CCSSO

Council of Chief State School Officers

Toward a "Common Definition of English Learner":

Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options

The Council of Chief State School Officers is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

Authors:

Robert Linquanti, WestEd

H. Gary Cook, Wisconsin Center for Education Research

Acknowledgments

CCSSO

One Massachusetts Avenue, NW, Suite 700 Washington, DC 20001-1431
Phone (202) 336-7000 Fax (202) 408-8072
www.ccsso.org

Table of Contents

Overview

The U.S. Department of Education (USED) requires states participating in any of four federally-funded assessment consortia to establish a "common definition of English Learner." This includes the two Race to the Top assessment (RTTA) consortia (Smarter Balanced Assessment Consortium [Smarter Balanced] and Partnership for Assessment of Readiness for College and Careers [PARCC]), as well as the two Enhanced Assessment Grant (EAG) English language proficiency assessment consortia (WIDA's Assessment Services Supporting English Learners through Technology Systems [ASSETS] and CCSSO's English Language Proficiency Assessment for the 21st Century [ELPA21]). Specifically, each consortium "must define the term in a manner that is uniform across member states and consistent with section 9101 (25)[1] of the ESEA" (USED, 2010, p. 20).

Although the two federally-funded consortia developing alternate assessments based on alternate achievement standards (AA-AAS)[2] are *not* required to develop a common definition of English learner (EL), their member states largely overlap with the assessment consortia mentioned above, and they will include eligible ELs in these assessments. Having a common EL definition for these consortia that agrees with the definition adopted by the other consortia is therefore clearly desirable.[3]

Fulfilling the requirement for a common EL definition among consortia member states is neither simple nor straightforward. In fact, this requirement presents substantial challenges that call for a carefully coordinated, multiyear effort within and across consortia member states. Such an effort will need to proceed in stages and encompass several critical decisions that need to be informed by student performance outcomes on the assessments that are to be developed. The purpose of this paper is to define key issues involved, and provide guidance that consortium member states can use to move toward establishing a common EL definition in ways that are theoretically-sound, evidence-based, empirically-informed, pragmatic, and sensitive to the many policy, technical, and legal issues.

The Council of Chief State School Officers (CCSSO), under the sponsorship of the Carnegie Corporation of New York and in conjunction with Understanding Language of Stanford University's Graduate School of Education and the WIDA Consortium, have convened a cross-consortium English Language Learner Assessment Advisory Committee, which consists of technical staff and leadership from the above-mentioned federally funded consortia, EL researchers and technical assistance experts, and policy advisers and other stakeholders. Two committee members were commissioned to produce this report and an accompanying policy brief (Linquanti & Cook, 2013) as starting points for a wider, informed discussion of the issues and opportunities afforded by addressing the common EL definition requirement.

This paper briefly outlines key issues in defining English learners and discusses specific policy and technical options by using a four-stage framework that encapsulates the key criteria and process of

---

1   Described below

2   Dynamic Learning Maps Alternate Assessment System Consortium (DLM) and the **National Center and State Collaborative Partnership (NCSC)**

3   Doing so will be complicated as communication issues are inherent in many of the disabilities of students with significant cognitive disabilities. Data from 18 states (Towles-Reeves et al., 2012) indicate that approximately 13% (range between 3% to 36%) of students with significant cognitive disabilities are ELs.

defining ELs: 1) identify a student as a *potential* English learner; 2) classify (confirm/disconfirm) a student as an English learner; 3) establish an "English-language proficient" performance standard on the state/ consortium English language proficiency (ELP) test against which to assess ELs' English-language proficiency; and 4) reclassify a student to former-EL status through the use of multiple exit criteria. This framework can help consortium member states systematically review and document existing practices, and discuss ways to align instruments, policies, and practices within a given state; across states within a given consortium; and across permutations of consortia in which multi-consortium states may be configured (e.g., academic and ELP assessment consortia).

Setting the context: What defines an English learner?

The current federal definition of English learner has legal and policy roots which merit a brief review. The educational rights of "national origin-minority children" are well-established in Title VI of the Civil Rights Act of 1964 (Public Law 88-352). Specifically, Section 601 declares

> "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving Federal financial assistance."   (42 USC Sec.2000d.)

In particular, these children's right to equitable educational opportunities including, when deemed necessary, effective English language development services is supported by the Equal Educational Opportunities Act of 1974 (Public Law 93-380), which requires states to ensure an education agency "take[s] appropriate action to overcome language barriers that impede equal participation by its students in its instructional programs" (20 USC Sec.1703(f)).

These educational rights have been upheld by the U.S. Supreme Court's 1974 *Lau v. Nichols* ruling and subsequent federal civil rights case law (e.g., *Castañeda v. Pickard*, 1981). In particular, the *Lau* decision (interpreting Title VI of the Civil Rights Act) affirmed that students unable to benefit from English-medium instruction were effectively foreclosed from "a meaningful opportunity to participate in the educational program" and were entitled to receive English language development instruction in order to access academic content instruction. The *Castañeda* ruling (interpreting the Equal Educational Opportunities Act) elaborated on determining "the appropriateness of a particular school system's language remediation program" noting the need for EL educational services to be based on a sound theory, meaningfully implemented, and evaluated to determine their effectiveness, with the requirement to modify the services as needed to ensure effectiveness. The *Castañeda* ruling also identified the need for educators to support EL students' linguistic and academic development so they acquire English language proficiency, do not incur "irreparable academic deficits" while doing so, and "attain parity of participation with other students" within a reasonable length of time.

In legislation and case law, there is clear recognition of the relationship between EL students' English language proficiency and their ability both to benefit from instruction in English, and to demonstrate

knowledge, skills, and abilities in academic content assessed using English.[4] There is also an implication that English learner status is meant to be temporary, and that students are expected to leave the EL category as a result of quality instructional services they are entitled to receive.

These federal sources affirm the educational rights of national origin-minority or language-minority students, yet they do *not* offer a specific definition of what constitutes "limited English proficiency," nor how to determine when it has been adequately addressed. It is the Elementary and Secondary Education Act (ESEA) – first in its 1978 reauthorization, and further refined in 1994 (IASA) and 2001 (NCLB) – which provides an explicit definition of what constitutes a "Limited English Proficient" student.[5] As defined in section 9101(25) of the law (italics and bold emphases have been added)

> The term limited English proficient, when used with respect to an individual, means an individual — (A) who is aged 3 through 21; (B) who is enrolled or preparing to enroll in an elementary school or secondary school; (C)(i) who was not born in the United States or *whose native language is a language other than English*; (ii)(I) who is a Native American or Alaska Native, or a native resident of the outlying areas; and (II) *who comes from an environment where a language other than English has had a significant impact on the individual's level of English language proficiency*; or (iii) who is migratory, *whose native language is a language other than English, and who comes from an environment where a language other than English is dominant*; and (D) **whose difficulties in speaking, reading, writing, or understanding the English language** *may* **be sufficient to deny the individual** — (i) **the ability to meet the State's proficient level of achievement on State assessments described in section 1111(b)(3)**; (ii) **the ability to successfully achieve in classrooms where the language of instruction is English**; or (iii) **the opportunity to participate fully in society**.

The emphases focus on the linguistic and academic dimensions of the current federal EL definition. Specifically, the italicized text signals students with a native language other than or in addition to English, or who come from an environment where a non-English language either is dominant or has affected a student's English-language proficiency. The bold sections posit that difficulties in these students' facility in and use of English *may* deny them the ability to perform proficiently on academic content assessments mandated under ESEA Title I; to achieve in English-medium classrooms; or to participate fully in society.

Based on this definition, determining what constitutes sufficient English language proficiency involves three tasks: 1) examining the relationships between EL students' assessed English language proficiency and their content assessment performance in order to determine an English-language proficient performance standard; 2) understanding and judging English-proficient EL students' linguistic capacities to engage

---

4   Strong research evidence (e.g., Slavin & Cheung, 2005; Francis, Lesaux, & August, 2006) supports use of bilingual instructional methods to facilitate access to early literacy development, as well as academic content instruction, while EL students develop English-language proficiency. Substantial research evidence also highlights equivalent or greater long term attainment of academic achievement results in English using well-implemented bilingual instructional programs. Several states have also instituted "seals of biliteracy" to signal their valuing of students' ability to perform academically in more than one language. Such bilingual methods are acknowledged as legitimate educational options, but are not evaluated per se in these legal cases.

5   The term "limited English proficient" (LEP) has been criticized for conveying a deficit view of students' non-English language abilities. The term English learner (or English language learner) has largely replaced LEP at federal, state, and local levels, but the older term remains operational in ESEA.

in English-medium classroom-based interactions, tasks, and activities; and 3) establishing and judging English-proficient EL students' linguistic capacities for wider social and occupational opportunities.

In summary, in the K-12 school context, current law and regulation call for state and local education systems to 1) identify national origin- or language-minority children whose current language use or language environment includes a language other than English; 2) determine which of these students needs additional specialized linguistic and academic support services; 3) provide appropriate linguistic and academic support services to ensure English-language skills are sufficiently developed such that these students are not denied the ability to learn, demonstrate learning, achieve in classrooms, and participate fully in society, using English; and 4) ensure that these services are provided in a timely and effective way.

We now turn to key issues and opportunities that states in the four federally funded assessment consortia face in moving toward and adopting a common definition of English learner. This includes a brief discussion of state consortium participation, and then proceeds to a review of issues and suggested guidance that consortium member states may consider to address options using the four-stage framework that encompasses the EL definitional process and criteria.

State consortium participation

As illustrated in Figure 1 below, there are different permutations of state assessment consortia participation. Rows represent participation in academic assessment consortia, while columns illustrate participation in ELP assessment consortia. Both the bottom row and the far right column (designated as "Stand-Alone") contain states not participating in an academic or ELP consortium, respectively, and the bottom right cell contains those states currently not participating in any consortium.

