeconomic status may have different level of exposure with computers).

While this issue may be viewed as a possible source of threat to the validity of computer-based system, we believe that by the time of implementation of the system, the factor will have less impact due to

increased access to and familiarity with computer-based assessments by teachers, students and parents.

To assist those sites in transition to a computer-based mode, a paper-and-pencil version will be available.

Addressing the Priorities and Selection Criteria

The project addresses three of the five Absolute Priorities for the Enhanced Assessment Grant (EAG) program by:

- (1) *Collaborating with institutions of higher education, other research institutions, or other organizations to improve the quality, validity, and reliability of state academic assessments beyond the requirements for such assessments described in section 1111(b)(3) of the ESEA.* The consortium's leadership team includes national experts in the assessment and accommodations of LEP students from high ranking institutions of higher education and nationally known research organizations and experts from the ELAAS member states. These organizations include CCSSO; CDE; The George Washington University Center for Equity and Excellence in Education; University of California, Davis; WestEd; TESOL; Metrica Research Associates; and the National Center on Educational Outcomes. The leadership team directly addresses issues concerning the content and psychometric quality of the assessment system.
- (2) *Charting student progress over time to provide data to inform teachers in their instruction and curriculum planning and to provide diagnostic and summative data for accountability purposes over time.* For example, LEP students will be assessed annually to measure their progress toward acquiring ELP (Annual Measurable Achievement Objective (AMAO1)) as well as their attainment of ELP (AMAO2).
- (3) *Developing an English language proficiency assessment system.* The ELAAS which will be developed under this project will provide valid, reliable, and fair assessments for measuring LEP students' level of English proficiency in four domains of English language.

Below is a detailed description based on the *Selection Criteria* defined in the Application for the Enhanced Assessment Grant Program under the "Assessment Design" section.

Number and Types of Assessments

The ELAAS consists of two assessments, a screener/diagnostic and a summative assessment. This comprehensive system has many advantages over the existing ELP assessments developed based on ESEA Title III guidelines. The proposed ELAAS will produce assessments that serve multiple purposes. The outcomes of these assessments can be used to improve curriculum planning and instruction. The use of grade cluster assessments that are vertically articulated will provide additional information to teachers and administrators. Using the vertically articulated ELP assessments will add the potential of using the results of summative assessments to inform teachers about students' areas of strengths and weaknesses. Results of assessments that are vertically articulated can also be used not only for curriculum planning but may also be used by teachers to plan their instruction based on individual student's knowledge and understanding of English. The outcomes of ELP assessments that are used diagnostically can help identify areas that LEP students need assistance in to succeed in core academic subjects and the corresponding assessments. Use of screener/diagnostic ELP assessment outcomes become increasingly important in helping teachers understand LEP students' instructional needs.

Measuring Student Knowledge and Skill

The ELAAS will measure student knowledge and skills against the full range of ELP standards that are common across the participating states and that correspond to the college- and career-ready standards in mathematics, English/ language arts and literacy in history/social studies, science, and technical subjects. The participating states in the ELAAS consortium, including California as the lead state, have had well established and operationally defined ELP standards for many years. Our preliminary review of the ELP standards from the participating states as well as the national TESOL ELP standards shows a substantial commonality among state ELP standards. The first step in assessment development will be to define a common set of ELP standards in collaboration with the participating states, and based on the work of researchers, and other organizations working on this project who are defining national ELP standards. This common set of ELP standards will be used as the basis for developing the test blueprints and later as the basis for item development.

Students who are fluent in basic interpersonal communication skills (BICS) but not proficient in cognitive academic language proficiency (CALP) may be at the risk of academic failure and may not be college- and career-ready (Abedi, 2008b; Bailey & Butler, 2003; Cummins, 2000). In the context of assessments, language proficiency tests could vary in the extent they gauge CALP. Therefore, it is essential that ELP tests correspond to the college- and career- ready standards and measure the type of language proficiency needed to be successful in the academic content classrooms. The term *correspondence* indicates that ELP assessment items should include the language that facilitates content and learning not to be aligned with the technical academic content. Thus, the ELP assessments should not be a measure of students' content knowledge (e.g., math, science, English-language arts) but rather a measure of English proficiency that facilitates content learning.

The test blueprint will be developed based on the ELP and correspond to college- and career-ready standards. The blueprints will detail the standards appropriate to assess as well as the number and types of items.

The ELAAS consortium will develop a common set of performance level descriptors (PLD) to describe what students know and can do at each achievement level by grade and domain. The PLDs will be developed using input from teachers, administrators and linguistic experts. The PLDs will later be used to inform the standard setting process and may at that time be updated.

Required Student Performance Data

Student performance data will be produced by the ELAAS in different forms including raw and scale scores as well as performance levels. A standard setting process (e.g., bookmark and map mark) will be utilized to establish cut scores for determining achievement levels (e.g., Pre- Emergent, Basic, Intermediate, and Proficient) (National Research Council, 2011). The standard setting process will include a diverse and representative group of teachers, administrators and other key stakeholders from the member states who have expertise in areas such as language acquisition and serving LEP students with disabilities. The cut scores obtained through the standard setting process will be validated using external

criteria such as students' scores on states' current ELP and reading language arts assessments and teacher evaluations of students' ELP levels.

Furthermore, reliability of both the ELAAS Summative and the screener/diagnostic assessments in all four domains will be examined using internal consistency and parallel form approaches. The use of the internal consistency approach for estimating reliability is based on the unidimensionality assumption. Therefore, the dimensionality of both the ELAAS components will be examined by conducting principal components analyses. To further examine the dimensionality of the test items and the domains of the ELP assessments, a confirmatory factor analytic model will also be used. The concept of unidimensionality of the ELP assessment is important, not only as a requirement for computing internal consistency reliabilities but, as a major condition in creating an overall ELP composite score. If the data support a unidimensional trait of English proficiency, a *compensatory* model in which a composite of all four domains will be used. However, if the subscale scores are not unidimensional, then a *conjunctive* model based on both subscale scores and total scores will be used.

Student Data to Inform Instruction Interventions, and Professional Development

The ELAAS diagnostic assessment will be given to students who were identified as potential LEP students through information gathered on a home language survey (HLS). The outcomes of this assessment will be used for initial identification and placement of LEP students. The outcomes of this assessment will also be made available to teachers to improve instruction and curriculum planning. The ELAAS diagnostic assessment may be re-administered during the academic year to provide additional information to teachers.

The ELAAS summative assessment will be administered at the end of the academic year. Results of this assessment will be used as a measure of student learning, accountability, and reclassification. For example, the summative assessment results will be used to measure LEP students' progress toward acquiring ELP (AMAO1) as well as their attainment of ELP (AMAO2).

Types of Assessment Data

Both the ELAAS assessments (screener/diagnostic and summative) will provide data on student achievement in the four domains along with a *Comprehension* score (a composite of reading and listening) and an *overall* composite (a composite of all four domains). The data from the ELASS diagnostic assessment will be used for identification and placement. The ELASS summative assessment will be used for accountability, end of year assessment purposes, and reclassification.

Research has shown that in a compensatory model, LEP students may reach proficiency by being proficient in some but not all of the ELP domains of reading, writing, speaking, and listening. For example, LEP students who are above proficient in listening and speaking but below proficient in reading and writing are deemed proficient in English because their high scores in listening and speaking compensate for their low scores in reading and writing (Abedi, 2008a). This consortium will study methodologies for creating composite scores and adopt a policy based on the results of that study to ensure the validity of reporting ELP assessment results.

Uses of Data for Determining Student Achievement and Progress and Informing Teaching, Learning, and Program Improvement

Data from the ELAAS will serve multiple purposes. The data will be used to determine students' progress in learning English as a language to access academic content and inform decisions on student placement and reclassification. As appropriate, student data may also be used as one of multiple measures to evaluate individual principal and teacher effectiveness.

The data will also be used to inform teachers in developing more effective instruction for LEP students and for school officials to design professional development that will allow teachers to better help students master the four English language domains. For example, results of studies nationwide suggest that LEP students are more proficient in listening and speaking than reading and writing (Linguanti, 2011). Such information may help teachers to put more emphases on areas that students need more help. The project intends to share results of ELP assessments with the parents in ways that could help them understand their children's progress. Focus groups will be conducted with parents to explore the best

format to present ELP assessment results. The ELAAS will provide professional development modules for all teachers (including academic content teachers) on how to best use the ELAAS assessment results.

Frequency and Timing of Administration

As indicated earlier, the ELAAS includes two assessments: (1) a diagnostic assessment will provide information for identification and placement and (2) the summative assessment will be used to monitor student progress, accountability, and reclassification. The ELAAS diagnostic assessment will be administered at the start of a school year and may be re-administered as needed. As indicated earlier, information from the ELAAS diagnostic section could provide valuable information for all teachers. Therefore, multiple administrations of the diagnostic assessment may be necessary to provide information for planning and improving instruction. Based on current research and industry standards, the ELAAS consortium will analyze and recommend suggested frequency and timing of administration of the diagnostic assessments. The summative assessment only will be administered once a year.

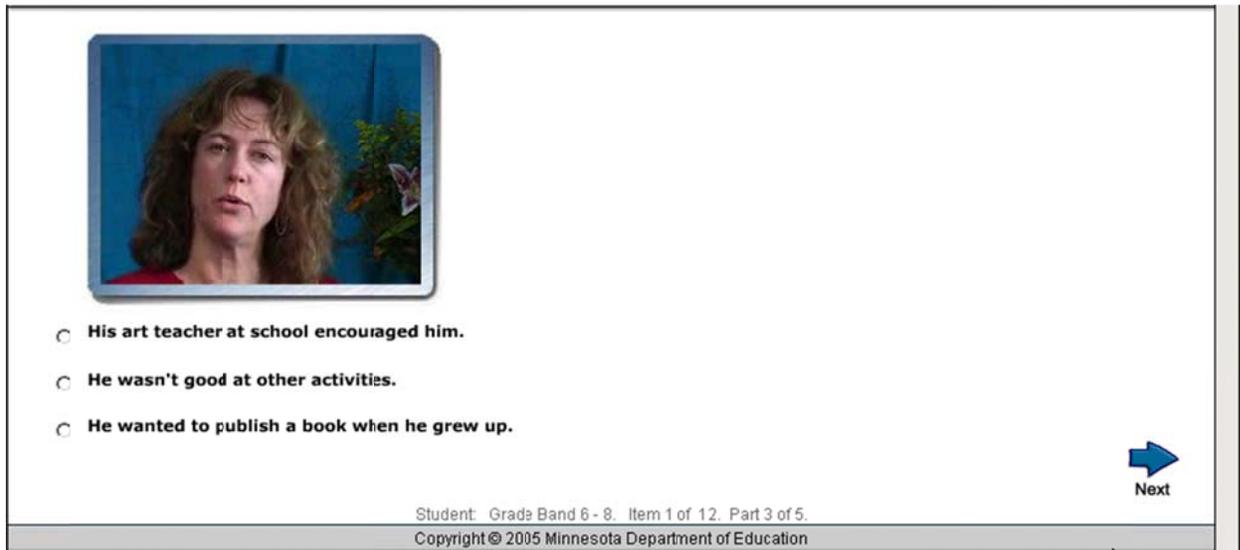
The Number and Types of Items

Different ELP domains may need different item types including, multiple-choice and constructed response items. To the extent possible, performance-based items (i.e. constructed response items and dichotomous-constructed response items) will be used across all four domains. A recent publication (Abedi, 2010) describes the importance of performance assessments (PA) and how they increase the validity of the assessment. Performance-based items may provide more opportunity for LEP students to present what they know and are able to do. These types of items encourage more participation and reduces guessing. The performance-based item development will be guided by current research on the assessment of LEP students.

