Assessment and Accountability
Comprehensive Center

AACC ● A WestEd and CRESST partnership

Framework for High-Quality English Language Proficiency Standards and Assessments

NoChild
LeftBehind

Framework for High-Quality English Language Proficiency Standards and Assessments

Suggested citation:

> Assessment and Accountability Comprehensive Center. (2009). *Framework for High-Quality English Language Proficiency Standards and Assessments.* San Francisco, CA: WestEd.

Questions and information about the Framework

Questions and requests for additional information about this Framework should be directed to:

> Dr. Edynn Sato
> WestEd, Assessment and Accountability Comprehensive Center
> esato@wested.org

A related brief and its references can be obtained at

> www.aacompcenter.org.

Acknowledgments

Contents

Purpose of the Framework

The *Framework for High-Quality English Language Proficiency Standards and Assessments* (Framework) was conceived as a critical tool in states' efforts to ensure that their English learner (EL) students achieve English language proficiency (ELP) and, also, achieve at high levels academically.[1] Building on the best knowledge from relevant research and practice, the Framework provides criteria for high-quality ELP standards and aligned assessments. It is essential that a state's system of standards and assessments accurately reflect the state's values, priorities, and needs — all of which may change and evolve over time.[2] Thus, it is important that the system be reviewed regularly and, if needed, refined. Intended primarily for state departments of education, the Framework suggests a cross-disciplinary process for using the criteria to either 1) engage in an evaluation of your state's existing ELP standards and assessments, and their implementation, or 2) oversee the development and implementation of new ELP standards and assessments.

The Framework includes a variety of worksheets for inventorying your evidence of quality and decisions and actions associated with judgments of quality as you evaluate, or develop and implement, your ELP standards and assessments. This documentation serves two important and related purposes: to help ensure that all criteria for quality are carefully considered and, once decisions have been made, to serve as evidence that all aspects of quality have been addressed.[3]

Recognizing that states are at different stages in developing and implementing their system of ELP standards and assessments, the Framework assumes "multiple points of entry." For example, depending on what work they have already done or feedback they have already received, states might focus their evaluation or their oversight efforts on both standards and assessments, on just one or the other, or, even, on one particular aspect of standards or assessment quality, such as utility. The tables in Appendix A give an overview of the criteria and related considerations by relevance to specific aspects of quality (Table 1) and by possible sources of evidence about how a state is addressing them (Table 2). Among states using the criteria to evaluate an existing ELP standards-and-assessment system, some may find that their standards do not meet the quality criteria; others may find that, while their standards are fine, more work needs to be done on implementing them with fidelity; and still others may find that their assessments are not sufficiently aligned to their standards. Thus, for those evaluating existing ELP standards and assessments, the worksheets are also the place to document next steps.

Background

The *No Child Left Behind Act of 2001* (NCLB) requires, for the first time, that states establish ELP standards and aligned assessments that reflect the language necessary for proficiency in English and for academic achievement. Although most states had already implemented standards and assessments for English language development, NCLB requires a more refined and research-based approach, oriented not just to English language proficiency in general but, more specifically, to proficiency that would enable students to achieve academically. As states have begun their efforts to comply with this requirement, many have requested assistance from the U.S. Department of Education (ED) on how to evaluate their current standards and assessments related to English learners. Key issues for which they have sought assistance include how to establish the rigor and quality of their standards; the technical quality of their assessments and valid use of assessment results; and alignment of their ELP standards, both to their ELP assessments and to the language requirements inherent in achievement of challenging academic content standards in the core subject areas required to be assessed under NCLB. In October 2006, ED made a priority the provision of resources to help states with development and implementation of their ELP standards and assessments.

This Framework is such a resource. It extends the strong knowledge base and research that exists on quality standards and assessments in general, as well as from the body of research and practice related to English language acquisition.[4] (For more information about the Framework's development, see Appendix B.)

Organization of the Framework

The "Getting Started" section of the Framework (pp. 5–9) presents a cross-disciplinary approach to evaluating or overseeing the development and implementation of ELP standards and assessments. A cross-disciplinary approach is necessary because high-quality ELP standards and assessments require input from a variety of disciplines (e.g., standards, assessment, language testing, psychometrics, applied linguistics, professional development, evaluation); a knowledge of Title I, Title III, and related programs; and the perspective of those who will be using the system. It then encourages states to carefully consider three fundamental questions, the answers to which influence the quality of a state's ELP standards and assessments: What is the intended purpose of our ELP standards and assessments, and how do we expect them to be used? Who are the EL students in our state and what are their relevant characteristics (e.g., languages, experiences, backgrounds)? How do we define the domain of English language proficiency (ELP) that we are teaching and testing, and what are the relevant characteristics (e.g., knowledge, skills, functions, modalities, register) of the ELP content?

The body of the Framework consists of two sections, the first focusing on ELP standards (pp. 10–20) and the second on ELP assessments (pp. 21–36). Each section outlines critical *criteria* for developing and implementing state ELP standards and assessments, respectively. The criteria are specific conditions that states should satisfy as they develop and implement their standards and assessments. For each criterion, the sections also present key *considerations,* that is, factors or information that states should account for when deciding how to address the particular criterion or evaluating the degree to which the criterion has been met. The sections also identify how each criterion relates to the overall quality of the standards or assessments.

Finally, the Framework includes two tables that are intended as quick references as you evaluate or oversee the development and implementation of your state's ELP standards and assessments. The first table (pp. 37–41) provides an overview of ELP standards and assessment criteria by relevance to quality. The second table (pp. 42–44) provides an overview of the criteria with examples of potential sources of documentation that may contain evidence related to each criterion.

As noted earlier, the Framework also includes worksheets to use in inventorying and documenting where your state is in meeting the criteria presented in the Framework. While the language in some of the worksheets is more oriented to using the criteria in an evaluation process, the language can be easily adapted to use in the development oversight process.

Key Characteristics of High-Quality Standards and Assessments

Irrespective of their domain (e.g., ELP, mathematics, history), to be considered high-quality, standards and assessments should be both *valid* and *reliable*, should be free from *bias* and *sensitivity* issues, and should have *utility*. All design, development, review, and implementation processes for state ELP standards and assessments should be set up to ensure these qualities, and those who are involved in the processes should establish and document evidence of these qualities, for example, through technical reports, state administrative code or law, meeting minutes from relevant committees (e.g., bias committee, technical advisory committee).

While these terms are commonly used in the field of standards and assessment development, there are slight variations in how they are defined and interpreted. The Framework's definition of these characteristics follows.

Validity of standards is the degree to which given standards appropriately define (for students, teachers, schools, and the community more broadly) what students should know and be able to do and, in turn, the degree to which the standards guide both the opportunities to learn and the resources that students are given, so as to achieve the particular knowledge and skills reflected in the standards. *Validity of assessment* is the degree to which an assessment measures what it is intended to measure and the extent to which the inferences made and actions taken on the basis of the assessment outcomes are accurate and appropriate. Generally, states should have evidence of content and construct validity, as well as evidence of the validity of interpretation and uses of results and of comparability of results across groups and time.

Reliability of standards is the degree to which standards and related documentation lend themselves to consistent understanding and implementation of defined skills and knowledge vis-à-vis both the targeted construct(s) and the standards' intended purpose and use. *Reliability of assessment* is the degree to which assessment results are dependable and consistently measure particular knowledge and skills. An assessment that is not reliable cannot be valid. Generally, states should have evidence of consistency across different tasks that are intended to measure the same knowledge and skills (item reliability, internal consistency) and, also, evidence of rater consistency (i.e., both intra- and interrater reliability). Additionally, states should have evidence of the precision of the assessment at cut scores, consistency of student-level classification, and degree of generalizability of results.

Bias and *sensitivity* are both related to fairness and equity. *Bias* is unfairness that results from any of various factors (e.g., related to culture, ethnicity, geography, gender, socioeconomic status) that interfere with the learning and performance of a group of students. *Sensitivity* is unfairness that results from any of various factors (e.g., death, violence, emotionally laden historical events) that offend or cause distress to a group of students and, consequently, may interfere with the learning and performance of that group. A state's standards and assessments should not advantage or disadvantage any group of students. Standards and assessments that are free of bias and sensitivity issues enable all students to access,[5] learn, and demonstrate the knowledge and skills they need for attainment of English language proficiency and the language necessary for achievement in the academic content areas. At minimum, all standards and assessments should be reviewed to identify and eliminate elements that may favor one group (e.g., language, culture, ethnicity) over another and, also, for their potential to inhibit inclusion of EL students with disabilities.

Utility of standards and assessments refers to the degree to which the standards and assessments are clear and easy to understand and the facility with which they can be consistently and accurately used for their intended purposes.

Some Cautions in Using the Framework

The Framework provides criteria intended to systematically guide the evaluation — or the oversight of development and implementation — of state ELP standards and assessments. The following cautions are intended to prevent misinterpretation and misapplication of the Framework.

- The Framework reflects an evolving area of knowledge; therefore, the information presented in the Framework — and any decisions made based on this information — may need to be refined over time, according to new research and emerging practices.

- The Framework's criteria are important to consider, but are not necessarily sufficient for ensuring the quality of state ELP standards, the technical adequacy of state ELP assessments, and full compliance with federal legislation.

- The Framework is not intended to prescribe specific methodologies for developing standards and assessments; rather, the information it generates and the criteria it provides are intended to inform decisions and activities relevant to the assurance of high-quality ELP standards and assessments.

- The Framework focuses on ELP standards and assessments; while both of these have implications for instruction, instruction is not a focus of this Framework, nor does the Framework suggest or assume any particular instructional approach or program for English language acquisition.

- The Framework was not developed and is not intended for use as a monitoring tool for compliance purposes.

Getting Started

The first step in planning an evaluation of your ELP standards and assessments, or the oversight of their development and their implementation, is to decide who should be involved in the process. Certain essential expertise needs to be represented, as does the range of perspective found in key user and stakeholder groups (e.g., site administrators, teachers, parents, business community). Because particular characteristics of the EL population and of the ELP domain should be considered in order to ensure the quality of the standards and the technical adequacy of the assessments, it is essential to have input from experts knowledgeable about Title III and Title I accountability requirements; English language acquisition and applied linguistics; language testing; psychometrics; standards, curriculum and instruction; professional development; and the state's EL population. Such expertise may be found in researchers, technical assistance providers, state department staff or consultants, policymakers, and classroom teachers. Bringing together this broad range of expertise and perspective provides an added value of creating a collaborative whose members share responsibility for deepening their understanding — both as a group and as individuals — of the foundations for high-quality ELP standards and assessments, particularly in terms of the domain the standards and assessments are intended to address and the student population they are intended to serve. In addition to helping ensure the quality of the evaluation or oversight work, a deepened understanding enhances the likelihood that the standards and assessments will be effectively implemented.