Figure 1. Permutations of Cross-Consortium Participation as of June 21, 2013

Academic/ELP	ASSETS	ELPA21	Stand-Alone
Smarter Balanced	DE, ID, ME, MI, MO, MT, NV, NH, NC, ND*, PA*, SC+, SD, USVI, VT, WI, WY	IA, KS, OR, SC+, WA, WV	AK, CA, CT, HI
PARCC	CO, DC, IL, KY, MA, MD, MS, NJ, NM, ND*, OK, PA*, RI, TN	AR, FL, LA, OH	AZ, GA, IN, NY
Stand-Alone	AL, MN, UT, VA	NE	TX

*Currently advisory states in Smarter Balanced and PARCC
+ Listed as ASSETS and ELPA21

Sources: Smarter Balanced; Achieve, Inc.; WIDA; and CCSSO

States in *any* of the four consortia must address the "uniform manner" EL definition requirement *within* and — where applicable — *across* their respective consortia. This implies, for example, that Smarter Balanced states either in different ELP assessment consortia – or with their own, stand-alone ELP assessment – will need

to move toward aligning their definitions of English learner. Similarly, ASSETS states across different academic assessment consortia – as well as those ASSETS states using stand-alone academic assessments – will also need to align their EL definitions. Conceivably, even current stand-alone states not participating in any consortium will, for reasons of cross-state comparability, be interested in utilizing some of the data analytic tools and participating in consortium discussions related to moving toward a common EL definition.

Clearly, states need to define a roadmap of processes and a timeframe for decision points so that efforts can be coordinated within states, as well as within and across consortia. The remainder of this document presents a framework for a common EL definition to identify key issues at each stage of the EL definition process; guidance on options for state and consortium action; and a proposed timeline and high-level workplan for addressing issues and options.

An organizing framework for a common English learner definition

In order to define and present the key policy and process tasks that might be undertaken by states in consortia, we present a framework for a common EL definition (see Figure 2, below) that will be used to organize our discussion. This framework is adapted from a recent National Research Council (NRC) panel report (NRC, 2011) that explored comparability issues in state procedures for identifying, classifying, and reclassifying English learners related to federal funds allocation for state EL instructional programs. The figure illustrates key stages in the basic trajectory of how a student becomes an English learner, and how they are exited from that status at the appropriate time. As indicated, all states typically define their policies, instruments, and procedures around four key stages: 1) identify a student as a *potential* English learner; 2) classify (i.e., confirm/disconfirm) a student as an English learner; 3) establish an "English-language proficient" performance standard on the state/consortium ELP test against which to assess ELs' English language proficiency; and 4) reclassify a student to former-EL status through the use of multiple exit criteria.

Figure 2. Framework for a Common EL Definition (adapted from NRC, 2011, p. 78)
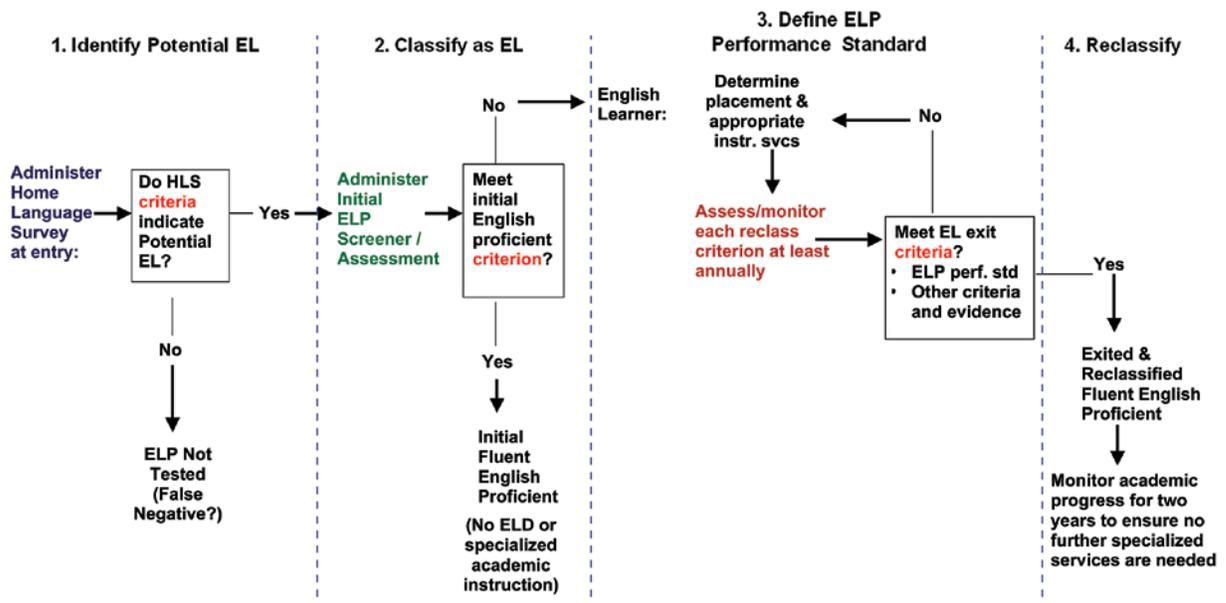
Consortium member states need to explicitly define their approaches to these four key steps, and establish mechanisms to move toward a sufficient degree of "convergence" or alignment over time. We now describe issues and opportunities for consortium member states to consider in each of these four steps.

1. Identifying Potential English Learners

In this first step, states identify from the overall K-12 student population the group of students that is *potentially* English learner. As noted in the discussion of the ESEA definition of an English learner above, this requires states to identify students with a native language other than English, or who come from an environment where a non-English language either is dominant or may have affected a student's English-language proficiency. It is this population that is *potentially* EL and will require assessment to determine if they are *in fact* EL and in need of specialized language and academic support services to which they are entitled. Note that this stage does *not* identify English learners *per se*, but only the pool of students who are potentially EL. Confirming (or disconfirming) EL status occurs at the next stage.

Issues

States currently use a variety of methods for identifying potential EL students. Home language surveys (HLS) are primarily used for this purpose in all but four states, but there is substantial variation in survey questions' phrasing, content, and application across states. As Bailey & Kelly (2012) note in their comprehensive review of HLS practices in the initial identification of potential ELs, this stage is crucially important as it is the first filter that determines the pool of prospective EL students. Home language surveys are usually administered to the adults enrolling a child as part of initial registration with the school district. Depending upon the number and nature of questions asked, surveys can "cast their net" widely or narrowly in defining the pool of students for further assessment. This has implications for the resources states must expend at this and subsequent stages. It also affects the potential for generating "false positives" (students wrongly identified as potential EL when they are in fact not), as well as "false negatives" (students not properly identified as EL because they are omitted from initial assessment).

As Bailey & Kelly also note, there are substantial validity concerns associated with current HLSs. These include the following:

> a) *The HLS's purpose may be unclear, ambiguous, or problematic, leading to inaccurate survey information.* Those administering and responding to the HLS may be confused by a lack of a clearly stated purpose. For example, Abedi (2008) found that adults may associate the survey with a determination of U.S. citizenship status, or with a predetermination of educational opportunities that may be perceived as unequal or undesirable. Factors such as these can compromise the accuracy of survey responses.

> b) *HLS questions may vary locally, and be irrelevant to the construct under examination.* Questions that ask about the language in which the student first began to speak, or that ask about the language used among adults in the home may have little relevance to the student's current language dominance or overall exposure to English. For example, an entering kindergartener may

have a non-English speaking grandparent at home, with whom one of the parents communicates in the primary language. Yet this same child may speak only in English with both parents, and have attended an all English-language preschool program. Conversely, monolingual-English-speaking parents may have adopted a non-English-speaking school-age child, and the parents' language use may have little or no bearing on the child's current English proficiency.

c) *Survey administration is inconsistent and decision-making procedures are not specified.* Instructions for delivering HLSs can be unclear, and may be implemented differently across sites within a given state or district. For example, some sites may translate questions provided only in English, or they may read questions to adults who are not literate in the language of the survey. However, survey questions can be difficult to translate, and if these are not consistently provided, they will likely lead to misinterpretation. In some states, making decisions on survey results are left to local educators, without guidance on how to weight question answers, or to resolve potentially ambiguous or contradictory answers.

Guidance

Drawing largely from Bailey & Kelly's (2012) review, we propose the following minimal guidance to states in consortia for strengthening the reliability, validity, and usefulness of HLSs in moving toward a common EL definition.

a. **Clearly state the purposes and intended uses of the HLS to those who will administer and those who will complete the survey.** The purposes and intended uses of the HLS should be made explicit to those administering and those completing the survey, and survey questions should reflect the stated purposes and uses. There may also be a need to explicitly state what the HLS is *not* intended to do (e.g., confirm citizenship status or predetermine educational services).

b. **Ask survey questions that accurately target the relevant constructs under consideration.** We recommend initially evaluating potential HLS questions as construct-essential (i.e., their *exclusion* will likely generate false positives or false negatives); construct-associated (they may increase accuracy but may also increase error), or construct-irrelevant (i.e., their *inclusion* will likely generate false positives or false negatives). Following Bailey and Kelly, we encourage states to include as construct-essential those questions that focus on the child's *current language dominance* and *degree of exposure to English*. Such questions might include the following:

  - *Which language does your child most frequently speak at home?*
  - *Which language do adults in your home most frequently use when speaking with your child?*
  - *Which language(s) does your child currently understand or speak?*

c. **Establish clear guidelines for administering the survey, and clear decision rules for interpreting and acting on the information.** Guidelines should include brief instructions for those administering surveys; rigorous translation of survey questions into multiple languages, as well as guidelines for oral administration of the survey by fluent/native speakers of the

family language in cases where adult literacy is an issue. Also, explicit decision rules should be established to determine the weighting of questions; to resolve conflicting or family-altered surveys; and to analyze survey results and either advance the student to the next step in the EL-identification process or conclude the student is not of potential-EL status.

d. **Conduct validation reviews of existing and proposed HLS forms and questions.** Collecting and analyzing data on the effectiveness of different HLS forms or questions in identifying the target population are needed to improve the efficacy of the overall process. This can include studies to examine the over- and under-identification of ELs, including "hit rates" based on false positives (students identified as potential EL who upon further assessment are found to be initially fluent English proficient[6]), and false-negatives (students not identified as potential EL who are subsequently found by teachers to need the specialized linguistic supports provided to English learners). Such studies are needed to refine and enhance this critical first step of the process.