Sample Items

Multiple Choice

In the sample below, students are asked to listen to a recording in which information about a given topic is presented, and then choose the sentence below the picture that provides a true statement about the information they heard.



His art teacher at school encouraged him.

He wasn't good at other activities.

He wanted to publish a book when he grew up.

Next

Student: Grade Band 6 - 8. Item 1 of 12. Part 3 of 5.
Copyright © 2005 Minnesota Department of Education

Performance-Based

The sample below requires students to provide an oral narrative based on a sequence of events that are viewed via a short video clip. This item type has the potential of eliciting a quality student response that is an accurate measure of student proficiency due in part to the detail that may be provided in a video clip.

AUDIO: You are going to see a video. Watch the video. It tells a story. I will start the story for you. One day, Molly and John were in front of the school. Tell me what happened. [The student's verbal response will be recorded and scored.]

The Assessments' Administration Mode

The ELAAS project proposes a computer-based mode for several reasons: (1) the two Race to the Top Consortia (SBAC and PARCC) are committed to a computer-based mode for their content-based assessments and it is important for the ELP assessments to be consistent with the consortia that are

preparing content assessments; (2) given time and budget restrictions, it is important to focus on a mode of test administration that has greater potential in future generations of the assessments; (3) the computer-based mode offers more flexibility and more innovations in terms of test administration capabilities and the ability to present assessment in a more universally designed manner; (4) the computer-based system provides more opportunities for presenting multiple forms of common accommodations for LEP students with disabilities (Rivera, et al. (2006); Shafer Willner, et al., (2008). In addition, a computer based mode offers opportunities to provide common accommodations more consistently.

There are however, some limitations with the computer-based assessments. Issues raised by researchers on the concept of *digital divide*, students' and families' socio-economic background and students' access to computer at home are worthy of consideration (Goode, 2007; U.S. Department of Education, 2008). However, as access to technology continues to widen across all schools in the country, we believe these issues may have less impact by the time the ELAAS test is ready for administration. The ELAAS project team will also prepare a paper-and-pencil version of the test and study comparability between the computer version and the paper-and-pencil version with additional support later in the process.

Scoring

The computer-based mode will allow for selected-response items to be scored in real time. Other item types will be scored according to rubrics developed in conjunction with the development of those items. To ensure scalable, accurate, and consistent scoring of items, any person scoring the ELAAS (including teachers) must have completed the ELAAS scoring certification course. The course will be developed after rubrics are finalized and exemplar responses for each item type are determined. The course will include a calibration component requiring participants to obtain a minimal percent perfect agreement and not exceed a pre-set maximum score bias in either direction (scores on a reference set of responses will be set by a group of expert raters). Opportunities for retraining will be offered on a regular basis, and the course can be deployed in a secure on-line system.

Reports

Reports for these assessments will be generated for students, teachers and parents. A series of focus groups will be conducted to find the best format for reporting the data to maximize understanding and usefulness. In the focus groups, attention will be given to the readability of the report (e.g., the clarity of language and score reporting). The student and parent reports will be translated to those languages with the highest frequency of use by students.

Assessment Development Plan

The ELAAS is designed to provide valid, reliable, and fair assessment outcomes for LEP students. The ELAAS will recognize and will be sensitive to the heterogeneous and diverse nature of the LEP student population. The assessment will be designed to be a valid, reliable, and fair measure for the diverse LEP student population including a variety of subgroups such as students whose formal education has been interrupted.

Approach for Developing Assessment Items

To account for diversity within the LEP population, the assessments will use Universal Design Principles (e.g., Accessible Test Design) in order to maximize the measurement of targeted constructs for each student and support more valid inferences about student language proficiency based on test scores. Accessible Test Design will be consistent with the widely accepted seven elements of universally designed assessments identified and described by the National Center on Educational Outcomes (Thompson, Johnstone, Thurlow, 2002) as elaborated below.

1. **Inclusive assessment population.** During the pilot and field testing, the ELAAS will recognize the needs of diverse subgroups within the LEP population by including a variety of items with different item formats (i.e., multiple-choice, constructed-response, performance-based), at different levels of difficulty (to provide discrimination power for groups of LEP students at different levels of proficiency), and by including a broad range of students with a variety of linguistic and cultural background. Specifically, the ELAAS field test sample will include all subgroups within the LEP population that are expected to take the ELAAS operational forms,

including LEP students with a range of disabilities who may participate in the content assessments with or without appropriate accommodations.

2. **Precisely defined constructs.** The ELAAS will measure student knowledge and skills against the full range of ELP standards that are common across the participating states. A preliminary review of the ELP standards from the participating states as well as the national TESOL ELP standards shows a substantial commonality among state ELP standards. The first step in assessment development will be to define a common set of ELP standards in collaboration with the participating states, and based on the work of researchers, and other organizations working on this project who are defining national ELP standards. This common set of ELP standards will be used as the basis for developing the test blueprints and later as the basis for item development. . The ESEA Title III requires ELP assessments to include four domains of English proficiency: reading, writing, speaking, and listening. It is important to note at this point that in the process of defining the focal construct for the ELAAS, all construct-irrelevant sources will be identified and removed, as defined in the Standards for Educational and Psychological Testing, “the degree to which test scores are affected by processes that are extraneous to its intended construct” (AERA, 1999; Thompson, et al, 2002).
3. **Accessible, non-biased items.** The ELAAS will use the principle of accessibility of assessments as elaborated by Elliott, Kettler, Beddow, and Kurz (2011) by incorporating the bias and sensitivity review process into the assessment development phase, by revising items that are affected by students’ linguistic and cultural background (irrelevant to the ELAAS focal construct) and by conducting differential item functioning to identify items that may introduce bias into the assessment process. After identifying items that differentially perform across subgroups of LEP students (by different backgrounds mentioned above), the items will be carefully reviewed and will be revised if possible or removed if not revisable.
4. **Amenable to accommodation.** Research has identified many sources of threat to the effectiveness and validity of accommodations for students with disabilities (Thurlow, Lazarus,

Thompson, & Robey, 2002) and LEP students (Abedi, 2007). Due to such validity issues, some LEP students with disabilities may be excluded from ELP assessments even under the Universal Design Principles. Thompson, et al. (2002) indicated that “The goal of universal design in such cases is to facilitate the use of appropriate accommodations and to reduced threats to validity and comparability of scores” (p.11). Based on this recommendation, the ELAAS will create assessment items in a manner to be amenable to accommodation. For example, as suggested by Thompson, et al. (2002), to make assessment amenable to Braille, in the ELAAS test development process, the “use of construct irrelevant graphs or pictures, use of vertical or diagonal text, keys and legends located to the left or bottom of the items...” will be avoided. The provision of extra time and multiple breaks as two of the most commonly used accommodations for students with disabilities can easily be incorporated into the ELAAS computer-based system. The authors also indicated the “Performance-based items that require the use of manipulatives can be more amenable to accommodations if alternate response modes are designed from the beginning and included in field test: (p.12). This is quite applicable to the ELAAS as it will include a variety of performance-based items.

5. **Simple, clear, and intuitive instructions and procedures.** As a computer-based system, the ELAAS will provide standardized, clear test instructions. At the pilot testing phase, a series of think-aloud and cognitive lab sessions will be conducted to ensure the clarity of instructions and procedures for the ELAAS test administration. The instructions will be set in a clear manner so that students can work independently throughout the assessment. As recommended by the Standards for Educational and Psychological Testing, (AERA, 1999, p. 47), sample questions and a representative item identified with each major area (listening, speaking, reading, and writing) will be provided to test takers prior to the administration of the assessment.
6. **Maximum readability and comprehensibility.** In the ELAAS item development process, efforts will be made to make assessment items clear and readable to the extent such efforts will not affect the focal construct. In its item development process, the ELAAS will include a series of think-

aloud and cognitive lab sessions in which representatives of the test takers (from all K-12 grades and from all subgroups of the LEP population) will assist in identifying test items that are not clear and readable. The outcomes of these sessions will help test item writers to revise items that are unclear and are not linguistically accessible.

7. **Maximum legibility.** Thompson, et al. (2002) defines *legibility* as the capability of an assessment being deciphered with ease. It includes legibility of text, legibility of graphs, tables, and illustrations and legibility of response format. The ELAAS will follow recommendations from research focusing on legibility within the Universal Design Principles (Thompson, et al., 2002; Hanna, 2005) to improve legibility of the items to the extent it does not interfere with the ELP construct measured by the assessment.

Application of the Evidence-Centered Assessment Design in the ELAAS Test Development Process

The ELAAS test development process will be guided by the principles of Evidence-Centered Design (ECD). ECD is an evidence-based comprehensive model for the design of assessments and is defined by three components: (1) student or claims, (2) evidence, and (3) tasks. (Almond, Steinberg, & Mislevy, 2002; Mislevy, Almond & Lukas, 2004). The claims component asks “What claims will be made about student performance on the ELAAS?” Similarly, the evidence component asks “What evidence is needed to support those claims?” and the task component asks “What tasks can be designed to gather the evidence needed to support those claims?” The *Delivery System Model* then brings the three components together and explains the collection of student evidence, task, assembly, and presentation models necessary for the assessment and how they will work together (Mislevy, Almond & Lukas, 2004).