Composition of individual committees may need to differ depending on which aspect of standards and assessment development or implementation (e.g., domain definition vs. alignment vs. scoring vs. equating) a given committee is expected to evaluate or oversee. When convening committees, consider the following:

- Committee members should have deep expertise in areas relevant to standards, assessments, or both, as appropriate for their given development or implementation task.

- As a group, committee members should reflect an appropriate balance of representation and perspectives (e.g., from key experts, users, and stakeholder groups).

- Committee members should be appropriately oriented to and, if necessary, trained for the task(s) at hand.

- Committee members should have adequate time for carrying out the task(s) they are asked to complete (e.g., developing, reviewing, providing input or guidance).

Remember: The language in some of the Framework worksheets is oriented to the evaluation process. If you are using the Framework as part of the oversight process for developing and implementing new ELP standards and assessments, you may want to slightly reword some of the worksheets.

Use the following worksheet to consider and document the degree to which your state has ensured the appropriate range of expertise and perspective on the committees involved in the development and implementation of your ELP standards and assessments.

Committee Composition

What were the state's recruitment and membership selection criteria and protocols for the committees convened during development and implementation phases of its ELP standards and assessments?	What was the range of expertise and perspectives represented on each of the committees?	What documentation does the state have showing membership and the range of perspective and expertise represented on each committee?	What are related next steps for improving our ELP standards and assessments? (e.g., additional criteria/protocol/processes, additional committee members, additional documentation)?

Regardless of whether you are evaluating an existing ELP standards-and-assessment system or overseeing development of a new one, and whether you are starting with standards or with assessments, your answers to the following three questions should inform your work. The care with which they have been answered will influence the quality of the standards and assessments.

- What is the intended purpose of our ELP standards and assessments, and how do we expect them to be used?

- Who are the English learner (EL) students in our state and what are their relevant characteristics (e.g., languages, experiences, backgrounds)?

- How do we define the domain of English language proficiency (ELP) that we are teaching and testing, and what are the relevant characteristics (e.g., knowledge, skills, functions, modalities, register) of the ELP content?

The three worksheets that follow will help guide your review of how your state has answered these questions.

1. What is the intended purpose of our ELP standards and assessments, and how do we expect them to be used?

Statement of purpose and use for state's ELP standards and assessments:

What information did the state use to define the purpose and use of its ELP standards and assessments? Who was involved in defining the purpose and use?	To whom has the state communicated the purpose and use?	How has the state communicated the purpose and use? (Identify relevant documents)	What are related next steps for improving our ELP standards and assessments (e.g., develop more details regarding purpose and use, additional communication)?

2. Who are the English learner (EL) students in our state and what are their relevant characteristics (e.g., languages, experiences, backgrounds)?
See pages 10 and 21 of this document for additional related information.

Note: Determining who the students are does *not* mean that different standards or achievement expectations should be developed for different students; only one set of ELP standards and assessments should be developed for all EL students in a state.

Definition of the state's EL student population:

On what information, including statistical data, was this population definition based? Who was involved in defining the population?	To whom has the state communicated information about its EL student population?	How has the state communicated information about its EL student population? (Identify relevant documents)	What are related next steps for improving our ELP standards and assessments (e.g., gather more information about the population, create a refined population definition, generate additional communication)?

3. How do we define the domain of English language proficiency (ELP) that we are teaching and testing, and what are the relevant characteristics (e.g., knowledge, skills, functions, modalities, register) of the ELP content?

See pages 10–11 and 21 of this document for additional related information.

Definition of the ELP domain, including information related to the breadth, depth, and range of complexity of language skills and knowledge in the four recognized language modalities[6] of listening, speaking, reading, and writing, as well as information about how the modalities interrelate and contribute both to English language proficiency in general and, more specifically, to meeting the language requirements inherent in challenging state academic content:

On what (a) theory and (b) research was this domain definition based? Who was involved in defining the domain?	To whom has the state communicated information about the ELP domain?	How has the state communicated information about the ELP domain? (Identify relevant documents)	What are related next steps for improving our ELP standards and assessments (e.g., gather additional information about the domain, develop a refined domain definition, generate additional communication)?

Section1: Standards

This section presents critical *criteria*[7] for developing and implementing state ELP standards. The criteria are specific conditions that states should satisfy as they develop and implement their ELP standards. Key *considerations* are presented for each criterion. Considerations are factors or information that states should account for when making decisions about how to address a criterion or whether a criterion has been met. Information related to how each criterion contributes to overall quality also is presented (see p. 3 for definitions of key terms), as are examples of possible sources of evidence (e.g., technical reports, state administrative code/law) to examine when evaluating the quality of a state's ELP standards.

Issues Relevant to the Development and Implementation of High-Quality Standards for English Language Proficiency

Three overarching issues influence the development and implementation of high-quality ELP standards and, therefore, should be kept in mind as states review and improve their standards.

Issue 1: The English Learner Population

English learner students are diverse in their education history, level of literacy, background experiences, sociocultural practices, and socioeconomic status. This diversity[8] may affect the ways in which different groups of EL students access and interpret academic content in curriculum, instruction, and assessment.[9] Some EL students also will be eligible for a range of special education services, if they have been identified as having learning or other disabilities. In some cases, students may have disabilities that affect their development and proficiency in English.

As states evaluate or develop and implement their ELP standards (e.g., draft the standards, review them, plan professional development regarding the standards and related instruction), they will want to involve experts (e.g., consultants, teachers, researchers) who have familiarity with the range of diversity represented in the state's EL student population. These experts should be included with other technical experts (e.g., in applied linguistics, second-language acquisition, English language development) at appropriate points in the evaluation process or the development

oversight process for the ELP standards. Addressing the diversity of the EL student population does not mean there should be different standards or achievement expectations for different EL subgroups; only one set of rigorous ELP standards should be developed for all of a state's EL students. The critical point here is that *because* standards are major levers for improving student achievement (by articulating goals and focusing instruction), they should be accessible to all students. This means that a state's ELP standards should be conceptualized and designed to appropriately guide implementation — that is, to appropriately guide instruction that can be adapted and differentiated, as necessary, to meet the varying needs of the state's EL students and provide them with the support they require to successfully attain English language proficiency and the language necessary to achieve in the academic content areas.

See page 8 of this document for additional related information.

Issue 2: The Nature of the Domain: English Language Proficiency

Clear models (based on both theory and research) of *how EL students acquire language* and clear definitions of *what constitutes proficiency in English* are necessary for appropriately defining the domain of ELP. A clear definition of the ELP domain and the multiple constructs it encompasses is, in turn, essential to evaluating or developing standards that can help drive aligned curriculum, instruction, and assessment, and that can inform the appropriate interpretation of assessment results.[10]

This Framework encourages states 1) to examine the theoretical and research bases (e.g., theories of and research findings about second-language acquisition, English language development/acquisition, sociocultural development, sociolinguistics)[11] for their definition of English language proficiency, and 2) to evaluate the breadth, depth, range of complexity, and articulation of the associated skills and knowledge expected of students as they progress toward and achieve "proficiency." In defining the domain of English language proficiency, it is important to specify the required language skills and knowledge in the four recognized language modalities of listening, speaking, reading, and writing. It also is important to consider how the four modalities interrelate and contribute to English language proficiency, both in general and as required for

students to learn and demonstrate their knowledge of challenging academic content. Additionally, states should examine the degree of overlap between ELP knowledge and skills and the knowledge and skills required in their English language arts (ELA) standards, and should consider whether the existing overlap is intentional and appropriate. They should also examine their respective expectations for proficiency in the two different, but related, domains of ELP and ELA.

Clarity of the theoretical and research-based underpinnings of a state's definition of English language proficiency, as well as the careful and purposeful use of this foundation to inform its standards development, is critical to ensuring high-quality ELP standards. See page 9 of this document for additional related information.

Issue 3: State Context Factors: History, Policies, Resources

A number of political and resource issues influence the development and implementation of a state's standards. Among these contextual factors are system readiness; familiarity with the population; past and current policies, practices, and programs; and structures and resources to support the instruction and assessment of the population. How these particular factors play out will vary from state to state.

For example, states that historically have had large numbers of EL students, such as California, Texas, and Florida, are likely to have more mature policies, practices, and structures in place and have different levels of resources available to support EL students' achievement compared to states that have only recently experienced significant increases in their EL student population, such as South Carolina, Kentucky, and Indiana. These and other "rapid-growth" states experienced an increase of more than 400 percent in their PreK–12 public school EL student population from school year 1994–95 through school year 2004–05.[12]

This Framework encourages states to identify and minimize any limiting influences of such contextual factors so that the effectiveness of their ELP standards and implementation will not suffer and *real* student achievement will be realized more fully. Individuals with knowledge of and experience with such contextual factors (e.g., policymakers; state and local administrators; state specialists in standards, assessment, curriculum, and instruction; professional development and technical assistance providers; researchers) should be involved at critical points in the development, review, and implementation of the state's ELP standards.

Standards Development Phase
CRITERIA AND RELATED CONSIDERATIONS

1.1 Organization or structure of the standards

1. The structure of the state's ELP standards (e.g., format, organization/hierarchy, levels of detail) is appropriate for the standards' instruction- and assessment-related (including reporting) purposes and uses.

1.2 Number of standards

1. The number of standards is appropriate for the depth and breadth of the ELP domain, as it is defined by the state.

2. The number of standards allows for appropriate coherence and consistency of skills and knowledge across modalities (i.e., listening, speaking, reading, writing), as defined by the state.