Additional Options

Researchers (e.g., Bailey & Kelly, 2012; Abedi, 2008) suggest enhancing the HLS with additional questions, or even augmenting it with additional interview, student background, or observation protocols. While resource constraints may limit the feasibility of multiple measures at this stage, states and consortia may want to test the utility of questions focusing on the student's exposure to literacy practices at home, as well as to language and literacy practices outside the home (e.g. preschool and the wider community). While this extends beyond the home, it may yield more useful information related to the students' potential EL status.

By enacting the minimum guidance recommended above, member states within and across consortia can strengthen HLS instruments and procedures, and move toward standardizing and validating their home language surveys. With accumulated empirical evidence regarding validity and utility, there may be justification to move eventually to a single, commonly-used HLS instrument and set of procedures within a consortium, or even across consortia.

2. Establishing Initial English Learner Classification

As noted above, the HLS does not determine if students are English learners and should be afforded specialized language instructional services, nor does it identify levels of English language proficiency. The HLS only identifies whether a student is a potential EL. Once identified as such, states use a variety of measures to confirm (or disconfirm) EL status and establish initial EL classification. The intent of these assessments is to identify if a student should be classified as an EL and, if so, to determine the student's current ELP level. Following classification, ELs can then be appropriately placed into language instruction educational programs. Those found to have sufficient initial English proficiency to not require specialized support services are considered as initially fluent English proficient (or I-FEP) and are not classified as EL.

---

6   The term "initially fluent English proficient" is defined in the *Establishing Initial English Learner Classification* section below.

Issues

According to a recent report by the National Research Council (NRC, 2011), 27 states use a screener/placement test for EL classification.[7] Seventeen states allow school districts to select the language proficiency assessment used for initial classification, though they provide a list of tests from which the district can select. Four states use their current ELP test for the initial proficiency screening,[8] while two states[9] allow districts to choose between the state ELP test and a screener. As is apparent, there is substantial variability in how states initially classify ELs. This variability makes cross-state comparisons difficult where the classification tool differs or the methods of applying the same classification tool vary. Currently, there is a strong potential that students in one state who are classified as an EL or I-FEP might not be so classified in another. In states with local control of EL classification tools and procedures, this variation may occur across school districts within a state (NRC, 2011, p. 86). Of course, as with initial determination of potential EL status, the confirmation or disconfirmation of EL classification may generate errors that include false positives (wrongly identified as EL) and false negatives (wrongly identified as I-FEP).

Guidance

States within a given consortium (ELP or academic) need to have consistent initial EL classification tools and procedures, or, in the case of states in overlapping (ELP and academic) consortia, demonstrate that their tools and procedures lead to comparable initial EL classification results.

a. **Identify common EL classification policies, tools, and procedures *within* each ELP assessment consortium.** The variability in how states – and in local-control states, districts within states – classify ELs makes identifying a common EL definition across states within a consortium practically impossible without some standardization. Local educational agencies (LEAs) within consortia states must be encouraged to use the same classification/placement instrument, be it a screener, a nationally normed ELP assessment, or the annual state ELP assessment. Guidelines, common policies and procedures, and training must be developed and implemented to assure the appropriate use of the identified classification/placement instrument. A monitoring process should also be put in place to support the standardization of policy, procedures, and practices within and across states.

b. **Establish comparability of EL classification results *across* the two ELP assessment consortia.** Once standardized classification policies, procedures, and practices are in place within each consortium, a method of establishing *comparability* between ELP assessment consortia should be established. At a minimum, either a linking or equating study should be conducted between ELP consortia's EL classification/placement instruments. This could be accomplished in a variety of ways. (For more on this procedure, see Kolen & Brennan, 2004.) The goal of this study would be to create comparable ELP classifications between consortium instruments. Additionally, an examination of the comparability of consortium classification and placement policies,

---

7   Of these 27, 18 use one of the screener tests developed by the WIDA Consortium (the W-APT or the MODEL); 3 use the LAS Links Placement test, 4 use their own screener; 1 uses the LAB-R; and one uses the Woodcock Muñoz Language Survey.
8   Alaska, Arizona, California, and Florida
9   Connecticut and Nevada

procedures, and practices should be conducted. The goal is to demonstrate the commonality of these policies, practices, and procedures.

c. **Establish comparability of EL classification results for "stand-alone states" in each of the *academic* assessment consortia not participating in either ELP assessment consortium.** To move toward a common EL definition requires that classification tools and procedures of "stand-alone states" within academic assessment consortia be appropriately included. A comparability study between consortia classification/placement policies, tools, and procedures and stand-alone states' classification/placement policies, tools, and procedures should be conducted. The procedures used to establish cross-consortia comparability as described above could be applied to stand-alone states.

Additional Options

Federal law requires local educational agencies to notify parents of a students' ELP status within 30 days (ESEA, Section 1112.(g)(1)(A)). Once students are identified as potential ELs and found through initial assessment to be ELs, they are classified as such and placed in an English language instruction educational program. What if a student is misidentified and misplaced as either an EL or I-FEP? Current federal law provides no guidance for these cases. Recognition of such misidentification/misplacement may occur well after a student has been placed (or not placed) into an English language instructional program. A common procedure may also need to be developed to detect and resolve these "false positive" and "false negative" cases with respect to student EL classification and placement in language instruction educational services.

One option might be for states or consortia to establish procedures for detecting and rectifying student misidentification/misplacement into or out of EL programs within a reasonable timeframe after initial placement. Tools to support this type of decision-making procedure might take the form of a checklist or rubric. EL educators could be trained to administer these tools to confirm detection of "false positive" or "false negative" classifications. Another option that may be farther off, as it would have national implications, is to have states and ELP assessment consortia explore and work toward the creation of a common EL screener. This screener could be used to support consistent EL classifications across states regardless of a state's particular ELP assessment.

3. Defining the "English proficient" performance standard

Federal law requires states to annually assess ELs in four domains: reading, writing, listening, and speaking (ESEA Section 1111(b)(7)).[10] The law also requires states to monitor EL students' progress in attaining English language proficiency in these domains and in comprehension. This requirement has motivated states to create domain (listening, speaking, reading, and writing) and composite (oral, literacy, comprehension, and overall) scores for their ELP assessments.

10   Section 1111(b)(7) states: "ACADEMIC ASSESSMENTS OF ENGLISH LANGUAGE PROFICIENCY – Each State plan shall demonstrate that local educational agencies in the State will, beginning not later than school year 2002–2003, provide for an annual assessment of English proficiency (measuring students' oral language [further clarified as listening and speaking by section 3121(d)(1)], reading, and writing skills in English) of all students with limited English proficiency in the schools served by the State educational agency…."

Issues

States are required by federal law to establish ELP standards that correspond to the challenging academic content and performance standards in the areas of reading/English language arts, mathematics, and science (ESEA, Section 3113.(b)(2)). Accordingly, states or states within consortia have developed or adopted a variety of different ELP standards. Each set of ELP standards operationalizes the English language proficiency construct in different ways. These differences pose challenges to establishing what "English proficient" means.

Virtually all states use some form of linear weighted overall composite score for ELP progress monitoring, attainment, and accountability. However, states create the overall composite score in different ways. For example, the California English Language Development Test (CELDT) currently weights each domain equally (0.25 x Listening + 0.25 x Speaking + 0.25 x Reading + 0.25 x Writing)[11] to create its overall composite score (California Department of Education, 2012). ACCESS for ELLs (the WIDA Consortium's assessment) weights its overall composite in favor of literacy skills (0.15 x Listening + 0.15 x Speaking + 0.35 x Reading + 0.35 x Writing) (WIDA, 2012), and the Texas English Language Proficiency Assessment System (TELPAS) (Texas Education Agency, 2012) weights its composite such that reading has substantial prominence (0.05 x Listening + 0.05 x Speaking + 0.75 x Reading + 0.15 x Writing).[12] In effect, what it means to be English proficient on CELDT, ACCESS for ELLs, or TELPAS, based on overall composite scores, is very different. In establishing an English proficient performance standard on ELP assessments, states have used a variety of judgmental and/or empirical standard-setting approaches. These different approaches also make comparing English-language proficiency assessment results a challenge.

Guidance

a. **Create common performance level definitions of English proficiency.** To establish common performance level definitions, it is necessary to compare English language proficiency descriptions across states. This first step is essential to ascertain key similarities and differences in various definitions of English language proficiency. For states sharing common ELP standards (e.g., WIDA), this is a straightforward process. For states not sharing common standards, a comparison is required. While this might seem straightforward, it is quite involved in its application. At a minimum, a body of expert stakeholders should meet and create a provisional set of common performance descriptions at key performance levels (e.g., beginning, intermediate, and advanced). There is an immediate need to accomplish this task since the PARCC and Smarter Balanced assessment consortia begin field testing in the spring of 2014. A common set of performance definitions, even in provisional form, could aid in determining the consistent application of appropriate EL assessment accommodations. Once these provisional common definitions are created, the ELP assessment consortia could refine the English proficiency performance definitions and anchor their assessment scores and levels to that refined version. An illustrative example of a common performance level descriptor is provided in Appendix A.

---

11   In grades 2–12. In grades K–1, Reading and Writing are weighted 0.05 each while Listening and Speaking are weighted 0.45 each. See the 2012-13 CELDT Information Guide (p. 30), available at http://www.cde.ca.gov/ta/tg/el/resources.asp.