Types of Personnel

Different aspects of the ELAAS test development process involve different expertise. The ELAAS test development process involves experts in the following areas in each component of test development:

1. Agreeing upon a common set of ELP standards that are derived from standards work currently underway across our participating states, research organizations and experts in the field.

2. Creating test blueprints: this phase involves assessment and psychometric experts, linguistics experts, academic content experts, experts in second language acquisition, cognitive scientists, experts in career readiness standards.
3. Developing assessment items: this phase includes teachers, assessment and psychometric experts, linguistics experts, academic content experts, experts in second language acquisition, cognitive scientists, experts on college- and career-readiness standards.
4. Pilot testing of the items: this phase needs experts in qualitative research (think aloud and cognitive labs), research methodology experts, assessment and psychometric experts, teachers, parents.
5. Field testing: research design and sampling experts, this phase includes ELP and content experts, psychometricians, and language development experts.

Approach and Strategy for Accommodations

Many of the students with different types of disabilities participate in the content assessment and accountability system with accommodations that are specified in their IEP or Section 504 plans. LEP students with disabilities are no exception. Therefore, the ELAAS will include accommodations for LEP students with disabilities. Research on accommodations for LEP students with disabilities reveal major issues on the effectiveness and validity of some of the accommodations currently used for these students (Thurlow, Lazarus, Thompson, & Robey, 2002). The decision on selection of accommodations for LEP students with disabilities in the ELAAS will be based on the current studies on accommodations. However, results of some of these studies are not conclusive. Therefore, the ELAAS will group the commonly used accommodations for these students into three categories: (1) USE which includes accommodations that are supported by research, (2) NOT SURE, accommodations for which the research findings are mixed and not conclusive, and (3) DO NOT USE, accommodations that clearly alter the construct of ELP. The ELAAS will recommend that participating states use the first category of accommodations (USE) whenever applicable, not use the third category of accommodations (DO NOT USE), and use the second category of accommodations (NOT SURE) with caution and a word of

acknowledgement. In case of no other alternative except the use of the third category of (DO NOT USE) accommodations, the ELAAS will recommend not combining assessment outcomes under such accommodations with the assessment outcomes from the content assessments.

ELAAS recognizes that a small percentage of ELLs with disabilities have significant cognitive disabilities which require that they participate in an alternate assessment based on alternate achievement standards (AA-AAS). Two General Supervision Enhancement Grant (GSEG) consortia currently are developing AA-AAS for reading and mathematics assessments, and will be sharing their knowledge about best practice in developing these assessments for students with significant cognitive disabilities, including those who ELLs. Because the ELAAS project is lagged one year behind the GSEG projects, it will benefit from the information disseminated by those projects. In addition, ELAAS plans to implement the following strategies for creating an ELP assessment appropriate for these students: (1) conduct an inventory of the numbers and characteristics of ELLs who have significant cognitive disabilities, (2) identify learning pathways or entry points for the English proficiency content standards, (3) determine best approaches for assessing students (i.e., performance tasks, body of evidence, etc.), including allowable accommodations, and (4) develop assessment materials based on approach(es) states agree on for the common alternate ELP assessment.

Approach and Strategy for Ensuring Scalable, Accurate, and Consistent Scoring of Items

The computer-based mode will allow for selected-response items to be scored in real time. Other item types will be scored according to rubrics developed in conjunction with the development of those items. To ensure scalable, accurate, and consistent scoring of items, any person scoring the ELAAS (including teachers) must have completed the ELAAS scoring certification course. The course will be developed after rubrics are finalized and exemplar responses for each item type are determined. The course will include a calibration component requiring participants to obtain a minimal percent perfect agreement and not exceed a pre-set maximum score bias in either direction (scores on a reference set of

responses will be set by a group of expert raters). Opportunities for retraining will be offered on a regular basis, and the course can be deployed in a secure on-line system.

For both the diagnostic and the summative assessments, all responses will be collected and stored as data for continuous improvement of the rater certification system, for validation research purposes, and also for the potential deployment of an automated scoring system. The diagnostic assessment will be scored by the ELAAS-certified test administrator including teachers, and those scores also collected and associated with the raw responses. This way, the results of the diagnostic assessment can be reported immediately, but also retained for future use.

The summative assessment will not be scored by the test administrator. Instead, the raw responses will be routed to a distributed scoring system where ELAAS-certified raters (including teachers) will score the responses. Responses will be independently scored. A rule for score resolution will be developed, and a percentage of responses will receive a second score to ensure inter-rater reliability. A procedure based on calibration samples for attending to potential rater drift will alert the scoring manager to any problems with a particular rater. Rater-item-blocks will be small enough to avoid potential halo effects which can create unwanted rater-by-respondent dependencies.

The Approach and Strategy for Developing the Reporting System

The ELAAS will include a comprehensive reporting system including report templates. Score reports for the ELASS summative assessment include vertically scaled scores across the K–12 grade span in each domain (listening, speaking, reading, and writing), and an interpretation of the scale score designated as an English proficiency levels. The decision on the number and description of levels will be made collectively by the ELAAS consortium. After this decision is made, the cut scores will be identified based on the standard setting procedures (e.g., bookmark or mapmark). The cut scores will then be validated by using external criteria such as states' current ELP assessments, content assessments, and teacher evaluations of student performance. The diagnostic and the summative assessment reports may include additional information, depending on the member states' needs. The decision on the number and type of composite scores will be made collectively by the members of the ELAAS consortium. The

ELAAS consortium will determine the relative weights of each of the domain scores going into the composite scores based on input from the technical advisory committee.

The reporting system will include information from the ELAAS assessments. The ELAAS reporting system will be designed based on the following sources of information: (1) sample reports from the participating states in the consortium, (2) information from focus groups and interviews with teachers, parents, school principals and students, and (3) information from national sources and national/international assessments such as NAEP, TIMSS and PISSA. Such information will be collected and will be used by the ELAAS reporting committee for developing a comprehensive reporting system.

The Overall Approach to Quality Control and Strategy for Field-Testing

Since the ELAAS is a computer-based assessment system, the quality control process in the ELAAS includes procedures that must be completed by the test developer and/or by the participating state assessment staff or local educational agencies before the assessments are administered to ensure that schools' computer facilities are set for the proper administration of the system. The quality control system also includes assurances that test administrators have proper training for administering the assessments.

Following the item development phase, test items will go through the process of pilot test and later the field test. Among the important components of the pilot tests are bias and sensitivity review and think-aloud and cognitive lab procedures. The outcomes of these procedures will help identify test items that may be culturally and/or linguistically biased or may not be accessible due to the construct-irrelevant sources thus will need to be revised. The field test will be conducted to learn about how the items function by collecting and analyzing item-specific data. A different collection of items for each domain (reading, writing, speaking, and listening) will be constructed in each of the five grade clusters (K, 1-2, 3-5, 6-8, and 9-12). Each collection will be assigned to one of two field-test forms (A, B) so that each form reflects, as closely as possible, the final test blueprints. Item specific psychometric characteristics such as DIF, index of item difficulty, and item discrimination will be obtained for test items in the two forms for each of the ELAAS assessments (screener/diagnostic and summative). These characteristics will help identify content and psychometric issues in the items. The field test data will also provide information on

the overall test reliability and criterion-based (concurrent) validity. Reliability of the ELAAS screener/diagnostic and summative assessments will be examined by the internal consistency approach.

Dimensionality of test items and ELP domains will be tested and evidence suggesting unidimensionality or multidimensionality of items and domains can help later decisions on creating composite scores and reporting the results. Given the limitations of internal consistency approach of being sensitive to multidimensionality of items within domains and across domains, a parallel form approach (Forms A and B) will also be used for estimating reliability of the ELAAS assessments. Data from the ELAAS field test will be used also for latent-class modeling. Five-class models will be fit to item response data from each of the two field-test forms. For each domain and grade-cluster form, we will estimate the proportion correct for each item within each latent class. To evaluate the validity of items for discriminating among the latent classes, we will calculate the differences in proportion correct between adjacent classes. This approach will shed light on the vertical articulation of the ELAAS components.

To ensure a representative field test sample of all LEP students the ELAAS consortium will use a stratified random sampling procedure in which LEP background characteristics will be considered. The stratification variables will include all those categories mentioned above plus new categories that will be identified by reviewing the research. A power analysis will be conducted to determine the sample size needed to address subgroup issues in the field test.

Summary of the Assessment Development Process

Below is a summary of our plans for the ELAAS test development:

Phase I: Development of an ELP Assessment Framework

An ELP assessment framework outlines the four basic domains of ELP (listening, speaking, reading, and writing), and the dimensions of performance within each of these domains. The process for developing a joint ELP assessment framework begins by collecting and reviewing ELP standards documents from each participating state. Based on participating states' ELP standards and other work being conducted by research organizations, a common set of ELP standards will be used as the foundation

of the assessment framework. The assessment framework will describe the correspondence between the ELP and the college- and career-ready standards.

Phase II: Development of a Table of Specifications and Item and Test Form Specifications

The blueprints for the ELAAS will describe the distribution of weighting for different ELP standards, which covers dimensions such as mastery of vocabulary and grammar, syntax and complexity of language production (for writing and speaking), and comprehension of different forms of spoken and written language. The weights assigned to ELP standards reflect the relative value of each standard in assessing ELP at each grade cluster. The Steering Committee, with input from the ELP language experts from each member state and the technical advisory committee, will determine these weights. Test blueprints will be developed for each assessment by grade cluster.

The ELAAS will consist of several item types. Before items are developed, content experts and psychometricians will draw up a scheme for a defensible weighting across item types for each standard. Typically, each selected response supports one point, and other item types can support a range of up to about 5 or 6 points. Consideration regarding adequate reliability, number of targeted cut points, and testing time constraints, will result in targeted numbers of items and therefore test length and test form point totals.

The ELAAS will provide an opportunity to use some item types that are new to ELP assessment such as technology enhanced items. Prior to production of these item types, small pilot studies will be conducted to ensure that the item types are collecting evidence of performance on the intended constructs. While such studies are best practice in assessment development, this work is all the more important on the ELAAS where students may be new to formal education or have limited exposure to large scale assessment, especially that which will use technology. We will develop study designs that attend to these and other construct validity issues to ensure that the item types provide the maximum – and valid – information.

In addition, the ELAAS will develop accommodations and access opportunities at item inception rather than after form development as is commonly the case in large scale assessment. Adhering to

principles of Accessible Test Design will ensure that the ELAAS remains committed to inclusion of all students in item and test design. As detailed in the Technology section of this proposal, such accommodations and access considerations at the beginning of development will lead to more valid inferences about students' ELP levels.