1.3 Level of specificity, or "granularity," of the standards

1. The state's ELP standards are described with sufficient clarity and definition to guide curriculum development, instructional planning, and assessment development for the EL population.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Do the structure, number, and level of specificity of the standards support the purpose(s) and use(s) of the standards in curriculum, instruction, assessment, and reporting?	• Validity • Utility	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
b. Are the structure, number, and level of specificity of the standards consistent with the theoretical and research bases on which the state has developed its standards? Note: Your standards should address both general ELP skills and knowledge and academic language[13] necessary for success in content areas.	• Validity	• _____ • _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ • _____
c. Are the relationships among the language modalities (i.e., listening, speaking, reading, writing) and how they contribute to overall English language proficiency clear? — The degree of overlap of language skills and knowledge versus delineation within and across language modalities should be purposeful.	• Validity • Utility	• _____ • _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials • Parent Resources	• _____ • _____
d. Are the expectations for increasing language proficiency levels in English and how they build upon skills in preceding levels and lead to skills in subsequent levels clear? — The number of proficiency levels is appropriate and lends itself to the articulation of progress toward and attainment of "proficiency." And, the degree of overlap of language skills and knowledge versus delineation within and across language proficiency levels, as well as within and across grade ranges, should be purposeful.	• Validity • Utility	• _____ • _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials • Parent Resources	• _____ • _____
e. Are the standards written with the appropriate level of detail? The level of detail (e.g., supporting skills, indicators, enablers, benchmarks) should be necessary and sufficient for ensuring accurate and consistent understanding and implementation.	• Validity • Reliability • Utility	• _____ • _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ • _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?

f. Is the wording of the standards statements precise, clear, and consistent enough to facilitate implementation and measurement of the standards?	• Validity
	• Reliability
	• Utility
		• _____
		  _____
		• _____
		  _____
			• Manuals/Guides
			• Technical Reports and Research Studies
			• Meeting Reports/Minutes
				• _____
				  _____
				• _____
				  _____

g. Do the standards appropriately differentiate between statements of expected student learning outcomes (the *what*) and suggestions for curriculum/instruction (the *how*)? — Standards should focus on the former.	• Validity
	• Utility
		• _____
		  _____
		• _____
		  _____
			• Manuals/Guides
			• Technical Reports and Research Studies
			• Meeting Reports/Minutes
			• Training and Professional Development Materials
				• _____
				  _____
				• _____
				  _____

Standards Development Phase, continued
CRITERIA AND RELATED CONSIDERATIONS

1.4 Alignment[14]

1. The state's ELP standards are articulated horizontally.

2. The state's ELP standards are articulated vertically.

3. The state's ELP standards are linked with the state's academic content standards.[15]

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?

a. To ensure *horizontal* articulation, does the state have a process for making sure that skills and knowledge called for in the standards are appropriate for the language proficiency levels and grade ranges for each language modality? More specifically, does the process ensure that
 	
• the levels of complexity of language skills and knowledge across language modalities are comparable within a grade range, unless differences are supported by theory and/or research?
 	
• any repetition of language skills and knowledge across the language modalities within a given language proficiency level and grade range is purposeful and meaningful?
	• Validity
	
	• Utility
		• _____
		  _____
		
		• _____
		  _____
			• State Administrative Code/Law/Policy
			
			• Technical Reports and Research Studies
			
			• Meeting Reports/Minutes
				• _____
				  _____
				
				• _____
				  _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?

b. To ensure *vertical* articulation, does the state have a process for making sure that, as skills and knowledge called for in the standards move from lowest language proficiency level and grade level or grade range to highest,
 	
• levels of language skills and knowledge appropriately increase in complexity, and the incremental increase is based on theory and/or research?
 	
• prerequisite language skills and knowledge appear as appropriate in lower language proficiency levels and grade ranges?
 	
• broader, deeper, and new language skills and knowledge appear in higher language proficiency levels and grade ranges (building from skills and knowledge in lower/prior levels) and are introduced at the appropriate language proficiency level and grade range?
 	
• language skills and knowledge that are expected to be acquired are explicitly stated at the appropriate language proficiency level and grade range?
 	
• the balance of representation — coverage and complexity of language skills and knowledge — shifts appropriately across language proficiency levels and grade range?
 	
• any repetition of standards (i.e., language skills and knowledge to be developed) is purposeful, and it is clear what English language skill/knowledge is to be acquired and when it is to be acquired?
	• Validity
	• Utility
		• _____
		  _____
		• _____
		  _____
			• State Administrative Code/Law/Policy
			
			• Technical Reports and Research Studies
			
			• Meeting Reports/Minutes
				• _____
				  _____
				• _____
				  _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
c. Does the state have a process for ensuring a match (linkage) between ELP language skills and knowledge in each of the language modalities (i.e., listening, speaking, reading, writing) and the language skills students need to have in order to achieve in the academic content areas? More specifically, does the process ensure that • all relevant aspects of grade-level academic content language are identified and considered: forms and functions, as well as language skills consistent with higher-order thinking skills and understanding necessary for achievement in the content areas? • the ways in which and degree to which language and content converge/diverge are clear[16]? • minimally, language skills and knowledge defined in the ELP standards are evaluated according to the degree to which they: » cover relevant aspects of grade-level academic content language (i.e., breadth, depth, complexity — as described previously) in the state's academic content standards; » cover the range of language complexity necessary for achievement in the content areas, including levels of language complexity consistent with higher-order thinking skills and understanding reflected in the state's academic content standards; and » reflect a balance of representation — coverage of language skills and knowledge and complexity of language skills and knowledge — consistent with the language needed for achievement of grade-level academic content reflected in the state's academic content standards?	• Validity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Standards Implementation Phase
CRITERIA AND CONSIDERATIONS

1.5 Training

1. The state provides guidance and training to local education agencies — for example, to teachers of English as a Second Language, bilingual teachers, content area teachers, special education teachers, school and district administrators — on the ELP standards, their purpose and use, and implementation strategies.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Does the state have a process and schedule for communicating the purpose and use of its ELP standards to local education agencies (for example, to teachers of English as a Second Language, bilingual teachers, content area teachers, special education teachers, school and district administrators)? To the degree appropriate and feasible, is the training coordinated? That is, does the training provided by the state build on/make use of existing structures and systems (at state and local education agency levels)?	• Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____
b. Do state materials illustrate/describe how the ELP standards are linked to language requirements (i.e., language skills and knowledge, including academic language) of grade-level content standards for curriculum and instructional purposes, including language skills consistent with higher-order thinking skills and understanding necessary for achievement in the content areas?	• Utility	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
c. Does state documentation include descriptions and/or examples of key language competencies for each language proficiency level and each language modality across the state's grade ranges, including examples of academic language necessary for achievement in the content areas? More specifically, are the expectations clear about what English language skills/knowledge are to be acquired and when they are to be acquired?	• Utility	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Training and Professional Development Materials • Parent Resources	• _____ _____ • _____ _____
d. Is there a clear distinction made about what language skills and knowledge are appropriately assessed at the state level versus the local level? Note that some standards may be difficult to assess in a technically defensible or practical manner on a statewide assessment, such as standards that require student production of complex language samples that necessitate a combination of skills across language modalities.	• Validity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Manuals/Guides • Training and Professional Development Materials	• _____ _____ • _____ _____

Standards Implementation Phase, continued

CRITERIA AND CONSIDERATIONS

1.6 Monitoring and evaluating

1. The state has systems and structures for monitoring and evaluating the implementation of its ELP standards in local education agencies, schools, and classrooms.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Does the state have a process and schedule for monitoring the implementation of the state's ELP standards, including a plan for evaluating the guidance and training it provides to local education agencies and teachers? To the degree appropriate and feasible, is this process and schedule a part of/combined with the state's existing monitoring and evaluation systems and structures?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/ Policy • Meeting Reports/Minutes	• _____ _____ _____ • _____ _____
b. Does the state rely on multiple sources of data/information (e.g., internal and external monitoring, qualitative data/ analyses, quantitative data/analyses) for evaluating the fidelity with which its ELP standards are implemented and the effectiveness of the implementation?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ _____ • _____ _____

Section 2: Assessment

Research suggests that during development and implementation of an assessment, different types of information are needed, specifically evidence of validity, reliability, and freedom from bias.[17] Section 2 focuses on ELP assessments and presents critical *criteria*[18] for a phase-by-phase approach to development and implementation. The criteria are specific conditions that states ought to satisfy as they develop and implement their ELP assessments. Key *considerations* are presented for each criterion. Considerations are factors or information that states ought to account for when making decisions about how to address a criterion or whether a criterion has been met. Information related to the relevance of each criterion to overall quality also is presented (see p. 3 for relevant definitions), as are examples of forms of evidence and documentation (e.g., technical reports, state administrative code/law, committee meeting minutes) related to the quality of a state's ELP assessment.

Issues Relevant to the Development of High-Quality English Language Proficiency Assessments

Five overarching issues influence the development and implementation of high-quality ELP assessments and, therefore, should be kept in mind as states review and improve their assessments.

Issue 1: The English Learner Population

English learner students are diverse in their education history, background experiences, sociocultural practices, and socioeconomic status. This diversity[19] may affect how different groups of EL students access and interpret academic content in curriculum, instruction, and assessment.[20] Some EL students also will be eligible for a range of special education services, if they have been identified as having learning or other disabilities. Addressing the diversity of the EL student population does not mean that states should have different assessments or achievement expectations for different EL subgroups. Rather, the state's ELP assessments should be sensitive to the nature of the individual differences within this diverse population, so that the assessments are accessible to all students. That is, the state's ELP assessment and its test items (e.g., graphics, cultural/situational references) should be developed with consideration for the varied backgrounds and experiences (e.g., linguistic, cultural, geographic) of EL students, avoiding any graphics, cultural or situational references, or other factors that could confuse students or otherwise interfere with their access to and performance on the assessment. States are encouraged, for example, to carefully design and execute their assessment pilot tests and field tests so that they include an appropriately representative sample of students, including EL students with disabilities. See page 8 and page 10 of this document for additional related information.

Issue 2: The Nature of the Domain: English Language Proficiency

Clarity of the theoretical and research-based underpinnings of a state's definition of English language proficiency, as well as the careful and purposeful use of this foundation to inform its assessment development, is critical to ensuring high-quality ELP assessments that fairly and accurately meet all intended purposes of the state's assessment system. See page 9 and pages 10–11 of this document for additional related information.

Issue 3: Similarities and Differences in the Applicability of Technical Criteria

Research suggests that some of the common methods used to develop assessments for general student populations may *not* directly transfer to developing assessments for special student populations, such as EL students, or, if they do, they may need to be considered in a different manner for establishing the technical adequacy of assessments for special student populations (e.g., ELP assessments).[21] Thus, procedures and criteria widely used for technical reviews of assessments for non-EL student populations, including establishing the validity of these assessments, may need to be adapted before they are applied to assessments for more specialized populations, such as EL students.