12   See the 2013 Texas Student Assessment Program Interpreting Assessment Reports for TELPAS, pp. 5.4-5.5, available at www.tea.state.tx.us/index3.aspx?id=3282&menu_id=793.

b. **Articulate a clear rationale for ELP assessment composite score weighting and provide evidence of the efficacy of the identified weighting procedure**. A clear articulation of what "English proficient" means on ELP assessments should be a minimum expectation for states or consortia. English proficiency, as identified on these assessments, is a function of the overall composite score. Composite score weighting articulates how states or consortia value each domain relative to its notion of English proficiency. Accordingly, states or consortia should examine their composite score weighting method. A rationale comprised of relevant theories, expert judgment, and evidence of efficacy should be articulated for ELP composite score weighting procedures. As PARCC and Smarter Balanced assessments become operational, a reanalysis of ELP assessments' composite score weighting procedures should be conducted (e.g., using multiple regression techniques). The intention is for ELP assessment composite score weighting procedures of different ELP consortia and stand-alone states to be brought more in-line using EL student results from these ELP assessments and PARCC or Smarter Balanced assessments.[13] Results from such studies should be part of the evidence collected to establish the validity of ELP assessments in these states.

c. **Conduct studies that examine English proficient criteria using relevant educational outcomes**. As noted above, the Elementary and Secondary Education Act (*No Child Left Behind Act,* 2001) defines an English learner (i.e., a student defined under ESEA as Limited English Proficient) as a student in elementary or secondary school (italics have been added for emphasis)

> whose difficulties in speaking, reading, writing, or understanding the English language *may be sufficient* to deny the individual (i) the ability to *meet the State's proficient level of achievement on State assessments* described in section 1111(b)(3); (ii) the ability to *successfully achieve in classrooms* where the language of instruction is English; or (iii) the opportunity to *participate fully in society.* Section 9101 (25)(D)

As illustrated in the italicized sections, the statute *does not* state that an English learner *must* be proficient on state content assessments, successfully performing in class, or fully participating in society in order to be considered English proficient. It only indicates that English proficient students *have the potential* to successfully achieve and participate as defined. When EL students' difficulties in speaking, reading, writing, or understanding English no longer deny them this potential, they can be considered English proficient.

Empirically, three types of studies should be undertaken to explore this. The first examines the relationship between students' English language proficiency and academic content performance on associated assessments. The goal of this research is to identify the ELP level(s) that is sufficient to no longer impede academic content performance, and therefore warrant removal of specialized language support services. Recent empirical research methods (Cook et al., 2012; Francis et al.,

---

13   Smarter Balanced is currently utilizing a language complexity rating tool (see Cook & MacDonald, 2012) to "tag" its assessment items and tasks. This can provide valuable information for research and development efforts by examining the performance of ELs at different levels of English language proficiency on content assessment items and tasks of different language complexity.

2011) allow for such analyses. (See Appendix B for a brief discussion of these methods.) The second study examines EL students' English language proficiency and their capacity to participate in the disciplinary practices of English-medium classrooms. The goal of this line of inquiry is to identify the ELP level that is sufficient for effective classroom participation. The last study attempts to operationalize the term "fully participate in society." This study examines how English language proficiency interacts with full societal participation, and could be operationalized by examining the language demands of deeper learning skills that reflect the cognitive, interpersonal, and intrapersonal capacities needed for 21st century life and work (discussed further below; see also the National Research Council's report [2012a] on this topic). All three studies can be leveraged to help validate over time the performance levels and score ranges where English learners might be identified as English proficient on ELP assessments.

d. **Conduct comparability studies between ELP assessments' English proficient scores**. There are a variety of acceptable procedures that could be used to establish cross-state or -consortia comparisons of English proficiency (e.g., equipercentile or regression approaches – see Gottlieb & Kenyon 2006; Kolen & Brennan, 2004; Crocker & Algina, 1986). At a minimum, consortia should agree upon an acceptable approach and apply it. Careful consideration of composite score weighting should also be part of this process.

Additional Options

Once the above guidelines have been enacted, a cross-state and -consortia concordance of ELP assessment scores representing English proficiency could be created. This concordance would take composite weighting into account and be anchored to the common performance level descriptions of English proficiency. The concordance could provide a common measure of English proficiency across states regardless of consortia membership.

4. Reclassifying English Learners

The act of reclassifying an English learner (i.e., exiting them from EL to "former EL" status) is significant because it signals that educators have determined an EL student no longer needs specialized linguistic and academic support services they are entitled to receive under civil rights law. Under federal law, once a student is exited from EL status, local educators are expected to describe the former EL student's annual academic progress for a two-year monitoring period (ESEA Title III, Sec.3121(a)(4)). States are allowed (but not required) to include the performance of former ELs in their Title I adequate yearly progress (AYP) calculations for the EL subgroup during the two-year monitoring period, but are not permitted to do so beyond that timeframe.

Issues

EL reclassification is complex from both technical and policy perspectives. For example, researchers have documented issues in using nonlinguistic and noncognitive criteria for reclassification decisions, as well as local criteria that are noncomparable within and across states (Linquanti, 2001; Ragan &

Lesaux, 2006; Wolf et al., 2008; NRC, 2011; Working Group on ELL Policy, 2011); in determining appropriate cutpoints of assessment-related criteria and timing of service removal (Robinson, 2011); and in reporting outcomes and holding educators accountable given an unstable EL cohort and temporary monitoring of the former EL cohort (Abedi, 2008; Saunders & Marcelletti, 2013; Hopkins et al., 2013). In effect, exit from EL status is a high-stakes decision because a premature exit may place a student who still has linguistic needs at risk of academic failure, while unnecessary prolongation of EL status (particularly at the secondary level) can limit educational opportunities, lower teacher expectations, and demoralize students (see Linquanti, 2001; Callahan, 2009; Robinson, 2011).

Additionally, there is a conceptual disconnect between meeting the "English proficient" requirement for ESEA Title III and exiting from EL status as defined under ESEA Title I. Specifically, the former is based solely on the state ELP assessment, while the latter usually involves multiple criteria that include the Title III ELP assessment result, academic achievement assessment results, and other locally identified criteria.[14] Thus a student can meet the "English proficient" performance standard under Title III, yet remain EL for one or more years beyond that point, which requires continued ELP testing under Title I. This very issue generated significant contention between the federal government and states as noted in the Federal notice of final interpretations regarding the interpretation of Title III provisions on annually assessing EL students (see Spellings, 2008, pp.61837-61838). Concerns have long been expressed about maintaining a student's EL status based on nonlinguistic performance criteria that could relate more to aptitude in mathematics or reading/language arts, which monolingual English speakers may also have difficulty demonstrating. Moreover, serious validity concerns arise when using academic achievement assessments that are neither designed nor intended to support inferences about EL students' English-language proficiency, as well as using course grades or grade point average (GPA), which regularly include noncognitive/nonlinguistic factors (e.g., attendance and homework submissions) and are rarely standardized.

The following guidance attempts to address these issues in three ways: 1) It moves toward greater standardization by conceptualizing reclassification (exit) criteria using the federal definition of an EL; 2) it attempts to address the *linguistic contribution* to academic performance *while not requiring a minimum level of performance on an academic content assessment for exit*; and 3) it allows for defensible, locally-defined exit criteria that can still be deemed comparable within and across states.

---

14   States use a variety of criteria in reclassifying (exiting) ELs to former EL status. According to data collected in 2006-07 school year (Wolf et al., 2008), over 70% (34) of 48 states surveyed use multiple (between two and six) criteria in reclassification decisions. Specifically
- 12 states use an ELP assessment only, while 2 states use only district-established criteria
- The remaining 34 states surveyed use multiple criteria
    - 11 consider the ELP test and one other criterion
        - 7 states additionally use content-area assessment achievement scores
        - 3 states additionally use district-level criteria
        - 1 state additionally uses school-level criteria
    - 23 states use the ELP test and *two to five* additional criteria, including those mentioned above as well as parent/guardian input and "other."
Adding to this variation, many states permit locally established criteria that vary *within* a state, thus leading to non-uniform, within-state definitions of EL.

a. **Conceptualize reclassification criteria using the federal definition of an EL.** Since consortia are required to use the federal definition of an EL in establishing their common definition, there is a compelling rationale for moving toward greater standardization in reclassification criteria based on this definition. As noted above, the current federal definition of EL rests on a student's linguistic minority background, non-English-speaking environment and language use, and associated difficulties in English reading, writing, speaking, and comprehension that *may* deny them the ability to *meet the State's proficient level of achievement* on State assessments; to *successfully achieve in classrooms* where the language of instruction is English; or the opportunity to *participate fully in society*. At the previous stage, we recommended that consortium member states consider multiple criteria for determining when a student is likely to be "English proficient." If it can be demonstrated empirically that meeting multiple criteria chosen to address these ELP dimensions of the federal EL definition very likely removes the linguistic barriers to EL students' *possibility* of meeting these expectations, then that would constitute evidence for justifying EL reclassification.[15] Such a conceptualization for reclassification criteria is illustrated in Figure 3 below.

Figure 3.