Phase III: Production of Test Form Drafts

It is useful at the beginning of a testing program to develop a base test form that includes all item types, rather than to develop many different versions of an item type in isolation and then later combine these into a test form. Test forms will embody all characteristics determined in the test blueprints (for ELP standards that correspond to college- and career-readiness standards) and in the test form specifications (for item types by ELP standard). Linguistic experts will make a determination of vocabulary and concepts that are not content specific but do facilitate content learning to inform item writing. A style guide will be developed for item writers to follow. This will help ensure a consistent look and feel between forms. It is likely that the style guide will need to be revised periodically to reflect decisions made after reviewing draft test forms.

Item writers will review all specifications, draft items using appropriate sources, following the style guide. A secure item authoring system will be used to ensure consistency of fonts and style, and track items and forms at different stages of development. Item writers will securely share drafts of test forms with technical working group members.

To ensure accommodations and accessibility are addressed upfront and throughout item development, item writers will consider appropriate accommodations and embed appropriate access strategies during the authoring of the items. They will be trained on protocols that will enable them to systematically incorporate a variety of presentation and response strategies that address particular student access needs. These protocols will reflect the accessibility options established in the Accessible Portable Item Protocol (APIP) standard described in the Technology section. In incorporating such training, the ELAAS may be the first national assessment to:

- Purposefully and systematically address the access needs of students at the item level

- Standardize the consideration of access as part of the item writing process
- Develop access options across a variety of students' access needs (e.g., physical, sensory, cognitive)
- Systematically apply a corpus of codes or tags associated with key elements of each test item (e.g., item type, embedded content support/accessibility strategy) that will be used to maximize interoperability, thereby supporting test/test content portability, enabling states to transfer their assessments from one technology platform to another, and supporting a consistent system of assessment delivery across state consortium members

Phase IV: Test Form Review and Acceptance

The ELAAS experts, including teachers, will review the fit between the draft items and the test framework. They will be assisted in this task by a team of ELP assessment experts. Comments will be distributed to the draft item writers for revisions as needed and the process will iterate until a final version is approved for each domain and each grade or grade cluster. That final version will represent the standard against which alternate or parallel forms are developed and evaluated. In this phase, small-scale studies may be commissioned from the Research and Evaluation budget to investigate the appropriateness and/or functioning of any novel item types or formats.

Research and Evaluation

The research and evaluation team is formed by collaboration between GW-CEEE and WestEd in coordination with the CDE and the technical advisory committee. This team will conduct the research and evaluation activities for the ELAAS, providing feedback to the project team and state members that will inform the test development effort by suggesting improvements throughout the project. Summative feedback at the end of the contract will be provided to determine if the program's Theory of Action has been met. The research evaluation team will document and report findings and recommendations to the Steering Committee and technical advisory committee on a regular basis.

Research and evaluation activities will be conducted within a validity framework (Koenig & Bachman, 2004) to ensure the assessment system and its key components are consistent with the Theory

of Action guiding their development and use and to best ensure that assessment outcomes will be meaningful and valid. The *Standards for Educational and Psychological Testing* (AERA, 1999) and the *NCLB peer review guidance* (USDOE, 2009) will be used to inform research and evaluation activities related to examining the rigor and validity of the ELAAS assessments. The *Framework for High Quality English Language Proficiency Standards and Assessments* (AACC, 2009) also will be used to inform research and evaluation activities because characteristics of the LEP population and the ELP construct necessitate particular considerations to ensure validity, meaningfulness, and rigor. Qualitative and quantitative data will be collected and analyzed and these data will be interpreted during formative phases of the system's development to refine and improve the system's elements as well as the system overall. Data also will be collected, analyzed, and interpreted near the completion of the project contract to verify the validity of the ELAAS. The major research activities include:

- Studies of innovative item types
- Comparability studies
- Validation studies
- Alignment of the assessments
- Technical adequacy
- Determination if the program theory of action has been met

Studies of Innovative Item Types

The ELAAS is committed to exploring and using innovation item types. Data to examine the innovative item types will be collected through cognitive interviews with students, teacher focus groups, and pilot tests. For the cognitive interviews, a purposeful sample of students representing each grade cluster, home language, and ELP level (including non-LEP students) will be asked (Almond, Cameto,

Johnstone, Laitusis, Lazarus, Nagle, Parker, Roach, & Sato, 2009; Ericsson & Simon, 1993; Paulsen & Levine, 1999; van Someren, 1994) about their understanding of the questions and what their cognitive processes are for the different item types corresponding to the four domains. The cognitive interviews also will provide the opportunity to assess whether students have sufficient familiarity with the assessed domains (e.g., opportunity to learn) and whether there is any disconnect between the way in which a domain is manifested in the assessment item and the manner in which the domain is presented in instruction. The research and evaluation team will conduct the review of innovative item types in collaboration with the selected test developer for the ELAAS to ensure the information is presented in a timely manner to inform test development.

Comparability Studies

Two comparability studies will be conducted. One comparability study will be conducted to determine the degree to which the two test delivery modes (computer-based and paper-and-pencil) yield similar results and then assess impact of students' computer familiarity (described earlier as "digital divide"). To determine the possible impact of mode and technology experience/access for students (in two languages to be determined) taking the computer-based version, a sample of LEP and non-LEP students across member states will take the assessment in both formats using a counter-balanced design. Additionally, this sample, when taking the computer-based version, will also take a translated or English version of the computer fluidity and computer literacy tests (Higgins, Russell & Hoffmann, 2005). Test facilitators will be surveyed afterwards about their observations. The research and evaluation team will review the outcomes and practical effects of any differences between the modes, and information from these comparability studies will inform the refinement of the assessments and their administration conditions and supports. The research and evaluation team will conduct these studies in a manner to influence the test development process.

Validation Studies

Construct validity will be examined through analyses of student data. More specifically, student field test results will be compared to other ELP measures and other state assessments. Construct validity will also be examined using a series of teacher focus group interviews, representing the consortium membership, to determine the degree to which the assessed domains are consistent with the constructs as currently enacted in classroom instruction. Outcomes of these analyses will inform refinement of the assessments as well as the interpretation of student performance on the assessment items.

The research and evaluation team will evaluate the common English learner definition and common assessment exit criteria to determine the appropriateness of decisions based on those criteria. The research team will conduct studies in member states to obtain another measure of identification, placement, and reclassification. Random samples of approximately 250 students in each grade cluster from member states will be rated by their ESOL/ESL teachers. These teachers will provide their independent ratings of student placement and ELP level of students taking the ELAAS diagnostic assessment field test. These same teachers will rate these students on their summative composite proficiency, reclassification, and proficiency in ELA and mathematics. These results will be correlated with actual student results and placement. Furthermore, correlations between the ELAAS domains, comprehension and composite scores, other ELP measures and other state assessments will be computed for evidence of convergent and discriminate validity, as much as possible during 2014-2015.

To assess predictive validity, a multiple regression approach will be used to investigate the relationships between student demographic factors, such as home language literacy and teacher judgment, and the ELAAS scores. For the summative assessment, the ELAAS summative score, in conjunction with teacher ratings, will be used to predict growth on state content area tests in ELA and mathematics. Predictions will be compared to actual results. (Note that the predictions cannot be developed until at least the second year of the operational program.) Furthermore, external validity will be judged by how well the results can generalize to the population. If the assessments predict better in some states or for some LEP populations, this will indicate that there may be generalizability issues. The research and evaluation

team will examine the results of the generalizability study (Brennan, 1983), identify policy issues, and make recommendations for improvements.

To collect consequential validity evidence, surveys for teachers and administrators will be developed to determine the actual and perceived uses of the tests. All surveys will be created using standard instrumentation development procedures such as those outlined by Fowler, 2009. Documentation from each state will be collected to establish consistency between defined test uses, stated uses within each partner state, and actual test uses based on teacher, administrator, and student experiences.

Alignment of the Assessments

The research and evaluation team will conduct an independent alignment study of the ELAAS assessments. The alignment study will determine if the assessments:

- A. Cover the full range of content specified in the common ELP standards
- B. Reflect the same degree and pattern of emphasis apparent in the common ELP standards
- C. Reflect the full range of language complexity described in the common ELP standards
- D. Yield results that represent all achievement levels specified in the achievement standards

Outcomes of these analyses will enable assessment developers to identify, for example, gaps in skill coverage and under/over-representation of certain skills compared to the breadth and depth of skills reflected in the assessed standards, as well as to make necessary refinements to the assessment items/item bank.

Additionally, PLDs will be reviewed for articulation and consistency across domains and grades. These descriptors will be validated against the relevant policy documentation that informed the definition and articulation of the levels.

Technical Adequacy

The research and evaluation team will review the technical adequacy documentation provided by the test development vendor including the reliability coefficients, standard error of measurement (particularly

around the cut scores and at the tails of the distribution to determine if the scores are precise), inter-rater reliability, item statistics (particularly distractor analysis), and equating models.

The research and evaluation team will examine the relationship of the scoring criteria to the items using the criteria outlined by Bennett, ND, Arter and McTighe, 2001 and Goldberg, 1996. A review of the consistency and reliability of the scorers' ratings will be conducted. Data from scorers' training, validity checks and set checks will be reviewed.

Review of the assessment system's fairness will include determining if the bias and sensitivity item reviews were heeded. Test results and item distractor data across different language groups disaggregated by test performance level will be examined, as will differential item functioning (DIF) results. Because the ELAAS is for the entire range of LEP students, some of whom may require test accommodations, reviewing reports and documentation for evidence of construct-irrelevant variance is necessary.

The research and evaluation team also will review the psychometric methods, dimensionality findings, and equating models and results as reported in technical notes and reports to ensure that the psychometric techniques chosen to scale and equate the diagnostic and summative assessments are suitable for initial identification and student placement and domain and composite score reporting. For the equating method selected, forms and administration guides will be reviewed for their appropriateness. If possible, any linkages to each state's existing scale and trends will be reviewed for accuracy and reasonability. Of particular interest will be the documentation and outcomes on the conjunctive and compensatory models for calculating comprehension and the overall composite scores, as this will have a direct effect on the AMAO 1 and AMAO 2 results.

