MOVING TOWARD A MORE COMMON DEFINITION OF ENGLISH LEARNER:
Collected Guidance for States and Multi-State Assessment Consortia

CCSSO
Council of Chief State School Officers

THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

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.

Moving Toward a More Common Definition of English Learner:
Collected Guidance for States and Multi-State Assessment Consortia

Authors:

Robert Linquanti, WestEd

H. Gary Cook, Wisconsin Center for Education Research

Alison L. Bailey, University of California, Los Angeles

Rita MacDonald, Wisconsin Center for Education Research

Acknowledgments:

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

Table of Contents

Preface

Dear Colleagues,

English Learners (ELs) are now one of the fastest growing populations of students – this is the case in schools, districts, and states across the country. We are learning so much about these students and how they gain English proficiency. We know that states' higher, clearer standards and assessments are providing a pathway to graduation and college and career success for them. But the pathway remains complex as students enter, navigate, and exit EL status – and this complexity lies in how specifically an EL is defined, not only within and across states, but within and across multi-state consortia as well.

The challenge before us lies in defining all phases and criteria of ELs as they enter – as potential ELs, classified ELs, level of English proficiency – and as they exit – as former EL and/or English proficient.

With the generous funding of the Carnegie Corporation of New York, CCSSO is collaborating with states, consortia, and EL education leaders to develop processes, recommendations, and tools to build a more consistent and more coordinated definition of an EL. Although the concept of a definition may seem straightforward and even simple, this is a multi-staged, multiyear, complex process incorporating new standards, new assessments, and an evolving and growing understanding of just how to best support a student learning English and gaining proficiency. To support states that are dealing with this challenge, CCSSO has sponsored a series of national working sessions and working papers that have been compiled into this publication.

States and consortia are actively engaged in this process, and we thank you and the Carnegie Corporation for joining us and for your work and dedication to these students. For more information on CCSSO's progress toward a common definition of EL, please visit http://www.ccsso.org/Resources/Programs/Effort_on_Moving_toward_a_Common_Definition_of_English_Learners.html.

Best,

Chris Minnich

Executive Director
Council of Chief State School Officers

Introduction

English Learners (ELs)—language-minority students whose English proficiency affects their ability to meaningfully participate and succeed in school—are expected to reach 25% of the total U.S. K-12 public school population by the year 2025. Yet, how states and school districts define this population varies widely, creating inconsistent and possibly inequitable services for students who move across state or even district boundaries.

States participating in any of four federally-funded assessment consortia (i.e., PARCC, Smarter Balanced, WIDA, and ELPA21)[1] have agreed to a U.S. Department of Education stipulation to establish a "common definition of English Learner" within their consortia. Fulfilling this common definition requirement is neither simple nor straightforward because individual states vary in their policies, tools, and practices for determining which students are ELs, what specialized services they receive, and what criteria and processes are used to exit them from this status. Additionally, federal and state statutes, case law, and regulations influence requirements for K–12 public school educators. These requirements call for a carefully coordinated, multiyear effort within and across consortia member states. Such an effort needs to proceed in stages and encompass several critical decisions informed by student performance outcomes on new assessments.

The Council of Chief State School Officers (CCSSO), with support from the Carnegie Corporation of New York and in conjunction with the Understanding Language initiative of Stanford University and the WIDA Consortium, established a cross-consortium *English Language Learner Assessment Advisory Task Force* in 2012 to address challenges presented by new college- and career-ready standards and assessments. The Advisory Task Force—with technical staff and leadership from the above-mentioned consortia, EL researchers and technical assistance experts, and policy advisers and other stakeholders—identified the common EL definition as a key issue.

The goal of this effort was to define the key issues involved and provide actionable 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, and pragmatic to the many policy, technical, and legal issues. This includes: 1) developing policy guidance and technical assistance activities to enhance state and district knowledge of critical issues; 2) identifying effective practices that can be utilized to establish a more consistent definition of ELs within and across states; and 3) disseminating research, guidance, and resources to states, school districts, assessment consortia, researchers, and advocacy associations.

From 2012–2015, Advisory Task Force members held meetings, facilitated three national working sessions with a broad representation of national, state, and local stakeholders, and produced a series of working papers that provide guidance on key policy and technical issues in defining ELs. Two members of the Advisory Task Force, Robert Linquanti of WestEd and Gary Cook of Wisconsin Center for Education Research, produced initial guidance featuring a 4-stage framework to assist states in analyzing issues and strengthening policies and practices for defining ELs:

---

[1]   Smarter Balanced Assessment Consortium (Smarter Balanced), Partnership for Assessment of Readiness for College and Careers (PARCC), WIDA's Assessment Services Supporting English Learners through Technology Systems (WIDA ASSETS), and English Language Proficiency Assessment for the 21st Century (ELPA21).