---
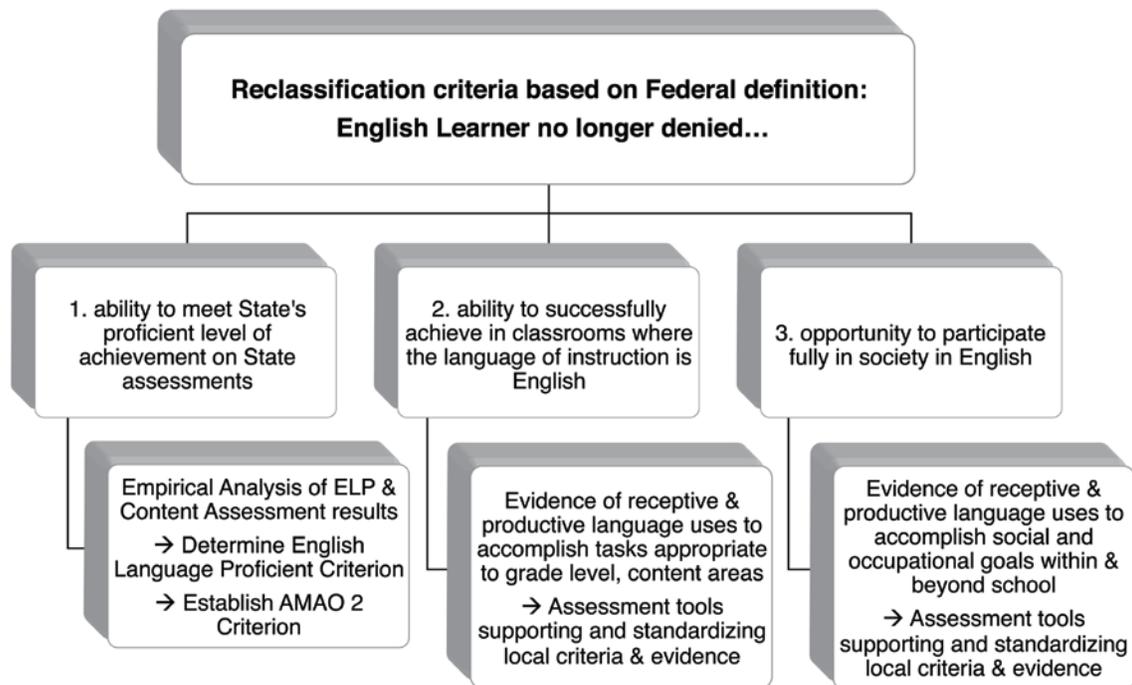
15   Reclassification could be granted provisionally during the two-year monitoring period to ensure that the decision is sound. The new common core state standards and next-generation science standards will require all teachers to build the capacity of all students to engage in sophisticated language uses to carry out more rigorous content area practices. Special attention can be paid during this time to ensure recently-exited EL students are supported to continue progressing toward and sustaining academic proficiency. In fact, several researchers have argued for monitoring former ELs' performance beyond the two-year period, and holding educators accountable for long-term outcomes of students that entered the school system initially EL. (See Hopkins et al., 2013; Saunders & Marcelletti, 2013; and the Working Group on ELL Policy, 2011, for discussions of stabilizing the EL cohort for monitoring and accountability purposes.)

Note that the English proficient performance standard on the state/consortium ELP assessment discussed in stage 3 above (and illustrated in 1. in the left-hand boxes) would constitute part of the reclassification criteria. Note also that the additional dimensions of English-language proficiency reflected in the federal definition (illustrated in 2. and 3. in the center and right-hand boxes, respectively), which we recommend be included in a multi-study validation effort of the English proficient performance standard on the ELP assessment, would also yield additional reclassification criteria that address these dimensions of the federal EL definition. Proceeding in this way would unify the criteria for defining "English proficient" with those for defining readiness for reclassification and exit from EL status.

b. **Consider reclassification criteria that address the *linguistic contribution* to academic performance and societal participation *while not requiring a minimum level of performance on an academic content assessment for exit*.** As noted above, recent methodological advances permit relating ELP assessment performance to academic content assessment performance in ways that allow policymakers and educators greater confidence in identifying a range of performance on ELP assessments that can be considered English proficient. Such approaches should obviate the need for explicitly incorporating academic content assessment performance in reclassification criteria. This is an important shift as several states currently require a minimum level of performance on their state content achievement tests for exit from EL status, despite the problematic nature of this practice for reasons of construct validity, measurement, and accountability (see Working Group on ELL Policy, 2011, pp. 5-7 for further discussion).

As explained earlier, ELP consortia members need to identify a theoretically sound, empirically informed performance standard or performance range on the shared ELP assessment. ELP stand-alone states participating in academic content assessment consortia could also participate in this process. Studies are needed to examine relationships of results from ASSETS, ELPA21, and stand-alone states' ELP assessments to the academic performance outcomes on the Smarter Balanced, PARCC, and stand-alone states' summative assessments. Such studies would help to identify the point at which EL students are identified as having sufficient English skills to be considered English proficient. These studies, which will need to be done over time using empirical data from several states, can provide helpful insights and recommendations for consortia policymakers' consideration. This can in turn lead to a much more comparable, aligned set of performance standards across consortia member states for identifying those students ready to be reclassified as former ELs.

Criteria relating to the other two dimensions of the federal EL definition (i.e., being able to successfully achieve in classroom-based practices using English, and to participate fully in society using English) have traditionally been operationalized locally. These are discussed next.

c. **Develop assessment tools that can be used to support and standardize local criteria for their relevance and construct/predictive validity for use in reclassification decisions.**

Given the additional dimensions of the federal EL definition, and the long-standing practice in many states permitting educators to use local criteria in exit decisions, ELP and academic consortia member states should consider collaborating to develop standardized assessment tools (e.g., rubrics or rating scales) that capture key receptive and productive language functions required for performing analytical tasks in different academic disciplinary practices used in the classroom. These in turn could be used to help teachers more consistently collect evidence and evaluate student performance along these relevant dimensions. These assessment tools supporting and standardizing local criteria and evidence can be used by states to monitor local reclassification practices, which in turn can be used by federal monitoring visits to examine state oversight of local practices.

Similarly, regarding the dimension of successfully participating in society, extended standardized assessment tools (rubric or rating scales) can help identify those receptive and productive language uses that support students to accomplish social and occupational goals within and beyond school. While this is currently the least operationalized aspect of the federal LEP definition, recent work by the National Research Council's Committee on Defining Deeper Learning and 21st Century Skills (NRC, 2012a) can help to shed light on salient language dimensions of transferable knowledge and skills needed for life and work in the 21st century. Some of these language uses are already found in new content and corresponding ELP standards. They can be situated within the cognitive, interpersonal, and intrapersonal capacities described by the NRC Committee as correlating with positive education, workplace, and health outcomes, and as potentially reducing disparities in educational attainment (NRC, 2012b, p. 2).

As noted above, studies need to be conducted on these additional reclassification criteria that lead to informed decisions. Taken together, these studies can provide states, regardless of consortia membership (ELP and academic content), with tools to reclassify ELs with respect to their English language proficiency.

Additional Options

States will likely face challenges in moving toward an integration of "English proficient" criteria with reclassification (exit) criteria. For example, several states have laws requiring the use of state academic achievement assessment results in determining an EL student's eligibility for reclassification to former EL status. It will take time to build understanding and consensus regarding the multiple criteria that adequately capture EL students' academic and social uses of language that could help validate the English-proficient performance standard, and justify a reclassification decision. Given these realities, an additional option might be for a group of states to engage in pilot R&D efforts to prototype the assessment tools mentioned for supporting and standardizing these

other criteria. These efforts can use the provisional common ELP descriptions described above, consult documents such as CCSSO's *English Language Proficiency/Development Framework* (2012), and consider consortium and state ELP standards documents.

Next steps for moving forward: A proposed strategic workplan

To support the process of developing a common definition of English learners, the following proposed workplan is provided in Table 1. A plan for each step in the process is shown with high-level tasks, proposed completion dates (year and quarter), and a span of participation. Three participation levels are identified: national, across-consortia, and within-consortia. Some tasks are best accomplished through national representation (i.e., consortia and non-consortia [stand-alone] states). For example, a national group of English language development stakeholders should include ASSETS and ELPA21 consortia state representatives and representatives from states such as California, Texas, and New York. These last three states do not currently belong to an ELP assessment consortium, yet are among the largest EL-enrolling states. In general, high-level tasks include a development phase, a validation phase, and an implementation phase. Substantial collaboration between stakeholder groups is assumed to be part of the process at each phase. Some tasks have immediate time constraints (e.g., development of a preliminary common performance level descriptor [PLD] for English proficient). This is of immediate concern as the Smarter Balanced and PARCC consortia require a common indicator of an EL's language proficiency classification in order to apply accommodations during field testing which is scheduled to occur in the 2013-2014 school year. Other tasks have a longer timeline (e.g., empirical analysis and validation of English proficient classification). This activity cannot occur until all consortia have operational assessment results available.

Figure 4: Proposed Workplan for Moving Toward Establishing a Common EL Definition

Activity	Description of High-level Task	Complete By	Span of Participation		
			National	Across-Consortia	Within Consortia
1. Identifying Potential ELs	Development of a common Home Language Survey draft protocol	2013 Q4	✔		
	Validation of common Home Language Survey protocol	2014 Q2		✔	
	Final protocol development, training, and implementation	2014 Q4			✔
2. Establishing initial EL Classification	Initial classification draft protocol(s)	2013 Q4		✔	
	Validation of classification protocol(s)	2014 Q2		✔	
	Final protocol development, training, and implementation	2014 Q4			✔
3. Defining English Proficient	Development of a preliminary common performance level descriptor (PLD) for English proficient	2013 Q4	✔		
	Communication & implementation of preliminary common PLD for English proficient	2013 Q4		✔	
	Empirical analysis and validation of English proficient classification	2015 Q4		✔	
	Development, implementation, and communication of final PLD for English proficient	2015 Q4		✔	
4. Reclassifying ELs	Development of reclassification prototyped tools	2013 Q4		✔	
	Validation of reclassification prototyped tools and recommendations	2014 Q4		✔	
	Final reclassification tools and recommendations	2015 Q3			✔

Conclusion

The complex policy and technical issues involved in developing a common EL definition are going to require a well-defined roadmap of processes and decisions for all consortia members to enact over time. Given the different permutation of states involved in the four consortia, this work is best engaged via close coordination and frequent communication within and across consortia. All phases and criteria — including initial identification as potential EL, EL classification, determination of English-proficient status, and reclassification — will need to be addressed, using all consortia member states' ELP and academic content assessments.