Determination if the Program Theory of Action Has Been Met

The ELAAS Research and Evaluation team will assess whether the assessment system is feasible and usable, and if systems are in place to ensure it will be implemented as intended. To determine if the system is implemented as intended, administration materials and training will be examined to determine

feasibility of implementation within each state. Feasibility includes computer-based testing in and teacher/administrator's ability to understand documentation such as administration and accommodation guidelines. Focus group interviews and surveys of teachers' and administrators' comfort and experience with the ELAAS implementation will be conducted after professional development training and after test forms have been administered. Furthermore, a sample of teachers will be observed during test administration for additional implementation data. The results will be used to suggest improvements to the ELAAS components and systems. To evaluate whether the theory of action is beginning to be realized, the team will gather evidence about the intended consequences, the validity of the system components, interpretive claims about the assessment results at the student, school and system levels (as appropriate), and potential unintended or negative consequences (Bennett, 2010). The theory and the evidence will be reviewed by the research and evaluation team as well as by the technical advisory committee. Evidence for changes in instructional materials ahead of and planned as the result of the operational assessment will be collected through surveys and interviews with instructional leaders in member states.

To assess whether the system is having the intended effects on individuals (e.g., students and teachers) and on institutions (e.g., schools and English language acquisition programs), evidence will be collected on how students, teachers, programs and schools are affected by changing to the ELAAS framework, professional training on the new system and field test results. A series of focus group interviews with students, teachers and school leaders will define the parameters of a larger survey for each population on consequences of the assessment. Questions will be asked about how the field test results are used and what the results mean for each group or program. Results will be compared to the intended consequences and uses. Discrepancies will be noted and revised materials and guidelines produced as a result. Of particular interest will be how domain weighting will affect state trends and AMAO calculations. In addition, as part of the evaluation of the ELAAS consequences for teachers, the

research and evaluation team will develop guidelines for conducting an evaluation study on the appropriateness of including these data as part of a teacher evaluation system.

Of note, the program evaluation will use research methods to inform decisions and provide evidence that is actionable by stakeholders. As mentioned, the validity framework provides a structure to display the relationships between outcomes and activities that will be monitored.

Professional Capacity and Outreach

This section details the ELAAS consortium of member states' plan for building the capacity of educators in each state to implement the ELAAS and use the assessment results to inform and improve initial identification and reclassification decisions, accountability reporting, and instructional practice. It also details the strategy and plan to provide outreach to involve and inform the public and key stakeholders (including teachers, administrators, families, legislators, and policymakers) in each member state within the consortium about the assessments. Outreach will be conducted in order to build support for the use of the ELAAS to facilitate the learning and achievement of LEP students.

The ELAAS professional capacity and outreach components will utilize the professional knowledge base for effective teaching and professional development (PD) activities that will be ongoing, comprehensive, collaborative, focused on student outcomes, and focused on enhancing the subject matter and assessment literacy of teachers (American Federation of Teachers, 2008; Berman, Desimone, Porter & Garet, 2000; Elmore, 2002). More specifically, the ELAAS professional capacity and outreach components will address the informational, content, technical, and logistical needs of teachers and administrators so they can administer the ELAAS with fidelity. In addition, the ELAAS professional capacity and outreach components will meet the needs of ESOL/ESL and general education teachers and administrators so that they can (1) use the results of the ELAAS to integrate and coordinate language and content instruction for LEP students to improve student achievement, (2) provide a forum for promoting communities of practice related to the ELAAS to support improved teacher and instructional quality, and

(3) provide the general public and key stakeholders (including teachers, administrators, families, legislators, and policymakers) in each member state within the consortium a clear understanding of the purposes and use of the ELAAS and the manner in which it is being implemented.

Plan for Supporting Teachers and Administrators and Developing Professional Capacity

To build the professional capacity of educators implementing the ELAAS, the consortium plans to design, develop, and administer four online training modules which would have both asynchronous and synchronous components. Modules 1 and 2, intended for teachers and administrators, will focus on administration of the ELAAS to assess the four language domains (speaking, listening, reading, and writing) and include information on the determination of appropriate accommodations. Modules 3 and 4 will focus on the use of the ELAAS data to make timely, informed decisions about students' instructional needs related to the four language domains and instructional planning required to facilitate students' progress in ELP. Modules 3 and 4 also will focus on school and district teams to promote and support a culture of discussion *within* each school and foster the creation of a larger, ongoing community of practice. These teams would include stakeholders such as ESOL/ESL and content teachers responsible for teaching core ELA and mathematics content to LEP students, administrators responsible for ESL program implementation, and school and district academic leaders. A general description of each module follows:

Module 1: *What do I need to know about the ELAAS?* This module will include a description of the purposes and intended uses of key components of the ELAAS (screener/diagnostic, summative); the assessed content; student eligibility criteria for the assessment; exclusion criteria; test accessibility features and accommodations; the development process including people/constituencies involved in the development of the ELAAS; administration considerations and the delivery platform; training and scoring considerations; and reporting features. Also, to support the implementation of the ELAAS, a framework to inform states' transitions from their current assessment systems to the ELAAS will be included. This framework will provide states guidance and tools to address the content, technical, and logistical considerations involved in a successful transition.

Module 2: *How is the ELAAS administered?* This module will provide technical and logistical

information about test delivery, administrator training, the administration and scoring of the ELAAS, and specific methods for ensuring that each component of the assessment is administered with fidelity. Module 2 also discusses supports available to teachers and administrators for the administration of the ELAAS (e.g., PowerPoint slides, related informational materials such as pamphlets for teachers and administrators, administration manual, technology support).

Module 3: *How do I interpret the ELAAS results?* This module will guide teachers and administrators through a step-by-step process for understanding the assessment reports and interpreting results. A structured method for reading the reports, determining student and school needs, and developing plans to use the ELAAS results to improve the language and content area instructional practices provided to LEP students will be included.

Module 4: *How do I use the ELAAS results?* This module will provide guidelines and examples of how results can be used to make data-informed decisions about instructional practice and policy (e.g., Means, Padilla, DeArger, & Bakia, 2009). This module will include a structured method for disaggregating data, evaluating student and school needs using the ELAAS results in conjunction with outcomes of other measures as appropriate, using the data to inform teaching that supports the achievement of LEP students in both language and academic content proficiency, and evaluating the use of results as well as to identify administrator (e.g., principals, assistant principals) and teacher professional development needs and to inform principal and teacher effectiveness.

The ELAAS online training project team (WestEd in collaboration with GW-CEEE and TESOL) will offer at least two webinars each year related to each of the four modules, for a total of at least eight training webinars across the four years of this project—these will comprise the “initial” training modules. The initial training modules provide a method for establishing the ELAAS communities of practice. To further support the fostering of these communities of practice, and as a forum for sharing information, experiences, and supporting resources, each initial training module also will have 2-3 optional follow up trainings that focus on key aspects of each module in greater depth. For example, initial training for Module 1 will present the full scope of information described above. The optional follow up trainings for

Module 1 would provide more detail and discussion of (1) the assessment's content, design, and development process, (2) characteristics of the students who will be assessed, eligibility criteria, and accessibility and accommodations considerations, and (3) administration, scoring, and reporting considerations. These optional follow up trainings will be offered once each year. Additional community of practice resources and discussions will be offered in a password-protected area. Materials for the four online training modules will include readings, case study examples, PowerPoint presentations, authentic student audio and video samples, sample guidelines and templates, and video screen casting demonstrations of tool, guideline, and template use. Successful module completion will be determined using a brief online, self-scoring assessment. An option for SEAs to provide state certificates of completion and if applicable, relicensure points to participants who complete each module (i.e., each of the four initial training modules) would contribute to enhanced motivation among module participants.

The ELAAS project online training development team (GW-CEEE in collaboration with WestEd and TESOL) will manage the design, development, and administration of the four training modules (initial and follow up training modules). The ELAAS project development team will work with instructional designers to create suitable module content for a web-based interface. To ensure integration of the ELAAS into state systems of support, the four ELAAS training modules will be designed following Sharable Content Object Reference Model (SCORM) standards and specifications for web-based e-learning requirements for compatibility in multiple learning system platforms, thereby allowing flexible distribution across online learning platform systems. These designs will build on existing online training designs for test administration which are used successfully by the ELAAS consortium member states, including the CELDT Scoring Training of Trainers (CDE, 2011) and the Minnesota Test of Academic Skills Training for Test Administrators (MDE, 2011).

All four modules will be designed in coordination with other elements of the ELAAS project. Modules 1 and 2 will be piloted during the initial development phase (2012–13), field tested (2013–14) during the full field test administration, and ready for full implementation with the launch of the operational ELAAS assessments (2014–15) to be used each subsequent year prior to administration of the

ELAAS. Module 3 will be piloted and ready for use the second half of year 2 (2012–13) of the project. Module 4 will be piloted and ready for use by the beginning of year 3 (2013–14) of the project. A group comprised of educators from consortium member states, including ESOL/ESL teachers, content teachers, and administrators will be convened each year of the grant to review module content and results and recommend improvements to the content, design, and delivery of the training modules. This group will help ensure that the content of the modules are meaningful and useful to educators and developed with broad stakeholder involvement.

Outreach Plan and Strategy

The goals of the outreach plan and strategy are to inform the public and build support among key stakeholders (including teachers, administrators, families, legislators, and policymakers). Beginning in the first year of the grant, the ELAAS consortium will use a variety of media and strategies to distribute information and build understanding about the ELAAS.

At the national level, the ELAAS project outreach team (TESOL in collaboration with GW-CEEE and WestEd) will coordinate with major organizations such as NCEO, and their national and state networks and affiliates to design a clear, consistent message about the purpose and use of the ELAAS. Outreach efforts will connect with existing within-state networks, especially state assessment and evaluation personnel, to distribute information on the assessment, using conference presentations, face-to-face meetings, websites, listservs, online community spaces, and online newsletters. The ELAAS project staff (TESOL, GW-CEEE, WestEd) will develop sample training documents which can be customized by the states for outreach purposes to state policy makers and district and school administrators. These outreach activities will focus on supporting public and stakeholder understanding, support, and implementation of the ELAAS.

The ELAAS project outreach team also will use online events to support understanding and build the capacity of the public and stakeholders on research and practice relevant to current issues related to ELP assessment. More specifically, TESOL and WestEd will use their online platforms (SchoolsMovingUp (SMU), a WestEd web-based forum, and the online TESOL Community) to offer a

series of presentations and discussions to support states' capacity to implement the new generation of ELP assessments such as the ELAAS. The TESOL Community and the SMU platform also will be used to help the ELAAS project staff conduct an annual survey of state, districts, and school educators to gather common questions and informational needs about ELP assessment more generally and the ELAAS and implementation processes more specifically to inform the content of these four webinar modules, and to best ensure that the content of the webinars are relevant and useful to the public and consortium stakeholder groups. TESOL will also sponsor webinars that focus specifically on the teacher quality aspect of the implementation of the ELAAS and the implications on instruction, built around the TESOL professional standards framework (TESOL, 2009). In addition, the TESOL Community platform will be used to support and facilitate a community of practice of those involved with ELP assessment.