This does not mean that current psychometric processes are insufficient. Rather, it means that some standard practices may need to be adapted because of the particular characteristics of this population and of the types of tests currently used to assess EL students for Title III purposes. Just as there are important differences in how to appropriately develop and validate assessments for general education students

without disabilities versus students with disabilities,[22] there also may be important differences in developing and validating assessments for EL students. For example, ELP tests typically have performance-based sections that can suffer from small sample sizes and affect statistical analyses due to bias or reliability estimates.[23]

This Framework encourages states to consider the specific characteristics of their EL student population and the nature of the ELP domain as they:

- articulate the purpose(s) and role of the ELP assessments in the state's assessment system;

- develop item and test design specifications and test blueprints;

- define the assessment's sampling/norming group and groups with whom they may want to pilot items/tasks;

- plan the state assessment's field test design and statistical analyses (including differential item functioning and other bias identification methodologies);

- establish criteria and protocols for bias and sensitivity reviews (e.g., linguistic, sociocultural) and specify the full range of experts and committee members, within and outside of the education community, necessary for such reviews;

- develop administration and scoring protocols and specify the qualifications of those needed for test administration and scoring;

- develop a standard-setting protocol and articulate language-proficiency-level descriptors consistent with the language skills and knowledge specified in the state ELP standards; and

- monitor and evaluate the implementation of the state ELP assessment, including examining the burden on teachers and local education agencies of the administration, scoring, reporting, and security protocols.

Many of the procedures and criteria that have been found to be appropriate and essential for establishing the technical adequacy of assessments for EL students are similar to those relevant to the design, development, implementation, and evaluation of state academic content assessments. Therefore, this Framework encourages states, whenever possible, to build on, or extend, their current systems (e.g., for Title I)

in order to address requirements for their ELP assessments. This will help support a more technically sound, coherent state assessment system. As states routinely review and revise Title I and Title III assessments, opportunities to create greater coherence between the two may occur.[24]

Issue 4: Access and Accommodations

Appropriate access[25] and accommodations[26] for EL students on ELP assessments should not result in the simplification or alteration of the test's targeted constructs or the misrepresentation of students' achievement of ELP. Rather, methods for providing access and selecting allowable accommodations for EL students should be based on sound theory and practice, as well as on research-based evidence (e.g., second language acquisition, English language development/acquisition, language testing, measurement, Universal Design). These methods should be sensitive to the specific characteristics and needs of this population of students and the nature of the assessed domain, in this case, ELP.

Diverse as its individual members may be, the EL student population as a whole has some unique characteristics that influence how to assess the population and validate the assessment results; these characteristics include, for example, level of familiarity with American cultural references and the effect of a primary or home language on the learning of English.[27] Additionally, the ELP domain itself has some characteristics that should be kept in mind when developing an assessment. It is important to carefully consider how this domain is both similar to and different from academic content domains, such as English language arts or social studies (e.g., the degree to which the domain is defined in terms of skills and functions versus topics and concepts, and the degree to which development in the domain is expected to progress along a continuum versus to be discontinuous). Furthermore, as EL students develop ELP, their language abilities often develop at different rates across language modalities (i.e., listening, speaking, reading, writing), providing another layer of complexity when striving to ensure that EL students have appropriate access to test content and valid measurement of their ELP skills and knowledge.

This Framework encourages states to consider factors that may interfere with their EL students' ability to access the tested content (e.g., unfamiliarity with assessment item

or task format or presentation, with the sociocultural context of an item or task, or with the graphics associated with an item or task). Appropriate access for EL students (whether with or without disabilities) maximizes their understanding of what is being asked of them and enables them to fully demonstrate their ELP skills and knowledge across the language modalities. Some efforts to improve access can compromise the integrity of an assessment. For example, the use of oral prompting during a speaking test requires listening comprehension, and this listening requirement could confound the results of the speaking assessment. Thus, states also should consider and clearly specify strategies that appropriately facilitate access in each language modality without compromising the integrity of the assessments and their targeted content/constructs.

With regard to accommodations on ELP assessments, some EL students will be eligible for special education services if they have been identified as having learning or other disabilities, and they will require assessment accommodations per their Individualized Education Programs (IEPs). Therefore, states are encouraged to carefully articulate their protocols guiding the provision of accommodations to those students who need them for their ELP assessments. It is critical that, as for other student subgroups, states allow no accommodations that would alter the assessed constructs. Additionally, states should consider and clearly specify modality-appropriate accommodations that facilitate access for EL students without compromising the integrity of the assessments and their targeted content/constructs across the language modalities (e.g., more time, larger print, customized glossaries for construct-irrelevant vocabulary).

Issue 5: State Context Factors: History, Policies, Resources

As is also true of standards development and as was discussed in Section 1 of this Framework, a number of political and resource factors affect the development of ELP assessments and their implementation. See page 11 of this document for additional related information.

Assessment Development Phase
CRITERIA AND CONSIDERATIONS

2.1 Specifications

1. Specifications (i.e., item, graphic/stimulus, passage, test form) accurately translate the state's intent of the standards with sufficient detail to guide the development of items and tests that assess the standards.

2. Item specifications accurately represent the state's intent for the standards with sufficient detail to ensure consistent understanding of the standards across key groups of participants involved in the test's development (e.g., teachers, test developers, review committee members).

2.2 Test blueprint

1. The test blueprint communicates the structure and content of the test (e.g., breadth, depth, range of complexity, emphasis, item formats) in a manner consistent with the intent of the state.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Do the test blueprint and specifications support the inclusion of appropriate language skills and knowledge at each grade range and language proficiency level, and for each language modality necessary for attaining English language proficiency *and* for the achievement of challenging state academic content?	• Validity	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Training and Professional Development Materials	• _____ _____ • _____ _____
b. Do the test blueprint and specifications support the measurement of language skills and knowledge described in the state's ELP standards and *not* knowledge, skills, or other characteristics that are not specified in the ELP standards for all grade ranges, language proficiency levels, and language modality expectations?	• Validity	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Training and Professional Development Materials	• _____ _____ • _____ _____

Assessment Development Phase, continued

CRITERIA AND CONSIDERATIONS

2.3 Alignment

1. The state's ELP assessments are aligned with the state's ELP standards.

2. The language skills and knowledge included in the state's ELP assessments are relevant to English language acquisition and English language proficiency and necessary for students' achievement of challenging state academic content.

3. The state's ELP assessments are aligned with the state's language proficiency level descriptors.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Has the state conducted an alignment study or studies of • the degree to which its ELP standards and assessments are aligned comprehensively, meaning that the assessments reflect the full range (breadth, depth, complexity) of the ELP standards, including the range of language skills and knowledge needed for achievement in the content areas? • the degree to which the pattern of emphasis of knowledge and skills is similar between the state's ELP standards and assessments? *Note: Independent alignment studies are recommended.*	• Validity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
b. Does the state's alignment study include criteria for identifying and defining the academic language in the state's ELP standards and its academic content standards in order to determine the degree to which they correspond in terms of the language knowledge and skills associated with higher-order thinking and understanding challenging academic content?	• Validity	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?

c. Is "Proficiency" — overall, as well as in each of the language modalities (as appropriate) — defined in a manner consistent with the ELP standards? That is, do the state's proficiency level descriptors[28] fully reflect its ELP standards for each grade range and language modality and describe the language skills and knowledge expectations for each language proficiency level from "beginning" through "advanced," including appropriate vertical and horizontal articulation across language modalities necessary for progress toward and achievement of English language proficiency?
 	
Additionally, does each language proficiency level descriptor clearly define the language skills and knowledge for the attainment of that level of proficiency? Are the language proficiency level descriptors consistent with the ELP standards in terms of wording, structure, and use? Do differences across language proficiency level descriptors reflect real differences in language skill and knowledge expectations (e.g., along the expected continuum of English language acquisition) vis-à-vis proficiency and progress toward proficiency as well as functional differences (e.g., for reporting purposes)?
	• Validity
	
	• Reliability
		• _____
		  _____
		• _____
		  _____
			• State Administrative Code/Law/Policy
			
			• Manuals/Guides
			
			• Technical Reports and Research Studies
			
			• Meeting Reports/Minutes
			
			• Training and Professional Development Materials
			
			• Parent Resources
				• _____
				  _____
				• _____
				  _____

Assessment Development Phase, continued
CRITERIA AND CONSIDERATIONS

2.4 Bias and sensitivity

1. The state's ELP assessments are fair and accessible to students, regardless of, for example, their gender, culture, ethnicity/race, socioeconomic status, geographical location, and primary language.

2. Bias and sensitivity issues have been examined through both qualitative analyses (e.g., expert judgment, cognitive interviews) and quantitative analyses (e.g., Differential Item Functioning).

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Does the state have both judgment-based (e.g., committee reviews) and data-based (e.g., DIF studies) procedures/processes at key points throughout assessment development and implementation to ensure that its ELP assessments are fair and accessible to all EL student subgroups for which they are intended, including EL students with disabilities? Does the state have documentation (e.g., training materials, meeting minutes) of the criteria and procedures used to judge the fairness and accessibility of its assessments, including relevant research to support the criteria and procedures?	• Validity • Bias/Sensitivity	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____
b. Does the state provide an appropriate variety of accommodations for EL students with disabilities who are taking the assessment? *and* Do state materials (e.g., training materials, administration and technical manuals) clearly define the policies and procedures for selecting accommodations for use with the state's ELP assessments, including specifications that clearly delineate which accommodations may be used for specific sections of the test (e.g., accommodations appropriate for each language modality tested)? *and* Does the use of allowable accommodations yield valid and meaningful scores?	• Validity • Reliability • Bias/Sensitivity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Manuals/Guides • Technical Reports and Research Studies • Training and Professional Development Materials • Parent Resources	• _____ _____ • _____ _____

Assessment Implementation Phase
CRITERIA AND CONSIDERATIONS

2.5 Administration

1. The state provides guidance and training to test administrators and coordinators.
2. The state has systems and structures for ensuring standardization and fidelity of administration.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Does the state have a system of training and monitoring to ensure that each person who is responsible for handling or administering any portion of its assessments (e.g., teachers, school and district administrators) does so in a way that protects the security of the assessments and maintains fidelity to and equivalence of administration conditions across students and schools?	• Validity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Manuals/Guides • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____
b. Does the state have a clearly stated test security policy and consequences for violation that are communicated to the public and to local educators?	• Validity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Manuals/Guides • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____
c. Does the state have a system of training for and monitoring those who will provide accommodations to students who qualify for them?	• Validity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____

Assessment Implementation Phase, continued
CRITERIA AND CONSIDERATIONS

2.6 Scoring

1. The state provides guidance and training to scorers.

2. The state has systems and structures for ensuring quality and accuracy of scoring.

2.7 Reporting

1. The state's ELP assessments yield coherent and valid information for its EL population, grade ranges, and language proficiency levels.