It is prudent to approach the issue of creating a common definition of an English learner as a multi-staged, multiyear, deliberative process. As assessments come online, teachers begin to teach to the Common Core State Standards, and educational systems align to the expectations of college- and career-readiness, a refined understanding of English language proficiency will emerge. States and the consortia to which they belong should plan now for this process. To that end, we have offered a review of key issues and opportunities, and proposed guidance on how states and consortia might respond in moving toward a common definition of English learner.

References

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practices, 27*(3), 17-22.

Bailey, A., & Kelly, K. (2012). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy,* first published May 18, 2012. Retrieved January 16, 2013, from http://epx.sagepub.com/content/early/2012/05/17/0895904811432137.

Callahan, R., Wilkinson, L., Muller, C., & Frisco, M. (2009). ESL placement and schools: Effects on immigrant achievement. *Educational Policy, 23,* 355–384.

California Department of Education. (2012). *Overview of the California English language development standards and proficiency level descriptors.* Sacramento, CA: Author. Retrieved July 23, 2013, from http://www.cde.ca.gov/sp/el/er/documents/overviewpld.pdf.

Castaneda v. Pickard, 648 F. 2d 989, U.S. Court of Appeals, 5th Circuit. 1981. Retrieved July 23, 2013, from http://scholar.google.com/scholar_case?case=16848723757397550913&hl=en&as_sdt=2,44&as_vis=1.

The Civil Rights Act of 1964. Public Law 88-352, July 2, 1964. Stat. 241, H.R. 7152. Retrieved July 23, 2013, from http://library.clerk.house.gov/reference-files/PPL_CivilRightsAct_1964.pdf.

Cohen, A., & Bolt, D. (2005). A mixture model analysis of differential item functioning. *Journal of Educational Measurement, 42*: 133–148.

Cook, H.G., Linquanti, R., Chinen, M., & Jung, H. (2012). *National evaluation of Title III implementation supplemental report: Exploring approaches to setting English language proficiency performance criteria and monitoring English learner progress.* Washington, DC: U.S. Department of Education, Office of Planning, Evaluation and Policy Development.

Cook, H.G., & MacDonald, R. (2012). *Tool to evaluate language complexity of test items.* Unpublished manuscript. Madison, WI: WIDA Consortium, Wisconsin Center for Education Research.

Council of Chief State School Officers. (2012). *Framework for English language proficiency development standards corresponding to the Common Core State Standards and the Next Generation Science Standards.* Washington, DC: Author.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory.* Belmont, CA: Holt, Rinehart, and Winston.

The Equal Educational Opportunities Act of 1974. Public Law 93-380, August 21, 1974. Stat. 20 USC Sec. 1701-1758.

Francis, D., Lesaux, N., & August, D. (2006). Language of instruction. In D. August & T. Shanahan (Eds.), *Developing literacy in second language learners: Report of the National Literacy Panel on language-minority children and youth* (pp. 365-413). Mahwah, NJ: Lawrence Erlbaum.

Francis, D., Tolar, T., & Stuebing, K. (2011, May). *The language-achievement connection for English Language Learners.* Presentation at the M3 Conference at the University of Connecticut. Retrieved June 23, 2013, from http://www.modeling.uconn.edu/handouts/D2S5_Francis_m3conference-presentation-ELLs.pdf.

Gottlieb, M. and Kenyon, D.M. (2006). *The bridge study between tests of English language proficiency and ACCESS for ELLs®.* WIDA Consortium Technical Report No. 2. Madison, WI: WIDA Consortium, Wisconsin Center for Education Research. Retrieved July 23, 2013, from http://www.wida.us/get.aspx?id=132.

Hopkins, M., Thompson, K.D., Linquanti, R., Hakuta, K., & August, D. (2013). Fully accounting for English learner performance: A key issue in ESEA reauthorization. *Educational Researcher, 42*(2), 101-108.

Improving America's Schools Act of 1994. Public Law 103-382. October 20, 1994. 108 Stat. 3518.

Kolen, M.J. and Brennan, R.L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd Ed). New York: Springer-Verlag.

Lau v. Nichols et al., 414 U.S. 563. 1974. Retrieved July 23, 2013, from http://caselaw.lp.findlaw.com/cgi-bin/getcase.pl?court=US&vol=414&invol=563.

Linquanti, R. (2001). *The redesignation dilemma: Challenges and choices in fostering meaningful accountability for English learners.* Policy Report 2001-1. Santa Barbara: University Of California Linguistic Minority Research Institute.

Linquanti, R., & Cook, G. (2013). *Toward a "common definition of English learner": A brief defining policy and technical issues and opportunities for state assessment consortia.* Washington DC: CCSSO. Retrieved July 23, 2013, from http://www.ccsso.org/Documents/2013/Common%20Definition%20of%20English%20Learner_2013.pdf.

National Research Council. (2011). *Allocating federal funds for state programs for English language learners.* Panel to review alternative sources for the limited English proficiency allocation formula under Title III, Part A., Elementary and Secondary Education Act, Committee on National Statistics and Board Testing and Assessment. Washington, DC: National Academies Press.

National Research Council. (2012a). *Education for life and work: Developing transferable knowledge and skills in the 21st century* (J.W. Pellegrino and M.L. Hilton, Eds.). Committee on Defining Deeper Learning and 21st Century Skills. Board on Testing and Assessment and Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

National Research Council. (2012b). *Education for life and work: Developing transferable knowledge and skills in the 21st century.* Report Brief (July 2012). Board on Testing and Assessment. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

The No Child Left Behind Act of 2001 (NCLB). Public Law 107-110, Jan 8, 2002. Stat.115.1425-2094.

Ragan, A., & Lesaux, N. (2006). Federal, state, and district level English language learner program entry and exit requirements: Effects on the education of language minority learners. *Education Policy Analysis Archives*, *14*(20). Retrieved July 23, 2013, from http://epaa.asu.edu/ojs/article/view/91.

Robinson, J.P. (2011). Evaluating criteria for English learner reclassification: A causal-effects approach using a binding-score regression discontinuity design with instrumental variables. *Educational Evaluation and Policy Analysis, 33*(3), 267–292.

Saunders, W. & Marcelletti, D. (2013). The gap that can't go away: The Catch-22 of reclassification in monitoring the progress of English learners. *Educational Evaluation and Policy Analysis, 35*(2), 139-156.

Slavin, R., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English-language learners. *Review of Educational Research*, *75*, 247-281.

Spellings, M. (2008). Title III of the Elementary and Secondary Education Act of 1965 (ESEA), as amended by the No Child Left Behind Act of 2001 (NCLB). Notice of final interpretations. *Federal Register*, Vol. 73, n.202. pp.61837-61838.

Texas Educational Agency (TEA). (2012). ELPS-TELPAS proficiency level descriptors. *ELPS-TELPAS Proficiency Level Descriptors.* Austin, TX: TEA.

Towles-Reeves, E., Kearns, J., Flowers, C., Hart, L., Kerbel, A., Kleinert, H., Quenemoen, R., & Thurlow, M. (2012). *Learner characteristics inventory project report (A product of the NCSC validity evaluation).* Minneapolis, MN: University of Minnesota, National Center and State Collaborative. Retrieved January 16, 2013, from http://www.ncscpartners.org/Media/Default/PDFs/LCI-Project-Report-08-21-12.pdf.

U.S. Department of Education (USED). (2010*). Race to the Top Assessment Program Guidance and Frequently Asked Questions.* Washington DC: Author. Retrieved January 16, 2013, from: http://www2.ed.gov/programs/racetothetop-assessment/faqs.pdf.

WIDA Consortium. (2012). *WIDA performance definitions - Speaking and writing grades K-12.* Madison, WI: WIDA Consortium, Wisconsin Center for Education Research. Retrieved July 23, 2013, from http://www.wida.us/get.aspx?id=543.

Wolf, M.K., Kao, J., Griffin, N., Herman, J., Bachman, P., Chang, S., & Farnsworth, T. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses* (Practice Review. CRESST Report No. 732). Los Angeles, CA: UCLA.

Working Group on ELL Policy. (2011). *Improving educational outcomes for English language learners: Recommendations for the reauthorization of the Elementary and Secondary Education Act. Questions and answers.* Stanford, CA: Author. Retrieved July 23, 2013, from http://ellpolicy.org/wp-content/uploads/QA.pdf.

Appendix A. Common English language proficiency performance descriptors based on an amalgamation of descriptors from CA, AZ, TX, and the WIDA Consortium

Performance Level Descriptors				
	1 **Beginning**	**2** **Expanding**	**3** **Intermediate**	**4** **Advanced**
Vocabulary	Uses high-frequency everyday words with clear, easily-demonstrated referents (*brother, hand, book, three*)	Uses expanded content vocabulary (*add, product, three, third, group*) supported by easily-demonstrated examples; common synonyms and antonyms; common affixes and roots; nominalized forms of verbs (*survive-survival*) and adjectives (*strong-strength*)	Uses broad content vocabulary (*exponent, function, similarities*); beginning to develop idiomatic language (*raise the rent, run up a bill*)	Uses content vocabulary large enough to construct and convey nuanced meaning; includes figurative language
Language Processing and Production Capacity	Can process and produce words and short, formulaic phrases (*'I'm fine, thank you.'*)	Can process and produce related sentences	Can process and produce related paragraphs and discourse	Can process and produce multiple paragraph texts and extended discourse
Cohesion (logical connections across sentences)		Able to employ loose cohesion, accomplished by repetition of words or phrases. • *The first thing I did was… The next thing I did was…* • *This figure has four sides, and the four sides are...*	Able to connect sentences in more varied ways to show sequence (*first, next, last*), comparison and contrast (*the same, similar, like all the others, unlike, in contrast*). May struggle with naturalness of phrasing.	Able to use wider range of vocabulary and more specific connectors between sentences and clauses to show logical relationships of cause and effect (*consequently, resultant, contributory*).
Conciseness; Density of information		Can follow conversational patterns with usually one idea per independent clause or sentence. Develops meaning by stringing ideas together in simple sentences or combining simple sentences with coordinating conjunctions (*and, but, for, or, nor, so, yet*).	Can use phrases and relative clauses to consolidate meaning into fewer clauses, but still has some redundancies and limitations.	Can construct concise, tightly packed clauses by nesting or embedding structures within one another.
Relevant forms and structures		Able to use coordinating conjunctions; some simple prepositional phrases	Able to use subordinating conjunctions; long noun phrases; prepositional phrases; adverbial phrases; relative clauses	Able to use phrases nested or embedded within one another in tightly packed clauses

Appendix B: Methods for Establishing an English Language Proficient Performance Standard

The *No Child Left Behind Act* (2001) defines an English learner ("Limited English Proficient" in law) as a national origin- or language-minority student in elementary or secondary school (italics have been added)

> whose difficulties in speaking, reading, writing, or understanding the English language may be sufficient to deny the individual (i) the ability to *meet the State's proficient level of achievement on State assessments* described in section 1111(b)(3); (ii) the ability to *successfully achieve in classrooms* where the language of instruction is English; or (iii) the opportunity to *participate fully in society*. Section 9101 (25)(D).