To inform parents, the ELAAS consortium will provide translations of relevant materials into the top five LEP student home languages among consortium member states. However, due to the number and range of LEP student native languages, as well as various state laws and regulations governing the translation of documents, states and districts will translate relevant parent materials into additional languages as required. States and districts will be encouraged to translate relevant parent materials into additional languages specific to their respective jurisdictions. Translation of materials into less commonly-spoken student native languages can be shared among member states.

The ELAAS professional capacity plan is cost-effective, feasible, and supports the proposed theory of action. The plan calls for a training system which combines online training modules with webinar events to ensure cost-effective delivery of test administrator training. The ELAAS plan provides guidance and opportunities for deeper analysis and application of the ELAAS results in schools and districts during the school year. The scheduling and design of trainings and coordinated strategies for promoting communities of practice allows for the development of an ELAAS community of practice that is feasible for educator participation and practical application to attend throughout the school year and yield practical results. The outreach strategy and plan supports large-scale and targeted distribution of information and discussions among existing state and national networks to ensure a clear, consistent

message, in compliance with federal regulations, surrounding the use of the ELAAS.

Technology Approach

Leveraging technology to measure efficiently and effectively student ELP will be key in this assessment system. The ELAAS will use a comprehensive technology approach that will allow for seamless transition of item development, administration, scoring, and reporting, as well as provide a ready-made solution for portability of the ELAAS products to other states or organizations.

The ELAAS is proposed to be developed primarily for computer administration of its assessments. While paper or teacher administration is expected to be necessary in the near term because of some districts' infrastructure capacity or students' lack of exposure to computer based testing, The ELAAS will be preparing for its first administration four years from the grant award, which likely will mean only greater saturation of technology in the schools and students ability to interact with computer-based administration. Further, the lead time will allow states and districts an adequate period to prepare for computer-based administration.

The ELAAS will address the critical need for efficiency in test design, item development, and test administration by adhering to an open-source standard developed specifically with this goal in mind. The Accessible Portable Item Protocol (APIP) Standard provides the technical specifications required to modify or develop item authoring and test delivery systems capable of delivering items that are wholly interoperable. This means that the ELAAS will require items developed for its assessments to be coded according to the APIP standard from conception, and these items will be available for use in any APIP-compliant test engine without re-engineering or revision. The ELAAS will thus be developed without the need to rely on a specific commercially developed software or proprietary architecture. Rather, the ELAAS will develop APIP-compliant items that can be used in any APIP-readable test engine: this will provide the ELAAS the ability to select the best test engine solution that suits its needs. Moreover, if member states are administering their content assessments through an APIP-compliant test engine, it may be possible to administer the ELAAS using the same test engine, thereby reducing the administrative

burden of school staff and decrease construct irrelevance for English learners if they have additional experience with a familiar test engine.

The ELAAS will address the critical need for effectiveness in test design, item development, and test administration using APIP as well. In addition to portability of item content, APIP has standardized the manner in which items can be made more accessible for a variety of learners with a variety of needs. By applying principles of Accessible Test Design, the ELAAS will provide more valid inferences about student language proficiency based on test scores for all students. Rather than developing assessment content for the general population of LEP students and then making post hoc changes to that content in order to accommodate the needs of sub-groups of these students, Accessible Test Design provides a framework for making careful decisions about the methods used to tailor test administration to maximize the measurement of targeted constructs for each student. In turn, the Accessible Portable Item Profile (APIP) Standards provide a tool for implementing Accessible Test Design during item development.

The APIP Standard empowers the ELAAS to solve three challenges. First, APIP provides a structure for specifying and storing the access needs of each student. Second, APIP provides a structure for augmenting item content with a variety of supplemental and alternate accessibility information designed to ensure that a test item functions properly for students with a variety of access needs, and these needs are obviously varied across the spectrum of LEP students. Third, APIP provides specifications for developing test delivery systems that can use a student access profile to tailor the provision of access tools and accommodations (e.g., magnification, color contrast, masking) and the presentation of supplemental accessibility information (e.g., audio, Braille, tactile, or signed versions of item content). Collectively, the tools provided by APIP enable next generation assessments such as the ELAAS to capitalize on the flexibility of digital technologies to maximize test validity for all students.

Rather than develop a new item content development system that is APIP compliant—a complex and expensive task—the ELAAS consortium will sub-contract with an organization that has already developed or modified an item development system that supports the creation of APIP compliant test items. Similarly, rather than developing a new test delivery system that is APIP compliant, the ELAAS

consortium will sub-contract with an organization that has developed or modified an APIP compliant test delivery system. This system will be used to pilot test items and tasks during the test construction process. It is expected that the APIP compliant test delivery system will be accompanied by a Student Access Needs Profile system that supports the creation of APIP compliant student personal needs profiles.

To keep fully informed of APIP and to influence APIP's evolution, the member states of this proposal have already forged a partnership with IMS Global Learning Consortium to serve as Contributing Members, participating in the APIP workgroup and QTI2/APIP Alliance Partnering as Contributing Members with IMS means that our consortium will be represented at the table in the continued development of the APIP standard so that the needs of LEP students are well represented. With equal partnership with test publishing companies such as ETS, Measured Progress, and Pearson, Questar our ELAAS development is poised to be the first consortium to this table to ensure the LEP student is considered in developing access priorities. If necessary, modification to this approach will be done after field testing of the assessments.

The ELAAS will leverage Question Test Interoperability to serve as its scoring and reporting functionality. QTI already has such features well developed and this standard is already frequently used in large-scale assessment. Moreover, as a Contributing Member in the APIP workgroup, our consortium will be able to work with the other organizations supporting APIP to ensure that QTI scoring and reporting are added to the APIP standard to support the needs of LEP students.

The implementation of the ELAAS as a computer-based system will create its own set of challenges, as such a project does any time that a new assessment begins or transitions from one mode to another. As the ELAAS plans to establish computer delivery as its main mode of administration, our consortium recognizes that there are infrastructure challenges ahead. A sufficient testing window that will allow all students sufficient time to move through a limited number of computer terminals remains a perennial challenge in computer-based testing. However, the unique nature of the ELAAS and the population it will serve – along with the anticipated adoption of computer-based testing in most states

under the Race to the Top Assessment consortia – will allow greater flexibility in a successful deployment.

In most districts, the number of students participating in the ELAAS will be considerably fewer than the general population as a whole. This means that capacity issues of both hardware and bandwidth will be lessened, computer-based delivery of the ELAAS could actually improve the efficiency and validity of the system.

Project Management

As of the submission of this grant proposal, 18 states will participate in the consortium led by the CDE. The CDE shall contract with the CCSSO in Washington, D.C. to provide all services designated as the responsibility of the management partner as required by the Enhanced Assessment Instruments Grant for the development of ELP assessments. CCSSO will work exclusively with the ELAAS consortium for the 2011 EAG grant competition. The CDE and CCSSO are referred to herein as the ELAAS project team and are responsible for overall management of the ELAAS project.

Consortia Structure

All 18 states have signed the Memorandum of Understanding (MOU) and have agreed to participate in the development and implementation of the ELAAS in accordance with the requirements of the U.S. Department of Education (USDOE). Each member state will nominate a representative to serve on a Steering Committee, chaired by Deborah Sigman, Deputy Superintendent, Curriculum, Learning, and Accountability Branch at the CDE or someone she designates to act in this role. California is referred to herein as the Lead State. A technical advisory committee will also be formed to advise the ELAAS project team and report to the Steering Committee upon request regarding assessment design and development, and technology. The Lead State will work directly with CCSSO, the Project Management Partner, and the USDOE to make sure that the deliverables are all produced in a timely manner and are aligned with what is proposed in the project narrative.

Each consortium member will play a key role on policy, operational and logistical decisions. Decisions will normally be made on a consensus basis; if consensus cannot be reached the decision will be determined by a majority vote. When a decision is needed, the Lead State Chairperson will convene the Steering Committee for a vote. Each member of the Steering Committee is entitled to one vote. Any decision voted on with less than two vote margin will be reopened for debate.

Protocols by which the consortium will operate, including protocols for member states to change roles in the consortium, for member states to leave the consortium, and for new member states to join the ELAAS Consortium are described in the MOU (Attached) and were agreed upon by each member state prior to signing on the project.

The ELAAS intends to use a common set of standards upon which to base the assessments, a common set of performance-level descriptors, a common set of English language achievement standards, common assessment administration procedures through a rigorous process of research and review. Each state in the ELAAS Consortium will have a say through their vote in the Steering Committee on how these will look. A common set of policies and procedures for accommodations and student participation will be developed as the operational assessment begins to take shape.

California's plan for funding the ELAAS assessments after this project concludes is based on current leveraging of legislative support for changing the ELP assessment system. During the development of the ELAAS, states will continue to administer their existing ELP assessments. For example, California the ELAAS lead state will continue to deliver its current ELP assessment, the California English Language Development Test (CELDT) in accordance with state law (Education Code sections 313 and 60810). In 2013, California will prepare a request for proposal (RFP) for the next test contract to incorporate the ELAAS assessments and discontinue the use of the CELDT as its statewide assessment for initial identification and annual assessment of LEP students as of the 2014-15 school year.

Member states will follow their own procurement rules to adopt the ELAAS once it is fully operational.

Details are further broken out in the timeline below.