2. The state's reporting system facilitates appropriate, credible, and defensible interpretation of its ELP assessment data that is meaningful and usable for a range of users (e.g., teachers, parents, state boards of education, federal government).

3. The state reports participation and assessment results for all EL students in its reports at the school, local education agency, and state levels, and the reports for any group or subgroup do not reveal personally identifiable information about individual students.

4. The state produces interpretive guidance following each administration of its ELP assessments.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Are the scoring and reporting structures consistent with the state's definition of English language proficiency, as well as with the definition of each language modality (i.e., listening, speaking, reading, writing) and comprehension, and their structures (i.e., item interrelationships are consistent with the framework from which the test arises)? Is the number of score points on the state's ELP assessments consistent with the representation of language skills and knowledge within and across each language modality?	• Validity • Reliability	• _____ • _____	• State Administrative Code/Law/Policy • Manuals/Guides • Technical Reports and Research Studies	• _____ • _____
b. Does the state have a process for producing itemized score analyses (e.g., by language modality, by language skill clusters) so that parents, teachers, and principals can interpret and address the specific English language needs of students?	• Validity • Utility	• _____ • _____	• Manuals/Guides • Training and Professional Development Materials • Parent Resources	• _____ • _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
c. Are the test scores related to external variables as intended (e.g., scores are correlated strongly with relevant measures of English language proficiency and are weakly correlated, if at all, with irrelevant characteristics/constructs)?	• Validity • Reliability • Bias/ Sensitivity	• _____ • _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ • _____
d. Does the state's ELP assessment system yield coherent and valid information for its EL student population, grade ranges, and language proficiency levels, and yield information necessary for Title III instruction and accountability purposes, including information needed to determine Annual Measurable Achievement Objectives (AMAOs)[29]?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ • _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ • _____
e. Do all score reports include error bands and guidance for proper and improper use of score reports?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ • _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials • Parent Resources	• _____ • _____
f. Does the state provide sufficient training for scoring, reporting, and interpreting results (e.g., to content teachers, ESL/bilingual teachers, school and district administrators, policymakers) to ensure consistency and accuracy of scoring and to ensure that reports for any group or subgroup do not reveal personally identifiable information about individual students, and to ensure that reports are accurately and appropriately interpreted?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ • _____	• Manuals/Guides • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ • _____
g. Does the state routinely collect information about the usability of its score reports, beginning with focus groups of the users (e.g., teachers, administrators, parents, policymakers) and including studies of proper and improper use?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ • _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ • _____

Assessment Implementation Phase, continued
CRITERIA AND CONSIDERATIONS

2.8 Standard setting

1. The state uses a validated standard-setting process[30] that results in language proficiency levels, descriptions of language competencies for each level, and cut scores that appropriately differentiate among English language skills and knowledge in each language modality and in each grade range. These language proficiency levels and associated descriptions of language competencies reflect a progression toward students' English language proficiency, as the state has defined it in its ELP standards and as is consistent with the theory and research on which the standards are based.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Is the standard-setting approach appropriate for the state's reporting and use of the ELP assessment results? For example, if separate language modality scores are reported with proficiency levels (i.e., performance categories), does the state use an appropriate standard-setting approach?	• Validity • Reliability • Bias/Sensitivity • Utility	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
b. Is the state's standard-setting approach properly implemented and documented? Are all participants fully and adequately trained, agendas of meetings maintained, and participant evaluation forms collected and reviewed?	• Validity • Reliability • Bias/Sensitivity • Utility	• _____ _____ • _____ _____	• Manuals/Guides • Technical Reports and Research Studies • Meeting Reports/Minutes • Training and Professional Development Materials	• _____ _____ • _____ _____
c. Do the state's ELP assessments have sufficient items/tasks (i.e., score points) at each language proficiency level to allow students to demonstrate the full range of language skills and knowledge specified in the state's ELP standards?	• Validity • Reliability • Utility	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?

d. Do the state's ELP assessments yield scores that reflect the full range of English language proficiency reflected by the state's ELP standards, and are assessment results expressed in terms of the language proficiency levels, not just scale scores or percentiles?
Do the state's ELP assessments yield scores that are clearly aligned with the state's language proficiency levels overall and across each language modality and grade range?
	• Validity
	• Reliability
	• Utility
		• _____
		  _____
		• _____
		  _____
			• State Administrative Code/Law/Policy
			• Manuals/Guides
			• Technical Reports and Research Studies
			• Training and Professional Development Materials
				• _____
				  _____
				• _____
				  _____

e. Is "Proficiency" — overall, as well as in each of the language modalities (as appropriate) — defined in a manner consistent with the ELP standards? Does the "proficient" designation represent the attainment of language skills and knowledge expectations for English language proficiency as defined by the state and based on relevant theory and research?	• Validity
	• Reliability
	• Utility
		• _____
		  _____
		• _____
		  _____
			• State Administrative Code/Law/Policy
			• Manuals/Guides
			• Technical Reports and Research Studies
			• Meeting Reports/Minutes
			• Training and Professional Development Materials
			• Parent Resources
				• _____
				  _____
				• _____
				  _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
f. Are state documentation and interpretive materials related to its language proficiency levels complete, accurate, and clearly written so that the range of users and the general public can readily understand the information? Does the documentation address • number of language proficiency levels and the labels for each level; • the relative contribution of each language modality and overall performance (total score) to judgments of proficiency; • the standard-setting methodology, including the method(s) used to determine cut scores; • descriptions of the individuals involved in the standard-setting process, including the articulation of the proficiency level descriptors; and • the validation of the state language proficiency levels against external measures (e.g., performance on content assessments)?	• Validity • Reliability • Utility	• _____ _____ • _____ _____	• Manuals/Guides • Training and Professional Development Materials • Parent Resources	• _____ _____ • _____ _____

Assessment Implementation Phase, continued
CRITERIA AND CONSIDERATIONS

2.9 Equating

1. The state uses a defensible equating methodology[31] that ensures that results are comparable across administrations and forms of the assessments.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Does the state use an equating model and methodology that are appropriate for the targeted purposes, population, and domain definition for its ELP assessments?	• Validity • Reliability • Bias/Sensitivity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
b. Does the state produce test forms and administration procedures that are appropriate for its selected equating model (e.g., common items, common students, placement of anchor items in test booklets)?	• Validity • Reliability • Bias/Sensitivity • Utility	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
c. Is the equating carefully reviewed, either using third-party validation or "real-time" review of equating results?	• Validity • Reliability • Bias/Sensitivity • Utility	• _____ _____ • _____ _____	• Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
d. Are the equating model and results reviewed periodically by the state's Technical Advisory Committee or other external groups?	• Validity	• _____ _____ • _____ _____	• State Administrative Code/Law/Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Assessment Implementation Phase, continued

CRITERIA AND CONSIDERATIONS

2.10 Monitoring and evaluating

1. The state has systems and structures for monitoring and improving the quality of its assessment, including a plan for ongoing procedures to maintain and improve alignment over time between the state's ELP assessments and ELP standards.

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
a. Is the state implementing ongoing quality control reviews to ensure that the system remains fully aligned over time?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/ Policy • Meeting Reports/Minutes	• _____ _____ • _____ _____
b. Does the state rely on multiple sources of data/information (e.g., internal and external monitoring, qualitative data/ analyses, quantitative data/analyses) for evaluating the quality and effectiveness of its ELP assessments? Does the state have a process for using the information gained through its series of studies related to validity, reliability, fairness/ accessibility, and alignment/linkage to eliminate gaps and address weaknesses, and does the state have a plan for regular quality review?	• Validity • Reliability • Bias/ Sensitivity • Utility	• _____ _____ • _____ _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____
c. Does the state have a process and schedule for monitoring the implementation of its ELP assessments and related consequences? Does the state help to ensure valid inferences and interpretation of assessment results? Do the state ELP assessments produce intended consequences, and have unintended consequences been considered and proactively and appropriately addressed?	• Validity • Bias/ Sensitivity • Utility	• _____ _____ • _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Consideration(s)	Relevance	What has the state done to address this consideration? How is it documented (specify document title and document date)?	Possible sources of evidence	What are next steps for addressing this consideration?
d. Does the state have a plan and schedule for routinely monitoring the extent to which accommodations used during ELP assessment administration are consistent with those specified in the IEP or 504 plans for eligible EL students with disabilities? Does the state have a process for examining its accommodations in terms of • their appropriateness vis-à-vis EL students with disabilities who are eligible for the state ELP assessment, and the degree to which the use of the accommodations is consistent with instructional approaches for each student, as determined by a student's IEP or 504 plan; • their impact on the assessed constructs (i.e., English language proficiency skills and knowledge) and the inferences based on student performance on accommodated ELP assessments; • the degree to which scores for EL students with disabilities that are based on accommodated administration conditions allow for valid inferences about these students' English language proficiency skills and knowledge and can be combined meaningfully with scores from non-accommodated administration conditions; and • the degree of effectiveness of specific accommodations for different groups of EL students with disabilities?	• Validity • Reliability • Bias/ Sensitivity	• _____ _____ • _____ _____	• State Administrative Code/Law/ Policy • Technical Reports and Research Studies • Meeting Reports/Minutes	• _____ _____ • _____ _____

Appendix A

Overviews of Criteria and Considerations by Relevance to Specific Aspects of Quality and by Possible Sources of Evidence About How a State is Addressing Them

Table 1: Overview of ELP Standards and Assessment Criteria by Relevance to Quality

	Criterion and Consideration	Relevance*			
		Validity	**Reliability**	**Bias/ Sensitivity**	**Utility**
S T A N D A R D S	1.1, 1.2, 1.3	X	X		X
	a.	X			X
	b.	X			
	c.	X			X
	d.	X			X
	e.	X	X		X
	f.	X	X		X
	g.	X			X
	1.4	X			X
	a.	X			X
	b.	X			X
	c.	X			

*There is an interrelationship among validity, reliability, bias/sensitivity, and utility. The notation in Table 1 for each criterion and consideration is intended to provide users of this Framework with a general notion of the relevance of each criterion and consideration to overall quality.