Note the three italicized sections above. An English learner's language proficiency *may* deny them 1) the ability to meet the state's proficient level on content assessments, 2) the ability to successfully perform in the classroom when instruction is in English, and 3) the opportunity to fully participate in society. It is important to recognize that the statute does not state that an English learner will be proficient on state content assessments, successfully performing in class, or fully participating in society when they are English proficient. It only indicates that English proficient students are no longer *denied the potential* to successfully achieve and participate. When students' difficulties in speaking, reading, writing, or understanding English no longer limit this potential, they can be considered English proficient. Empirically, studies could be envisioned that explore this potential. The first study could look at the relationship between students' English language proficiency and content proficiency as observed on state content and English language proficiency assessments. The goal of this research would be to identify the nexus of sufficient performances on both of these tests. The second study could look at students' English language proficiency and their success in classroom settings. This type of study could be qualitative in nature, the goal of which would be identify the English proficiency level that optimizes classroom success. The last study would need to operationalize the term *fully participate in society*. This research could examine how English language proficiency interacts with social and occupational participation. All three studies should be considered when determining the point(s) where English learners might be considered English proficient and reclassified as former English learners.

This appendix focuses only on the first type of study – the relationship between English language proficiency and content proficiency as seen on associated assessments. However, it should be understood that focusing on assessment results alone is insufficient to adequately identify the English language proficiency level that supports a reclassification decision.

Three methods are presented to examine the relationship between English language proficiency and content assessments: Equal Likelihood, Decision Consistency, and Comparable Distributions. The first two methods are described in detail in Cook, Linquanti, Chinen, & Jung (2012). The Equal Likelihood method presumes that the score on the ELP assessment where English learners are equally likely to perform proficiently on state content assessments is evidence of English language proficiency. Two approaches to examine equal likelihood are shown below. The first approach is a simple boxplot. The boxplots in Figure B.1 below display this distribution of scale scores on a state's English language arts

and mathematics assessment (y-axis) by English language proficiency level (x-axis). The line on the y-axis of these boxplots identifies the proficient score on the content assessments.

**Figure B1: Boxplots of ELA and Mathematics Performance by ELP Level[16]**
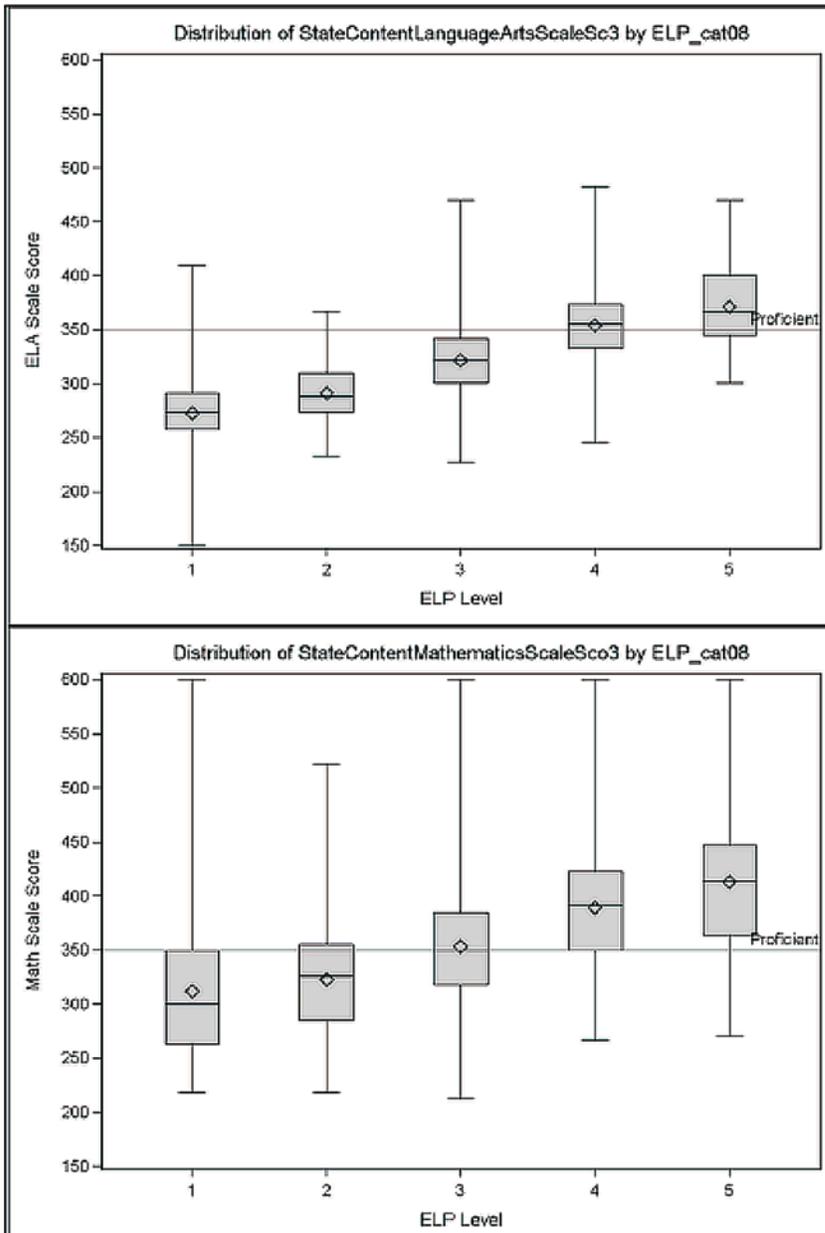
Notice that the lower the language proficiency level, the greater the distribution of students below the academic-proficient line. In the boxplots, the diamond inside each box is the mean score and the line in the center of the box is the median score. When the median score goes above the proficient line on the y-axis, half of the English learners at a given ELP level are performing proficiently on the content test (i.e., they are equally likely to be proficient than not). With the Equal Likelihood method, this is the level where discussions about English proficiency should begin. However, the

16   Graphic adapted from Exhibit 3 in Cook et al., 2012.

boxplots in the above figure show two different points. For English language arts (the top boxplot), the median value is above the proficient line at level 4, but for mathematics it is at level 3. This is not unexpected, as students participating in ELP assessments do so because they have limited English proficiency. Variability in performances is a function of several factors (e.g., the linguistic complexity of the assessments, the academic uses of language expected of students at different grades, etc.). That is why several methods are recommended. Those making decisions about the English-language proficient performance standard should use a preponderance of evidence (i.e., several methods) across grades and subjects to support decisions.

A second Equal Likelihood approach is logistic regression. Logistic regression estimates the likelihood that an EL will be proficient on a content assessment based on specific ELP assessment scores. Figure B.2 below plots likelihood results from a logistic regression analysis on the same data used in the boxplots above. The x-axis shows the English language proficiency assessment scale scores. The vertical lines represent English language proficiency level cutpoints. Scores to the left of the first vertical line are at level 1, and scores to the right of the last vertical line are at level 5. The horizontal line on the y-axis marks an equal (50/50) probability of meeting the proficient performance standard on the given academic achievement test. As English language proficiency scale scores increase on the x-axis, the likelihood that an EL will be proficient in English language arts (top graph) and mathematics (bottom graph) also increases. For English language arts, the 50% line is at a scale score in level 4. For mathematics, the 50% line is crossed at a scale score in level 3. The findings in the boxplot approach and logistic regression approach provide similar results. The benefit of the logistic regression approach is the ability to create confidence intervals, which provide a measure of precision at decision points on the English language proficiency scale.