Exhibit 1. Project Management Plan Timeline

Milestone	Date Milestone Completed
All member states approve MOU binding them to project	6/3/2011
CDE contracts CCSSO as Management Partner	10/3/2011
CCSSO contracts subcontractors	10/10/2011
CDE/ELAAS Work Group selects vendor(s) to develop test specifications	11/1/2011
Vendor(s) deliver ELP core items for field-testing	Spring 2012
Vendor(s) prepare a plan for development of test specifications aligned to English proficiency content standards.	Spring 2012
Vendor(s) prepare a plan for development of items aligned to English proficiency content standards.	Fall 2012
Vendor(s) prepare a plan for detecting, examining, and addressing any undue bias found in the tests prior to the administration of the test	Fall 2012
Vendor(s) prepare a plan for examining reliability of the tests using classical approach such as test-retest reliability and/or generalizability approach (to partition error variance into components by different subgroups).	Fall 2012
Vendor(s) prepare studies for Item-Level Analyses/Statistics: a) Item difficulty and discrimination, Mean, SID, item-total correlation, impact of individual items to the overall internal consistency index. b)Using spring 2012 field/census data, DIF analyses by gender, racial/ethnic, LEP status, non LEP status, language proficiency rating subgroup, and type of language/culture dependent on an adequate sample size c) IRT scaling and analyses using 2013 field data, equating and vertical linking	Fall 2012
Vendor(s) prepare a plan for IRT scaling and equating.	Fall 2012
Vendor(s) prepare a plan for paper pencil transition	Fall 2012
Vendor(s) prepare a plan for computer based test (CBT)	Fall 2012

Vendor(s) prepare a plan for standard setting.	Fall 2012
Vendor(s) prepare a plan for development and validation of scoring rubrics and inter-rater reliability information.	Fall 2012
Vendor(s) prepare a detailed plan for all field-testing, addressing sampling design, test format (common core vs. parallel forms), methodology/strategy for modifying problematic items, scoring, computer and paper base and analyses of data.	Fall 2012
Vendor(s) prepare a teacher-rated speaking scale aligned to the speaking PLDs, for those states planning to conduct census testing for the spring administration	Completed before the 2013 Field Trial
Vendor(s) prepare a summary of the main points and discussions of any meeting or conference call and provide them to CDE/CCSSO on behalf of the ELAAS consortium within 10 days after the event. This will be reviewed and prepared for the ELAAS consortium meeting/conference call summaries and any other notes or memoranda created by CDE/CCSSO and provide them to the ELAAS participants. Approval of these meeting notes will serve as the acceptance criteria for the project.	January 2013
<u>Milestone 1: Key Features</u> – ELAAS – Field Test Preparation and Implementation	
Vendor(s) prepare all materials specified below as necessary for a 2012-2013 field test of the ELAAS in hard copy and a computer format usable for the operations contractor including: All processes and materials listed in Milestone 2 through Milestone 4 necessary for 2012 and 2013 field tests.	2013
States conduct pilot tests and research studies	2013
<u>Milestone 2: Key Features</u> – ELAAS – Final Forms	
The vendor(s) ensure the ELAAS Final Forms contains:	Spring 2013
(A) Four assessment components, two focusing on written discourse, two focusing on oral discourse. The four assessment components will yield scores in 1) listening, 2) speaking, 3) reading, 4) writing 5) a comprehension score based on listening and reading, and 6) a composite score.	Spring 2013

(B) Three levels of assessment within each component to reflect language development across grades K-12. These levels will be K-2, 3-5, 6-8, and 9-12.	Spring 2013
(C) A paper-pencil assessment format except for speaking and listening will be developed for states while they transition to CBT. The speaking assessment will be provided on audio tape, or CD-Rom and a written script and the students' responses will be recorded on audio tape. The listening assessment will be provided on audio tape or CD-Rom and the students' responses will be written. However we will use any public domain technology for speaking and listening if financially viable.	Spring 2013
(D)The development of core test items measuring what are uniformly regarded as required standards that APIP compliance and can be utilized on both CBT and paper pencil..	Spring 2013
(E) IRT-derived item calibrations.	Spring 2013
(F)Assessment content aligned with to consortium and or college- and career- ready standards and benchmarks for the form and recommended bookmaps referred in 1.2 for each assessment component (e.g., writing) in each grade level span (e.g., grades 3 - 5) as well as item bank approaches to meet the differential needs of states adopting the assessment.	Spring 2013
CDE/CCSSO and vendors will have meeting to ensure that all the key features have been addressed to the satisfaction of the Steering Committee.	Spring 2013
<i><u>Milestone 3: Materials Development – ELAAS Final Forms</u></i>	
The ELAAS consortium determines requirements for 2012 such as the number of two field test forms producing the appropriate number of tests shall be developed for the all grade clusters. Item selection, form construction, scaling and equating design and procedures, and scoring process will be the responsibility of vendor(s). Responsibilities will also include the necessary reviews for any bias in the tests and the development and validation of a scoring rubric for open-ended and extended-open ended items.	Summer 2013

Vendor(s) will produce large-print, and loose-leaf test booklets and answer documents will be offered to accommodate students with disabilities participating in the ELAAS at all grade cluster levels.	Summer 2013
Vendor(s) develop test administration manuals providing the contents in an electronic format usable by the consortium. The manuals will contain general information about how to conduct the assessment and specific test instructions.	Summer 2013
Vendor(s) produce test Administration Manual – Scripted and computer directions for test administrators and procedures for test administration including security.	Summer 2013
Vendor(s) produce Pretest Workshop Training Materials. Training materials will include transparencies, handouts, a script, and PowerPoint computer based presentations that coordinate the use of the training materials. The Consortium will require copies of all manuals for all staff use.	Summer 2013
CDE/CCSSO and the vendor(s) conduct a minimum of two (3) regional training workshops that may employ the use of WebEx and videoconferencing.	Summer 2013
Vendor(s) develop and produce an Interpretive Guide in hard copy and electronic format to assist district and school personnel in reading and using the ELAAS reports.	Summer 2013
<u>Milestone 4: Production of Materials – ELAAS Final Forms</u>	
CCSSO/CDE approve the form of bookmaps and coded items (master copies)	Fall 2013
<u>Milestone 5: Distribution Procedures – ELAAS Final Forms</u>	
Vendor(s) and CDE/ CCSSO jointly approve final distribution procedures for the completed the ELAAS.	Fall 2013
Member states in the ELAAS consortium establish state and local procedures for distribution of the ELAAS materials once final publication is achieved.	Fall 2013
<u>Milestone 6: Procedures for Data Identification, Receipt Control, Scoring, and Editing that states will use once ELAAS is available.</u>	

Vendor provides up to 10 days of senior level consultation to support the process of scoring and reporting.	Fall 2013
<u>Milestone 7: Technical Reports</u>	
The vendor(s) produce a technical report in print and appropriate electronic formats on results of the spring 2013 field test of the ELAAS using data provided by the ELAAS consortium	Fall 2013

Approach to Identifying, Managing, and Mitigating Risks

As the lead state, California will be responsible for monitoring the contract. As part of the monitoring process, the ELAAS project team at the CDE will institute a review process that includes a systematic approach to review and approve all project deliverables at a fine-grain level. This approach ensures that the project deliverables are reviewed from start to finish and that formative feedback, including input from the ELAAS TAC, Steering committee and Research and Evaluation Team, is provided to the contractors responsible for each major task, so that the final products meet with the highest professional standards for test development, professional development, and technology.

Another inherent risk in project such as this is to test security. We plan to mitigate this risk by arranging a secure and reliable repository for the data. All data will be stored in a secure database at the GW-CEEE. Access to the database will be protected with secure passwords and encryption, and is accessible only on a need to know basis by the ELAAS Project Team staff assigned to test security. Any data for analyses performed by project consultants will have had all student identifying information removed. The newly proposed amendment to the Family Educational Rights and Privacy Act (FERPA), if approved, would give states flexibility to share student while safeguarding students’ personal information. Some student identities will be known to the project staff who interview them and are working with their teachers. Records that include subjects’ names or other identifying information, such as video, will be kept in locked file cabinets. All subjects will be assigned numbers, and these code numbers will be used on all data records instead of names. Project staff will respect teacher and student confidentiality in terms

of participation and comments related to performance. Project staff will not provide student scores or discuss test performance with any student or individual outside of the project. Likewise, they will not discuss any teacher's performance with anyone outside of school staff at any time. Any notes, forms, video, or protocols will be immediately stored to prevent accidental breaches in confidentiality. All staff members will abide by a code of confidentiality. No names will be used in conjunction with any reports of the results of this project. The CCSSO will assist the California ELAAS project team in review and oversight of project deliverables before distribution to or use by member states.

Adequacy of the Budget

The requested funds are necessary to support a coast-to-coast consortium that serves over half the nation's LEP students. California alone educates about one third of the nation's LEP students. The number of students to be served by the ELAAS consortium is estimated at 52.3 percent (2,376,524) of the 4,539,740 limited English proficient students in the nation.¹ We intend to deliver a top quality assessment while keeping within the parameters of the proposed budget.

The fiscal status of the project will be thoroughly monitored over the grant period to ensure that it is kept within the proposed budget. USDOE will also be provided with quarterly reports outlining the progress of the deliverables including costs to date for each deliverable. Bi-monthly reports to ensure that all deliverables are on course for successful completion will be produced and delivered to CDE to internally monitor this process as well. In addition, all contracts written for the project will include a cap that will not be adjusted and will fall within the proposed budget. By meticulous project management, we intend to have a fully operational assessment by 2014 created within the proposed budget.

Estimated Costs Associated with the ELAAS for States

¹ Source: *Summer 2010 EdFacts State Profiles* available online at <http://www2.ed.gov/about/inits/ed/edfacts/state-profiles/index.html>

Since states currently provide screening and summative assessments, a majority, if not all, of the operational costs for the ELAAS will not be new, but replacement costs. The implementation of a common set of assessments will provide some savings for each participating state, since revisions to the assessments, such as creation of new items, will be done for all member states at once, rather than individually by each state. CCSSO contacted states to provide the estimated costs for the ongoing administration, maintenance, and enhancement of the operational assessments. Of the states that responded to our request, the estimated average cost per student was \$14.25. The estimated average cost per teacher and administrator for training was \$26.25. States were also asked how they plan to fund the ELAAS after the project period is over. The overwhelming response from the states was that they will continue to pay for the amount of the test administration out of state assessment dollars. These figures show that the ELAAS will be affordable and self-sustainable for all the states once the project ends.

Project Personnel

As lead state, California will devote personnel to fulfill its leadership role in the ELAAS. The success of this program is critically important to the mission of the CDE and to the state legislature as California strives to serve its 1.5 million LEP students. In addition to the positions funded by the grant, there will be significant, in-kind support for the ELAAS as follows. Deborah V.H. Sigman², Deputy Superintendent, Curriculum, Learning, and Accountability Branch will provide leadership as she works with the legislature, finance, and advocacy groups to ensure that the pathway is clear for implementing the ELAAS operationally by 2014–15. Deputy Superintendent Sigman will serve as a liaison for the State Superintendent to other states and institutions in support of the ELAAS. Rachel Perry, Director, Assessment and Accountability Division, will provide leadership and support for the work of the ELAAS project team, as well as coordinate with the Directors of the Special Education Division and English

² CVs for all personnel listed are included as a separate attachment and further outline their qualifications

Learner and Curriculum Support Division to ensure that appropriate consultant support is available to assist with the implementation of tasks. Sharon Taylor, Director, Fiscal and Administrative Services Division, oversees the accounting, budgeting, and contracting functions at the CDE. She will ensure that California, as the fiscal agent handles, the grant funds in accordance with specified laws and regulations.