Table 1: Overview of ELP Standards and Assessment Criteria by Relevance to Quality, continued

	Criterion and Consideration	Relevance*			
		Validity	Reliability	Bias/ Sensitivity	Utility
S T A N D A R D S	1.5	X			X
	a.				X
	b.				X
	c.				X
	d.	X			X
	1.6	X	X	X	X
	a.	X	X	X	X
	b.	X	X	X	X

*There is an interrelationship among validity, reliability, bias/sensitivity, and utility. The notation in Table 1 for each criterion and consideration is intended to provide users of this Framework with a general notion of the relevance of each criterion and consideration to overall quality.

Table 1: Overview of ELP Standards and Assessment Criteria by Relevance to Quality, continued

	Criterion and Consideration	Relevance*			
		Validity	Reliability	Bias/ Sensitivity	Utility
A S S E S S M E N T	2.1, 2.2	X			
	a.	X			
	b.	X			
	2.3	X	X		
	a.	X			
	b.	X			
	c.	X	X		
	2.4	X	X	X	
	a.	X		X	
	b.	X	X	X	
	2.5	X			X
	a.	X			
	b.	X			X
	c.	X			

*There is an interrelationship among validity, reliability, bias/sensitivity, and utility. The notation in Table 1 for each criterion and consideration is intended to provide users of this Framework with a general notion of the relevance of each criterion and consideration to overall quality.

Table 1: Overview of ELP Standards and Assessment Criteria by Relevance to Quality, continued

	Criterion and Consideration	Relevance*			
		Validity	Reliability	Bias/ Sensitivity	Utility
A S S E S S M E N T	2.6, 2.7	X	X	X	X
	a.	X	X		
	b.	X			X
	c.	X	X	X	
	d.	X	X	X	X
	e.	X	X	X	X
	f.	X	X	X	X
	g.	X	X	X	X
	2.8	X	X	X	X
	a.	X	X	X	X
	b.	X	X	X	X
	c.	X	X		X
	d.	X	X		X
	e.	X	X		X
	f.	X	X		X

*There is an interrelationship among validity, reliability, bias/sensitivity, and utility. The notation in Table 1 for each criterion and consideration is intended to provide users of this Framework with a general notion of the relevance of each criterion and consideration to overall quality.

Table 1: Overview of ELP Standards and Assessment Criteria by Relevance to Quality, continued

	Criterion and Consideration	Relevance*			
		Validity	Reliability	Bias/ Sensitivity	Utility
A S S E S S M E N T	2.9	X	X	X	X
	a.	X	X	X	X
	b.	X	X	X	X
	c.	X	X	X	X
	d.	X			
	2.10	X	X	X	X
	a.	X	X	X	X
	b.	X	X	X	X
	c.	X		X	X
	d.	X	X	X	

*There is an interrelationship among validity, reliability, bias/sensitivity, and utility. The notation in Table 1 for each criterion and consideration is intended to provide users of this Framework with a general notion of the relevance of each criterion and consideration to overall quality.

Table 2: Overview of ELP Standards and Assessment Criteria by Possible Sources of Evidence for How a State is Addressing Specific Criteria and Related Considerations

	Criterion and Consideration	Possible Sources of Evidence*					
		State Administrative Code/Law/Policy	Manuals/Guides	Technical Reports and Research Studies	Meeting Reports/ Minutes	Training and Professional Development Materials	Parent Resources
S T A N D A R D S	1.1, 1.2, 1.3		X	X	X	X	X
	a.		X	X	X	X	
	b.		X	X	X	X	
	c.		X	X	X	X	X
	d.		X	X	X	X	X
	e.		X	X	X		
	f.		X	X	X		
	g.		X	X	X	X	
	1.4	X		X	X		
	a.	X		X	X		
	b.	X		X	X		
	c.	X		X	X		
	1.5	X	X	X	X	X	X
	a.	X			X	X	
	b.		X	X	X	X	
	c.		X	X		X	X
	d.	X	X			X	
	1.6	X		X	X		
	a.	X			X		
	b.	X		X	X		

* There likely will be variation in state documentation. The notation in Table 2 for each criterion and consideration is intended to provide Framework users with a general notion of the relevant documentation that may contain evidence/information related to a given criterion and consideration.

Table 2: Overview of ELP Standards and Assessment Criteria by Possible Sources of Evidence for How a State is Addressing Specific Criteria and Related Considerations, continued

Criterion and Consideration	Possible Sources of Evidence*					
	State Administrative Code/Law/Policy	Manuals/Guides	Technical Reports and Research Studies	Meeting Reports/ Minutes	Training and Professional Development Materials	Parent Resources
2.1, 2.2		X	X		X	
a.		X	X		X	
b.		X	X		X	
2.3	X	X	X	X	X	X
a.	X		X	X		
b.			X	X		
c.	X	X	X	X	X	X
2.4	X	X	X	X	X	X
a.			X	X	X	
b.	X	X	X		X	X
2.5	X	X	X	X	X	
a.	X	X		X	X	
b.	X	X		X	X	
c.	X	X	X	X	X	
2.6, 2.7	X	X	X	X	X	X
a.	X	X	X			
b.		X			X	X
c.			X	X		
d.	X		X	X		
e.		X	X	X	X	X
f.		X		X	X	
g.	X		X	X		

(Left column spanning label: ASSESSMENT)

* There likely will be variation in state documentation. The notation in Table 2 for each criterion and consideration is intended to provide Framework users with a general notion of the relevant documentation that may contain evidence/information related to a given criterion and consideration.

Table 2: Overview of ELP Standards and Assessment Criteria by Possible Sources of Evidence for How a State is Addressing Specific Criteria and Related Considerations, continued

	Criterion and Consideration	Possible Sources of Evidence*					
		State Administrative Code/Law/Policy	Manuals/Guides	Technical Reports and Research Studies	Meeting Reports/ Minutes	Training and Professional Development Materials	Parent Resources
A S S E S S M E N T	2.8	X	X	X	X	X	X
	a.			X	X		
	b.		X	X	X	X	
	c.			X	X		
	d.	X	X	X		X	
	e.	X	X	X	X	X	X
	f.		X			X	X
	2.9	X		X	X		
	a.	X		X	X		
	b.			X	X		
	c.			X	X		
	d.	X		X	X		
	2.10	X		X	X		
	a.	X			X		
	b.	X		X	X		
	c.	X		X	X		
	d.	X		X	X		

* There likely will be variation in state documentation. The notation in Table 2 for each criterion and consideration is intended to provide Framework users with a general notion of the relevant documentation that may contain evidence/information related to a given criterion and consideration.

Appendix B

**Background —
Development of the Framework**

Public Meeting and Comment

On June 6, 2007, the U.S. Department of Education (ED) published a Federal Register notice announcing a series of public meetings and a comment period during which stakeholders were invited to submit recommendations regarding the content of the *Framework for High-Quality ELP Standards and Assessments*.

Specifically, ED invited comment on the following four questions:

1. What are the critical elements that States should examine to ensure that their ELP standards promote effective instruction to raise LEP students' level of English proficiency? (Section 3113(b)(2))

2. What are the critical elements that States should examine to ensure that their ELP assessments provide a valid and reliable assessment of English language proficiency? (Section 3122(a)(3)(A)(ii))

3. What are the critical elements that States should examine to ensure that their ELP standards are aligned with their ELP assessments? (Sections 3113(b)(2) and (3)(D) and 3122(a)(3)(A)(ii))

4. What are strategies that States can use to ensure that their ELP standards are aligned with the achievement of challenging State academic content standards and student academic achievement standards they have adopted under Title I? (Section 3113(b)(2))

In addition to inviting comment, ED's public meetings featured roundtable discussions with experts in the fields of standards and assessments, English language acquisition, applied linguistics, and language research. A list of the roundtable participants, as well as transcripts of the roundtable discussions and public meetings, is available at http://www.ed.gov/about/inits/ed/lep-partnership/.

Solicited Feedback

The first draft of the Framework was shared with experts and practitioners, as well as with states, for review and comment in August and September 2007. For the period of October 2007 through January 2008, the Assessment and Accountability Comprehensive Center (AACC), working with ED, received valuable feedback and recommendations not only from technical experts in standards, assessment, psychometrics, English language acquisition, applied linguistics, and language research, but also from parents, teachers, administrators, researchers, and others with experience and expertise in Title I and Title III programs. A number of regional centers in the Comprehensive Center system convened their states for an in-depth, facilitated review and discussion of the draft Framework. Feedback was submitted directly to ED and the AACC (ED: LEP.Partnership@ed.gov; AACC: lepframework@ aacompcenter.org and esato@wested.org).

Pilot

In February—June 2008, states were asked to volunteer to undertake a self-review of their state ELP standards and assessments using the draft Framework. The purpose of this self-review was to serve as a means of improving the Framework as a tool for states, not to judge the quality of the states' standards and assessments as part of any formal review. Six states participated in the pilot, and each was assigned a consultant to assist it in the pilot process. Dr. Jack Levy and The National Clearinghouse for English Language Acquisition coordinated this effort. The states and their consultants are listed below:

- Colorado: Dr. Edynn Sato and Dr. Stanley Rabinowitz
- Indiana: Dr. Ellen Forte
- Michigan: Dr. Gary Cook
- Tennessee: Dr. Ed Roeber
- Washington: Dr. Frances Butler
- Wisconsin: Dr. Charlene Rivera

Findings from the pilot were used to revise the Framework document.

Notes

1 The Elementary and Secondary Education Act (ESEA) refers to the Framework's target student population as "limited English proficient" (LEP). The authors recognize that many researchers and practitioners prefer the term "English language learner" or "English learner (EL) student". Consistent with the more common usages, the Framework uses the latter terms. According to Title IX, Section 9101, EL students are (a) 3 to 21 years of age, (b) enrolled or preparing to enroll in elementary or secondary school, (c) either not born in the United States or have a native language other than English, and (d) owing to difficulty in speaking, reading, writing, or understanding English, not able to meet the state's proficient level of achievement to successfully achieve in English-only classrooms or not able to participate fully in society.