Figure B.2: Logistic Regression of ELA and Mathematics Performance by ELP Level
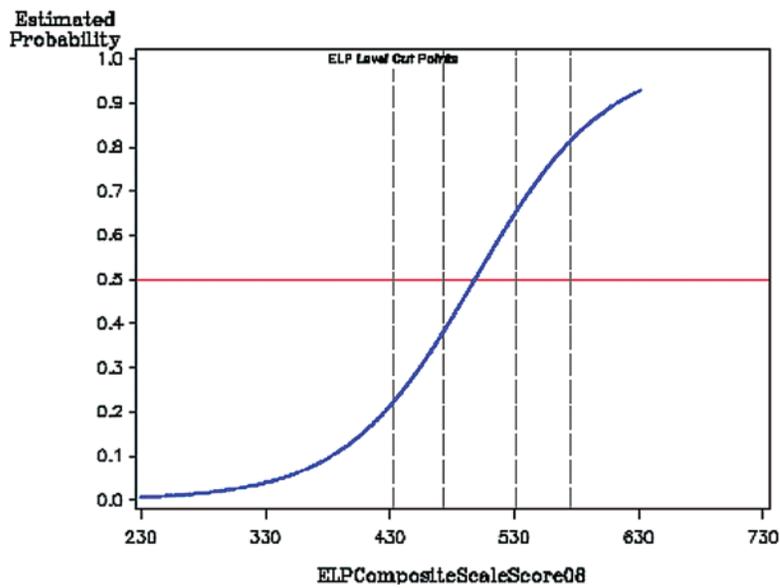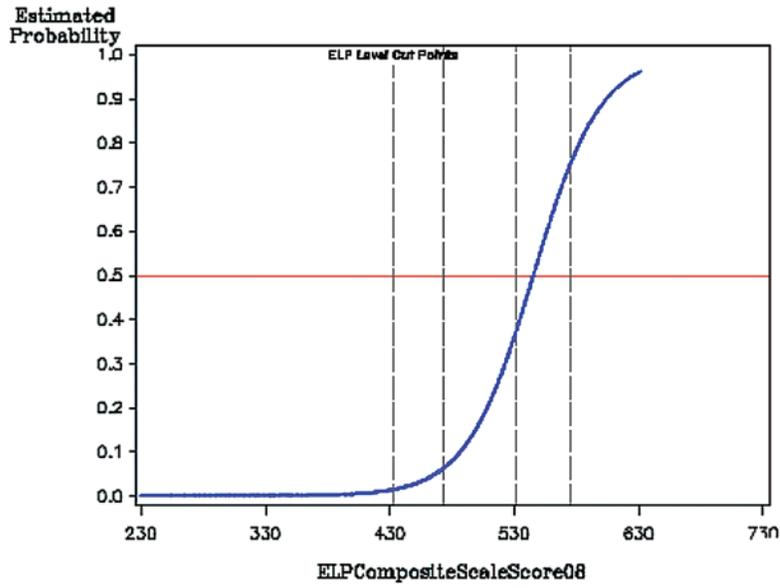

The second method is Decision Consistency. This method examines the decisions made with each assessment. Consider the following matrix in Figure B.3 below.

Figure B.3: Decision Consistency Matrix

		English Language Proficiency (ELP) Cut-Score (*TBD*)	
		Not Proficient	Proficient
Academic Content Assessment Cut-Score (*Given*)	Proficient	Quadrant I	Quadrant II
	Not Proficient	Quadrant III	Quadrant IV

This matrix compares two assessment-related decisions. Arrayed in rows are proficient and not proficient categories on state content assessments. In columns, the proficient and not proficient categories for the English language proficiency assessment are displayed. Four cells are created when these two decision categories are crossed, each represented by a quadrant. Quadrant I refers to English learners who receive a proficient score on the state content test but not on the language proficiency test. We would like the numbers in this cell to be minimized. Quadrant II reflects students who score proficient on both the state content and language proficiency assessment. Quadrant III identifies students who receive non-proficient scores on both the state content and English language proficiency test. It is desired to maximize the numbers of student in Quadrants II and III. Quadrant IV contains students who do not attain a proficient score on the state content test but do score proficient on the English language proficiency assessment. We would like to minimize the numbers of students in this cell. Using the following formula, we can identify the percent of "consistent decisions" (i.e., made at the intersection of both content and ELP assessment cut-scores): (QII + QIII) / (QI + QII + QIII + QIV). Given that the content test's proficient performance standard is usually already established, a calculation of the percentage of consistent decisions at a variety of English language proficiency scores can be made. Figure B.4 below displays a plot of the percentages of consistent decisions between a state's English language arts and mathematics content assessment and its English language proficiency assessment.

Figure B.4: Decision Consistency Plot
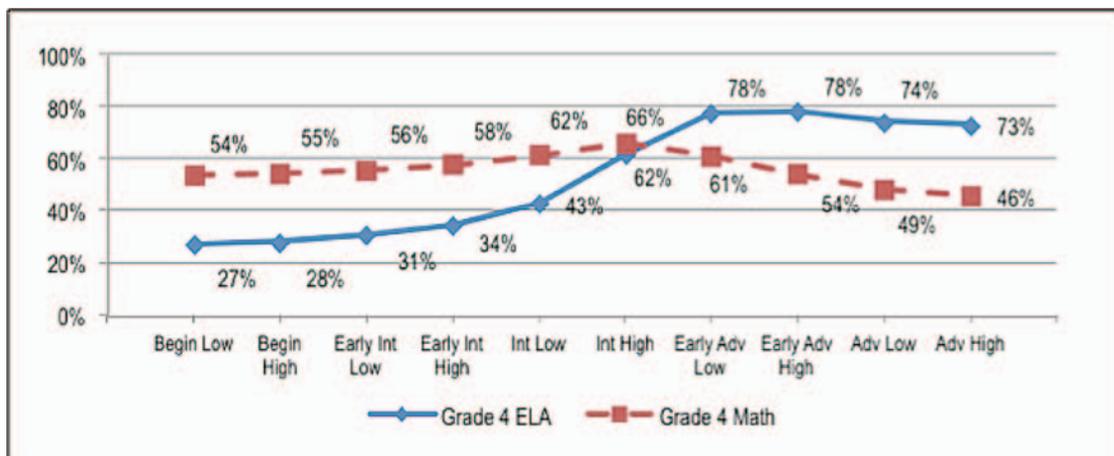
This figure assumes that the state's proficient scores on its content assessments are given and not adjusted. What is examined is the decision consistency when the English language proficiency scores are adjusted. The x-axis shows 10 different categories comprised of 5 proficiency levels (Beginning, Early Intermediate, Intermediate, Early Advanced, and Advanced) each split into a low and high category. The y-axis shows the percentage of consistent decisions by academic content area (ELA and Math). As ELP categories increase, the percent of consistent decisions also increases. This means that higher English language proficiency cut-scores increase consistent decisions regarding students' performance on the English language arts and mathematics assessments (i.e., more students are in Quadrants II and III). However, there is a point on the ELP assessment scale where consistent decisions regarding content assessments are maximized and then drop. For English language arts assessment, that is at the Early Advanced Low (Early Adv Low) category with a decision consistency value of 78%. For the mathematics assessment, it is at the Intermediate High (Int High) category with a value of 66%. It is at the maximum decision consistency value, or the point just before percentages drop, that discussion about establishing the English language proficient performance standard should occur. Data used to create the line graphs in Figure B.4 were the same as in Figures B.1 and B.2, and therefore findings from the Decision Consistency method corroborate those seen in the Equal Likelihood method. (See Cook et al., 2012 for further discussion.)

The final method examines performance differences on content assessments between English learners and non-English learners at the item level. This Equivalent Distribution method identifies the language proficiency score or level that does not differentiate item performances between English learners and non-English learners. (See Francis, Tolar, & Stuebing, 2011; and Colin & Belt, 2005 for more details and specific models.) The logic of this approach is displayed in Figure B.5 below.

Figure B.5: Conceptual Model of the Equivalent Distribution Method
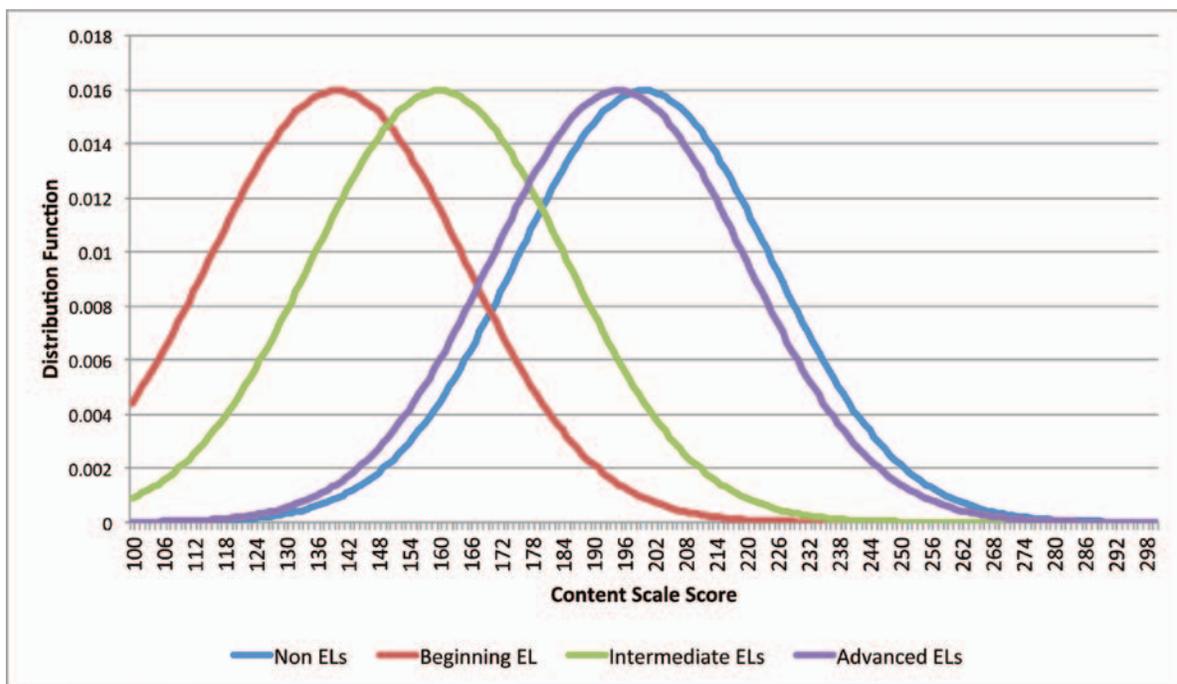
This figure illustrates four theoretical distributions of test/item performances on a hypothetical content assessment. The group performances are of English learners scoring at the beginning (Beginning ELs), intermediate (Intermediate ELs), and advanced (Advanced ELs) levels of English proficiency, as well as non-English learners (Non ELs). Notice that the distribution of test/item performances begin to overlap as the language proficiency level increases. The goal of this method is to identify the ELP level or score that does not differentiate between English learners and non-English learners. It is at that score range or level that discussions about the English-language proficient performance standard should begin.