Project Director (60%): Lily Roberts, Ph.D. will serve as the Project Director, of the ELAAS project team at the CDE. Dr. Roberts is highly qualified for this role given her education and professional experience in large-scale assessment, test development, standard setting, evaluation research, and program management. Dr. Roberts will be responsible for: overseeing the California ELAAS staff; coordinating with CCSSO to ensure that the ELAAS-related meetings are held, project deliverables are received, and timelines are met; overseeing the tracking and monitoring of the budget and contracts for all aspects of the ELAAS; reporting to the Steering Committee; providing the Steering Committee Chairperson with regular updates, not less than quarterly, on project progress; serve as primary liaison to the technical advisory committee, the Research and Evaluation team, and other contractors and collaborators.

Michelle Center, Educational Research and Evaluation Administrator (20%) will act as the liaison between the Project Director and the CDE's executive office to ensure the project is provided the necessary resources. She will keep the executives apprised of current progress on the project. She will provide support to the project on technical matters related to test development and implementation.

Eric Zilbert, Ph.D., Educational Research and Evaluation Administrator (10%), will assist with the periodic review of technical studies. One of his key roles will be to apprise the CELDT TAG of the development of the ELAAS and solicit their expertise in language acquisition and ELP assessment in the transition from CELDT to the ELAAS by 2014–15. He will supervise replication studies, review of technical reports from the research and evaluation team, and assist Dr. Roberts and her team as needed on other technical aspects of the ELAAS, such as CBT and standard setting processes.

Tavi G. Popp, Education Research and Evaluation Consultant (40%) in the ELAAS, CELDT and Title III Accountability Office, will assist the Project Director in monitoring the development of the ELAAS assessments and the research and evaluation studies. She will monitor all item development phases of the ELAAS, including development of the test blueprint and item specifications of the diagnostic and summative assessments, development of new item types, item writer training, item development review and approval, content and bias and sensitivity reviews, pilot testing that includes think aloud and cognitive labs, field testing planning and design, field test construction, development of accommodated testing protocols, scoring and range-finding of constructed-response items,. She will continue to work on the CELDT and serve as the liaison to the CELDT Technical Advisory Group to ensure technical support for the transition from the CELDT to the ELAAS by the end of the project period.

Jessica Barr, Education Programs Consultant, (50%) in the ELAAS, CELDT and Title III Accountability Office will assist the Project Director in monitoring the development of the technology aspects of the ELAAS, including the CBT, scoring and reporting, and development of ancillary materials such as scoring guides and directions for administration as well as the professional outreach to ensure that the member states and other stakeholders receive timely information on the ELAAS to ensure that the assessments can be readily moved into operational status in the states.

Dianna Gutierrez, Education Programs Consultant, (10%) in the Language Policy and Leadership Office, will coordinate with the ELAAS, CELDT and Title III Accountability Office in the development and dissemination of the ELAAS. She is bilingual and she works with all the Title III schools in the state.

Other Collaborators

Sharon Sáez is ELL and Inclusion Program Director with the Standards, Assessment, and Accountability Strategic Initiative at the Council of Chief State School Officers (CCSSO). She is also the collaborative advisor for both the English Language Development Assessment and English Language

Learners collaborative. Ms. Saez will coordinate the work of the consortium of state members, development partners, the principal investigators, project analyst, and the advisors, and oversee all work of the subcontractors and their activities across the member states.

Alison Bailey, Ph.D., is Professor and a former Division Head of the Psychological Studies in Education program of the Department of Education, University of California, Los Angeles, in addition to being a faculty associate researcher at the National Center for Research on Evaluation, Standards and Student Testing (CRESST).

Laurene Christensen, Ph.D., is a Research Associate at the National Center on Educational Outcomes. Dr. Christensen will coordinate project activities, overseeing the accessible test design and universal design contributions, as well as the review of test items. Dr. Christensen will provide liaison between NCEO and project partners.

Mark Wilson, Ph.D., is a Professor of Policy, Organization, Measurement, and Evaluation at the University of California Berkeley Graduate School of Education. He will advise the ELAAS project on issues of policy, organization, measurement, evaluation, cognition, and development.

Jamal Abedi , Ph.D., is a Professor in the School of Education at the University of California, Davis and a research partner at the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Dr. Abedi will provide his expertise by serving on the Technical Advisory Committee.

Kenji Hakuta, Ph.D., is a Lee L. Jacks Professor of Education at Stanford University. He is currently leading a national effort in the development of ELP standards that build on the linguistic supports necessary to attain the Common Core State Standards. Dr. Hakuta will be a member of the ELAAS Technical Advisory Committee and will ensure that the ELAAS is fully aligned to the ELP standards work.

Edynn Sato, Ph.D., is the Director of Research and English Learner Assessment in the Assessment and Standards Development Services (ASDS) program at WestEd and the Director of Special Populations for the Assessment and Accountability Comprehensive Center, a federally designated national center that provides technical assistance in the areas of assessment and accountability to regional centers and state departments of education.

Charlene Rivera, Ph.D., is the Director of George Washington University Center for Equity and Excellence in Education. Dr. Rivera will oversee all GWU aspects of the project, contribute to and review project elements, and provide executive guidance on research activities carried out by GW-CEEE staff. She will provide expertise with regard to the professional development, outreach, research and evaluation activities.

Rosa Aronson, Ph.D., is the Executive Director of the TESOL International Association (TESOL). Dr. Aronson will coordinate professional capacity and outreach activities on behalf of TESOL, lend her expertise in these areas, and oversee these activities that are carried out by TESOL staff and consultants.

William Lorie, Ph.D., is the Managing Director of Metrica Research Associates and President of EduMetrica, Inc. President EduMetrica. Dr. Lorie will advise the ELAAS consortium on issues of psychometrics.

Project Management Partner

CDE has identified the Council of Chief State School Officers (CCSSO) as their intended General Project management partner. Founded in 1927, CCSSO is a nonpartisan, nationwide, nonprofit organization of the public officials who head the departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. CCSSO seeks its members' consensus on major educational issues and expresses their views to

civic and professional organizations, federal agencies, Congress, and the public. Through its structure of standing and special committees, CCSSO responds to a broad range of concerns about education and provides leadership and technical assistance on major educational issues.

CCSSO is extremely qualified to assume the role of Project Management Partner for the ELAAS. CCSSO has a proven track record of managing Enhanced Assessment Grant (EAG) projects, serving as the primary management partner for six EAGs since the inception of the program. CCSSO, along with member states, has also successfully developed and managed an operational English Language Acquisition assessment through its English Language Development Assessment (ELDA) EAG. ELDA is a battery of tests designed to allow schools to measure annual progress in the acquisition of ELP skills among non-native English speaking students in grades 3-12.

CCSSO has a long standing history of supporting and facilitating collaborative work that empowers state education leaders to come together to address common challenges, share resources, and produce work that ultimately supports all states as they strive to improve the educational outcomes of all children. A majority of CCSSO consortia/collaborative bring states, thought leaders, and funders together in long-term multi-year work. These state consortia encompass content-specific, subgroup, and technically relevant interests in areas such as, accountability systems, assessment for LEP students, teacher and leadership standards development, comprehensive assessment system for ESEA Title I, and extended learning opportunities.

The ELAAS scope of work will connect with the Council's established on-going state collaborative efforts to include the College- and Career- Ready Standards, Implementing Common Core Standards (ICCS) representing thirty six state teams, the State Collaborative on Assessment and Student Standards (SCASS) system, and access to all fifty member Chiefs. These established connections create a ready mechanism for conducting and disseminating the work, provide administrative, meeting, and logistical support, and ultimately help to inform a larger transformational agenda in education policy and practice.

Summary

The ELAAS has the potential to improve how over half the U.S. LEP students' ELP level is assessed, improving their educational outcomes. This project is built around the solid foundation of collaboration between 18 states, leading organizations in the education of LEP students, assessment, and experts in the field to positively impact LEP student learning in ELP and in the content areas. The ELAAS will be aligned to ELP standards, correspond to the college- and career-ready standards, and complement the PARCC and SBAC. The ELAAS will be based on current research on the development, validation, and implementation of ELP assessments in the nation and around the world.

The system will generate a wealth of data that can be shared by the participating states for assessment, accountability, and research, and will provide important data for informing curriculum planning and instruction in order to meet LEP students' academic needs. The ELAAS will be sensitive to the heterogeneous nature of the LEP student population, adopting principles of Accessible Test Design. The ELAAS system will include accommodations for LEP students with disabilities and will carefully select accommodations for these students. The ELAAS system will include a comprehensive reporting system which will contain information from the diagnostic and the summative assessments. The research and evaluation team will conduct a review of the technical adequacy of the assessment and the assessment system as well as an evaluation of the development process and the theory of action. Using quantitative and qualitative methods, the research and evaluation effort will provide formative information to improve the assessments and processes as well as summative information on how well the assessments provide valid and reliable results and how well the system works.

The ELAAS professional capacity and outreach components will facilitate LEP student achievement by situating the assessment system within a targeted PD framework. These components will allow all teachers and administrators to use the results of the ELAAS to integrate language and content instruction for LEP students to improve student achievement, promote communities of practice, and provide the general public and key stakeholders in the consortium a clear understanding of the purposes

and use of the ELAAS and the manner in which it is implemented.

The ELAAS will be developed primarily for computer-based administration of its assessments, addressing the critical need for efficiency in test design, item development, and test administration by adhering to an open-source standard developed specifically with this goal in mind. The APIP Standard provides the technical specifications required to modify or develop item authoring and test delivery systems capable of delivering items that are wholly interoperable, thus developed without the need to rely on a specific commercially developed software or proprietary architecture. The ELAAS is poised to be the first consortium to ensure that LEP students are considered in developing access priorities.

The CDE will contract with CCSSO to provide all services designated as the responsibility of the management partner. The project's clear management plan and timeline will lead to the success of the assessment; the project personnel and collaborators' nationally recognized expertise only enhances this solid proposal. Thus, we are pleased to submit this proposal to ensure that assessment of LEP students will inform improvements in teaching and learning, and ultimately foster success in academic content classes.