2 Rabinowitz, Roeber, Schroeder, & Sheinker, 2006; WestEd, 2007.

3 Careful documentation also serves the added purpose of creating a record that can be used to inform and bring up to speed any new players who might join the process over time, due to staff turnover, for example.

4 For example, Abedi, 2008; American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Cummins, 1991; Gersten, Baker, Shanahan, Linan-Thompson, Collins, & Scarcella, 2007; Gottlieb, 2006; Green, 1998; Kane, 1992; Kopriva, 2008; Messick, 1989; Sato, 2008; Teachers of English to Speakers of Other Languages (TESOL), 2006; Walqui, 2000; Wolf, Kao, Griffin, Herman, Bachman, Chang, & Farnsworth, 2008; Wolf, Kao, Herman, Bachman, Bailey, Bachman, Farnsworth, & Chang, 2008.

5 **Access** refers to the minimalization or removal of conditions (i.e., sources of construct-irrelevant variance, such as aspects of presentation/format of test information, aspects of response requirements, sociocultural contexts or references that may disadvantage certain students, etc.) that may interfere with students' ability to meaningfully engage with content or demonstrate their content knowledge and skills (construct-relevant information). Appropriate access does *not* significantly change the targeted construct. Strategies that facilitate access are tailored to the particular needs of students (e.g., cognitive, linguistic, physical). When access is constrained, it can result in the measurement of sources of variance that are not related to the intended test constructs (i.e., construct irrelevance). Limited access can allow construct-irrelevant abilities to interfere with that student's ability to fully demonstrate what he or she knows and can do; consequently, test results underestimate that student's construct-relevant achievement. Inappropriate access can affect the construct such that the curriculum or assessment no longer sufficiently represents the targeted domain (i.e., under-representation) (Messick, 1989; Sato, Rabinowitz, & Gallagher, forthcoming).

6 The four language modalities of listening, speaking, reading, and writing are often also referred to in the field as language "domains." In this Framework, "domain" will be used to refer to the domain of "English language proficiency" and "modalities" will be used to refer to "listening, speaking, reading, and writing."

7 The criteria presented in the Framework are important for states to consider but may not be fully sufficient for ensuring the quality of state ELP standards or full compliance with federal legislation.

8 The important point is not that EL students are diverse, or heterogeneous, but, rather, it is *how* they are diverse. For example, EL students differ among themselves in their level of literacy and amount of formal education in their native language. Some EL students are newly arrived; however, most are second or third generation in the U.S. EL students speak and write languages that may have sound systems, grammatical structures, and writing systems that are different to varying degrees from those of English, linguistic differences that can affect the relative rate and ease of students' English language acquisition and their development of literacy in English. The nature of the diversity of this group may affect the implementation of standards (e.g., curriculum, instruction) and assessments.

9 Abedi, 2004; Abedi & Dietel, 2004; Kopriva, 2000; Solano-Flores & Trumbull, 2003.

10 Canale & Swain, 1980; Rabinowitz, Roeber, Schroeder, & Sheinker, 2006.

11 While there may not be just one accepted theory of English language development or acquisition (as there is not just one accepted theory for teaching mathematics, for example), much information relevant to English language development and acquisition has been validated by experts across several fields, and, therefore, needs to be considered when developing statewide systems of standards and assessments.

12 National Clearinghouse for English Language Acquisition, 2006.

13 **Academic language,** broadly defined, includes the language students need to meaningfully engage with academic content within the academic context. This should *not* be interpreted to suggest that separate word lists and/or definitions of content-related language should be developed for each academic subject. Rather, academic language includes the words, grammatical structures, and discourse markers needed in, for example, describing, sequencing, summarizing, and evaluating — these are language demands (skills, knowledge) that facilitate student access to and engagement with grade-level academic content. These academic language demands are different from cognitive demands (e.g., per Bloom's taxonomy). Although there may not be just one accepted definition of academic language, there are a good number of resources available that address the issue of academic language and may be considered in the development of state ELP standards and assessments. For example: Aguirre-Munoz, Parks, Benner, Amabisca, & Boscardin, 2006; Bailey, 2007; Bailey, Butler, & Sato, 2007; Butler, Bailey, Stevens, Huang, & Lord, 2004; Chamot & O'Malley, 1994; Cummins, 1980; Cummins, 2005; Halliday, 1994; Sato, 2007; Scarcella & Zimmerman, 1998; Schleppegrell, 2001.

14 Alignment of ELP standards to ELP assessments and to language proficiency levels is addressed in Section 2, the Assessment section.

15 According to NCLB Title III, Section 3113, the state ELP standards are to be "aligned" with the achievement of challenging state academic content and student academic achievement standards. As commonly used, **alignment** refers to relationships that tend to be *direct*; alignment models typically examine correspondence between standards and assessments for a single student population (e.g., general education, English language learners, students with disabilities) or for a single content area (e.g., English language arts, mathematics). **Linkage** refers to relationships that tend to be *developmental, foundational*, or *proximal*; thus, "linkage" is seen as the more appropriate term to use to describe the intent of NCLB Title III, Section 3113. The term "linkage" not only allows for correspondence of content area topics (e.g., figurative language, measurement, scientific inquiry), but also for the range of language knowledge and skills that students need to meaningfully engage with and achieve academic content (e.g., the words, grammatical structures, and discourse markers needed in describing, summarizing, inquiring, and analyzing). Models examining linkage often lend themselves to correspondence between standards and/or assessments developed for different student populations or different content areas (Bailey, Butler, & Sato, 2007; Cook, 2005; Flowers, Wakeman, Browder, & Karvonen, 2007; Sato, Lagunoff, Worth, Bailey, & Butler, 2005; Webb, Horton, & O'Neal, 2002; WestEd, 2004).

16 The distinction being made between *language* and *content* is that ELP language skills and knowledge are necessary but not sufficient to demonstrate proficiency in the academic content areas (e.g., ELA, mathematics, science, social studies). However, there is obviously more overlap (convergence) between the language and content of ELP and ELA/Reading than between ELP and other academic content standards because of the focus on English language in both ELP and ELA.

17 Rabinowitz, & Sato, 2005, 2006.

18 The criteria presented in the Framework are important for states to consider but may not be fully sufficient for ensuring the technical adequacy of state ELP assessments or full compliance with federal legislation.

19 The point is not *that* EL students are diverse/heterogeneous; rather it is *how* (i.e., the ways in which) they are diverse/heterogeneous. For example, EL students differ among themselves in their level of literacy and amount of formal education in their native language. Some EL students are newly arrived; however, most are second or third generation in the U.S. Some EL students speak and write languages that may have sound systems, grammatical structures, and writing systems that are different from those of English to varying degrees, which can impact the rate and ease of English language acquisition and English literacy learning. The nature of the diversity/heterogeneity of this group may affect the implementation of standards (e.g., curriculum, instruction) and assessments.

20 Abedi, 2004; Abedi & Dietel, 2004; Kopriva, 2000; Solano-Flores & Trumbull, 2003.

21 Marion & Pellegrino, 2006; Rabinowitz & Sato, 2005; Sireci, Li, & Scarpati, 2002; Solorzano, 2008.

22 For example: Gong & Marion, 2006; Marion & Pellegrino, 2006; Thompson, Quenomoen, Thurlow, & Ysseldyke, 2001.

23 Kopriva, 2000.

24 Rabinowitz, 2007.

25 Appropriate **access** for EL students on ELP assessments means that the assessment minimizes the effects of or contains no construct-irrelevant factors that could interfere with EL students' demonstration of English language proficiency across the language modalities (i.e., listening, speaking, reading, writing).

26 Appropriate **accommodations** for EL students on ELP assessments are testing condition changes (e.g., in presentation, setting, scheduling/timing) that are consistent with the requirements of a student's Individualized Education Program (IEP) and that are implemented to increase the student's access to test content without significant alteration to the assessed construct. Such changes are deemed fair and reasonable when standardized administration conditions do not provide an equal opportunity for all students to demonstrate what they know and can do. An accommodation is intended to minimize or remove the effects of construct-irrelevant factors on test performance. It is assumed that with or without the accommodation, the same construct is being assessed (Abedi & Lord, 2001; Butler & Stevens, 2001; Holmes & Duron, 2000; Rivera & Stansfield, 2001).

27 Liu, Anderson, Swierzbin, & Thurlow, 1999; Rabinowitz & Sato, 2005; Rivera & Column, 2004; Solano-Flores & Li, 2006; Solorzano, 2008.

28 **Proficiency level descriptors** are explicit definitions of what students must do (descriptions of knowledge and skill expectations) to demonstrate competence on academic content or ELP standards. **ELP standards** are statements that define — for students, teachers, schools, and the community more broadly — what students should know and be able to do and that guide the opportunities students have to learn and the resources students are provided to achieve particular knowledge and skills.

29 Title III requires that each state establish three Annual Measurable Achievement Objectives (AMAOs):

- AMAO 1: The number or percentage of EL students making progress toward English language proficiency (one level per year) until reaching proficiency.

- AMAO 2: The annual increase in the number or percentage of students attaining English language proficiency.

- AMAO 3: As a subgroup (per Title I), EL students' adequate yearly progress (AYP) toward meeting grade-level academic achievement standards in English language arts and mathematics.

30 For example: Berk, 1986; Cizek, 2001; Hambleton, 2001; Kingston, Kahl, Sweeney, & Bay, 2001; Lewis, Mitzel, & Green, 1996; Livingston & Zeikey, 1982.

31 For example: Hambleton, Swaminathan, & Rogers, 1991; Holland & Dorans, 2006; Kolen & Brennan, 2004; Lissitz & Huynh, 2003; Mislevy, 1992.

References

Abedi, J. (2004). The No Child Left Behind Act and English language learners: Assessment and accountability issues. *Educational Researcher*, *33*(1), 4—14.

Abedi, J. (2008). Linguistic Modification: Part I — Language Factors in the Assessment of English Language Learners: The Theory and Principles Underlying the Linguistic Modification Approach. Washington, DC: LEP Partnership.

Abedi, J., & Dietel, R. (2004). Challenges in the No Child Left Behind Act for English language learners. *Phi Delta Kappan, 85*, 782—785.

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*(3), 219—234.

Aguirre-Munoz, Z., Parks, J. E., Benner, A., Amabisca, A., & Boscardin, C. K. (2006). *Consequences and Validity of Performance Assessment for English Language Learners: Conceptualizing & Developing Teachers' Expertise in Academic Language*. CRESST/UCLA Center for the Study of Evaluation. Available at http://www.cse.ucla.edu/products/reports/ R700.pdf.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.

Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Bailey, A., Butler, F., & Sato, E. (2007). Standards-to-standards linkage under Title III: Exploring common language demands in ELD and science standards. *Applied Measurement in Education, 20(1), 53—78.*

Berk, R. A. (1986). A consumer's guide to setting performance standards on criterion-referenced tests. *Review of Educational Research, 56*(1), 137—172.

Butler, F. A., Bailey, A. L., Stevens, R., Huang, B., & Lord, C. (2004). *Academic English in fifth-grade mathematics, science, and social studies textbooks.* (Final Deliverable to IES, Contract No. R305B960002; currently available as CSE Rep. No. 642.) Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Butler, F. A., & Stevens, R. (2001). Standardized assessment of the content knowledge of English Language Learners K—12: Current trends and old dilemmas. *Language Testing*, *18*(4), 409—427.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1—47.

Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. New York, NY: Longman.

Cook, H. G. (2005). Aligning English language proficiency tests to English language learning standards. Paper prepared for Assessing Limited English Proficiency Students State Collaborative on Assessment and Student Standards. Washington, DC: Council of Chief State School Officers.

Cummins, J. (1980). The construct of proficiency in bilingual education. In J. E. Alatis (Ed.), *Georgetown University Round Table on Languages and Linguistics: Current issues in bilingual education* (pp. 81—103). Washington, DC: Georgetown University.

Cummins, J. (1991). Empowering minority students: A framework for intervention. In Minami, M., and Kennedy, B. (Eds.), *Language Issues in Literacy and Bilingual/Multicultural Education*, Series 22. Boston: Harvard Education Review.

Cummins, J. (2005). Language proficiency, bilingualism, and academic achievement. In Richard-Amato, P. A., and Snow, M.A. (Eds.), *Academic success for English language learners* (pp. 76—86). White Plains, NY: Longman.

Cizek, G. (2001). *Setting performance standards: Concepts, methods, and perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Flowers, C., Wakeman, S. Y., Browder, D. M., & Karvonen, M.(2007). *Links for academic learning: An alignment protocol for alternate assessments based on alternate achievement standards.* Charlotte, NC: National Alternate Assessment Center, University of North Carolina at Charlotte.

Gersten, R., Baker, S. K., Shanahan, T., Linan-Thompson, S., Collins, P., & Scarcella, R. (2007). *Effective literacy and English language instruction for English learners in the elementary grades: A practice guide* (NCEE 2007-4011). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved March 11, 2007, from http://ies.ed.gov/ncee.

Gong, B., & Marion, S. F. (2006). *Dealing with flexibility in assessments for students with significant cognitive disabilities* (Synthesis Report 60). Minneapolis, MN: University of Minnesota, National Center for Educational Outcomes.

Gottlieb, M. (2006). *Assessing English language learners: Bridges from language proficiency to academic achievement.* Thousand Oaks, CA: Corwin Press.

Green, D. R. (1998). Consequential aspects of the validity of achievement tests: A publisher's point of view. *Educational Measurement: Issues and Practices, 17*(2), 16–19.

Halliday, M. A. K. (1994). *An introduction to functional grammar (2nd ed.).* London: Edward Arnold.

Hambleton, R. K. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. Cizek (ed.), *Setting performance standards: Concepts, methods and perspectives* (pp. 89–116). Mahwah, NJ: Lawrence Erlbaum.

Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (ed.), Educational measurement (4th ed.). Westport, CT: American Council on Education and Praeger Publishers.

Holmes, D., & Duron, S. (2000). *LEP students and high-stakes assessment.* Washington, DC: National Clearinghouse for Bilingual Education, U.S. Department of Education.

Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin, 112,* 527–535.

Kingston, N., Kahl, S. R., Sweeney, K., & Bay, L. (2001). Setting performance standards using the body of work method. In G. J. Cizek (ed.), *Setting performance standards: Concepts, methods, and perspectives.* Mahwah, NJ: Lawrence Erlbaum.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: methods and practices (2nd ed.).* New York: Springer.

Kopriva, R. J. (2000). *Ensuring accuracy in testing for English language learners.* Washington, DC: Council of Chief State School Officers.

Kopriva, R. J. (2008). *Improving testing for English language learners.* New York: Routledge.

Lewis, D. M., Mitzel, H. C., & Green, D. R. (1996, June) *Standard setting: A bookmark approach.* Paper presented at the Council of Chief State School Officers Large-Scale Assessment Conference, Colorado Springs, CO.

Liu, K., Anderson, M. E., Swierzbin, B., & Thurlow, M. (1999). *Bilingual accommodations for limited English proficient students on statewide reading tests: Phase 1* (Minnesota Report 20). Minneapolis: University of Minnesota, National Center on Educational Outcomes.

Lissitz, R. W., & Huynh, H. (2003). Vertical equating for state assessments: issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research & Evalatuion, 8*(10). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=10.

Livingston, S. A., & Zeikey, M. J. (1982). *Passing scores: A manual for setting standards of performance on educational and occupational tests.* Princeton, NJ: Educational Testing Service.

Marion, S. F., & Pellegrino, J. W. (2006). A validity framework for evaluating the technical quality of alternate assessments. *Educational Measurement: Issues and Practice, 25*(4), 47–57.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: Macmillan.

Mislevy, R. J. (1992). *Linking educational assessments: Concepts, issues, methods, and prospects.* Princeton, NJ: Educational Testing Service.

National Clearinghouse for English Language Acquisition. (2006). *How many school-aged English language learners (ELLs) are there in the U.S.?* (NCELA FAQ No. 1). Washington, DC: Author. Retrieved October 22, 2007, from http://www.ncela.gwu.edu/expert/faq/01leps.html.

Rabinowitz, S. (2007). *Assessment of English Language Learners Under Title I and Title III: How One Testing Program Can Inform the Other.* Washington, DC: LEP Partnership.

Rabinowitz, S., Roeber, E., Schroeder, C., & Sheinker, J. (with CAS SCASS Study Group: Transitions in Assessment from IASA to NCLB). (2006). *Creating aligned standards and assessment systems.* Washington, DC: Council of Chief State School Officers.

Rabinowitz, S., & Sato, E. (2005). *A technical review of high-stakes assessments for English language learners.* San Francisco: WestEd.

Rabinowitz, S., & Sato, E. (2006). *The technical adequacy of assessment for alternate student populations: Guidelines for consumers and developers.* San Francisco: WestEd.

Rivera, C. & Collum, E. (2004). *An analysis of state assessment policies addressing the accommodation of English Language Learners.* Issue paper commissioned for the National Assessment Governing Board Conference on Increasing the Participation of SD and LEP Students in NAEP. Arlington, VA: George Washington University.

Rivera, C., & Stansfield, C. W. (2001). *The effects of linguistic simplification of science test items on performance of Limited English Proficient and monolingual English-speaking students.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Sato, E. (2007). *Rethinking alignment for English language learner assessments and standards: Issues and implications for extending current models.* Paper commissioned by the Council of Chief State School Officers.

Sato, E. (2008). *Linguistic Modification: Part II — A Guide to Linguistic Modification: Increasing English Language Learner Access to Academic Content.* Washington, DC: LEP Partnership.

Sato, E., Lagunoff, R., Worth, P., Bailey, A. L., & Butler, F. A. (2005). *ELD Standards Linkage and Test Alignment Under Title III: A Pilot Study of the CELDT and the California ELD and Content Standards.* (Final Report to the California Department of Education). San Francisco: WestEd.

Sato, E., Rabinowitz, S., & Gallagher, C. (forthcoming). *Access and Special Student Populations — the Similarities/Differences in the Needs of English Language Learners and Students with Disabilities: Implications for Standards, Assessment, and Instruction* [working title]. San Francisco: WestEd.

Scarcella, R., & Zimmerman, C. (1998). Academic words and gender: ESL student performance on a test of academic lexicon. *Studies in Second Language Acquisition*, *20*, 27—49.

Schleppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education, 12*(4), 431—459.

Sireci, S. G., Li, S., & Scarpati, S. (2002). *The effects of test accommodations on test performance: A review of the literature.* CEA Research Report 485. Amherst, MA: School of Education, University of Massachusetts.

Solano-Flores, G., & Li, M. (2006). The use of generalizability (g) theory in the testing of linguistic minorities. *Educational Measurement, Issues and Practice, 25*, 13—22.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English-language learners. *Educational Researcher, 32*(2), 3—13.

Solorzano, R. W. (2008). High-stakes testing: Issues, implications, and remedies for English language learners. *Review of Educational Research, 78*(2), 260—329.

Teachers of English to Speakers of Other Languages. (2006). P*reK—12 English language proficiency standards.* Alexandria, VA: Author.

Thompson, S., Quenemoen, R., Thurlow, M., & Ysseldyke, J. (2001). *Alternate assessments for students with disabilities.* Thousand Oaks, CA: Corwin Press.

Walqui, A. (2000). *Access and engagement: Program design and instructional approaches for immigrant students in secondary school.* McHenry, IL: Delta Publishing.
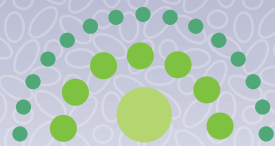
Webb, N. L., Horton, M., & O'Neal, S. (2002). *An analysis of the alignment between language arts standards and assessments for four states.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

WestEd. (2004). *Alignment Study: Louisiana Content Standards and the Louisiana Alternate Assessment (LAA)*. San Francisco, CA: Author.

WestEd. (2007). *Responses to critical questions: Oregon standards and assessments evaluation report*. San Francisco, CA: Author.

Wolf, M. K., Kao, J., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S. M., & Farnsworth, T. (2008). *Issues in Assessing English Language Learners: English Language Proficiency Measures and Accommodation Uses — Practice Review*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Wolf, M. K., Kao, J., Herman, J. L., Bachman, L. F., Bailey, A. L., Bachman, P. L., Farnsworth, T., & Chang, S. M. (2008). *Issues in Assessing English Language Learners: English Language Proficiency Measures and Accommodation Uses — Literature Review*. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).