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/consortium English language proficiency (ELP) test against which to assess ELs' English proficiency; and

4. **Reclassify** a student to former-EL status through the use of multiple exit criteria.

These four stages describe the basic trajectory of how a student enters, moves through, and exits EL status, and served as the basis for subsequent work.

Specifically, this framework was used to organize a series of national working sessions involving states and consortia, non-consortia "stand-alone" states, and several school districts, and to produce detailed guidance documents. As such, this collaborative effort lays out a roadmap of processes, options, and recommendations for moving toward a common definition so that efforts can be coordinated within states, as well as within and across consortia. All phases and criteria—including initial identification as potential EL, EL classification, determination of English-proficient status, and reclassification—need to be addressed, using states' ELP and academic content assessments.

This publication compiles the related guidance and working papers published by CCSSO from 2013 to 2015. It defines key issues and provides guidance that assessment consortium member states can use to move toward establishing a more consistent EL definition in ways that are theoretically sound, evidence based, empirically informed, pragmatic, and sensitive to many technical, legal, and policy issues. This work is also of value to states not in an assessment consortium that wish to strengthen EL definitional policies and practices within their states, as well as to foster equitability of services and comparability of outcomes for ELs across school districts within their states. Please take note of the publication dates for each of the working papers. Given the rapid pace of change and fluctuation in consortia membership, it is important to note that the working papers provide historically accurate portrayals of consortia membership at the time each document was produced. The publication date of each of the papers is indicated in the title page preceding each document.

The introductory guidance document provides context, describes several key issues, and lays out the organizing framework for a more common EL definition. Each of the subsequent working papers explores in depth one of the four stages identified in the introductory guidance informed by national working sessions composed of national, state, and local EL educators, researchers, policy advisers, and advocates as indicated in the appendix of each working paper. The following provides a brief description of each document.

Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options (Linquanti & Cook, 2013) outlines central issues and offers introductory guidance in moving toward a more common EL definition. In particular, this paper discusses policy and technical options for defining English learners using the suggested four-stage framework of key criteria and processes. It effectively sets the stage for the four working papers which follow.

Reprising the Home Language Survey: Summary of a National Working Session on Policies, Practices, and Tools for Identifying Potential English Learners (Linquanti & Bailey, 2014) is the first in the series of working papers that elaborate on the four-stage framework described in the introductory guidance. It summarizes discussions on the identification of potential English learners held by a national working session of state and consortium representatives, experts and stakeholders held on September 18, 2013 at the CCSSO offices in Washington, DC.

Strengthening Policies and Practices for the Initial Classification of English Learners: Insights from a National Working Session (Cook & Linquanti, 2015) examines issues and options associated with the second stage of the framework, initially classifying ELs, and develops ideas discussed at a second national working session held on May 23, 2014 at the CCSSO offices in Washington, DC.

Reference Performance Level Descriptors: Outcome of a National Working Session on Defining an "English Proficient" Performance Standard (Cook & MacDonald, 2014) explores a critical facet of the third stage (defining an "English proficient" performance standard) of the framework and presents a set of Reference Performance Level Descriptors developed with guidance provided at a national working session held on September 17, 2013 at the CCSSO's Washington, DC offices.

Finally, *Re-examining Reclassification: Guidance from a National Working Session on Policies and Practices for Exiting Students from English Learner Status (Linquanti & Cook, 2015)* examines issues and options associated with reclassifying ELs to fluent English proficient (R-FEP) status. The report focuses on the final stage of the framework. It summarizes and further develops ideas discussed at a national working session held on September 23-24, 2014 at the CCSSO's Washington, DC offices. In particular, it recaps significant issues and tensions surrounding current EL reclassification policies and practices; provides a review of current reclassification criteria for the 50 states and Washington, DC; and offers guidance to districts, states, and multi-state consortia for moving toward more common EL reclassification criteria.

Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options

Robert Linquanti, WestEd
H. Gary Cook, WCER
2013

Overview

The purpose of this paper is to define key issues 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. 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 defining ELs:

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/consortium English language proficiency (ELP) test against which to assess ELs' English 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.[2] 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

---

[2]   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.

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.[3] 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 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.

---

[3]   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.

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 work plan 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)
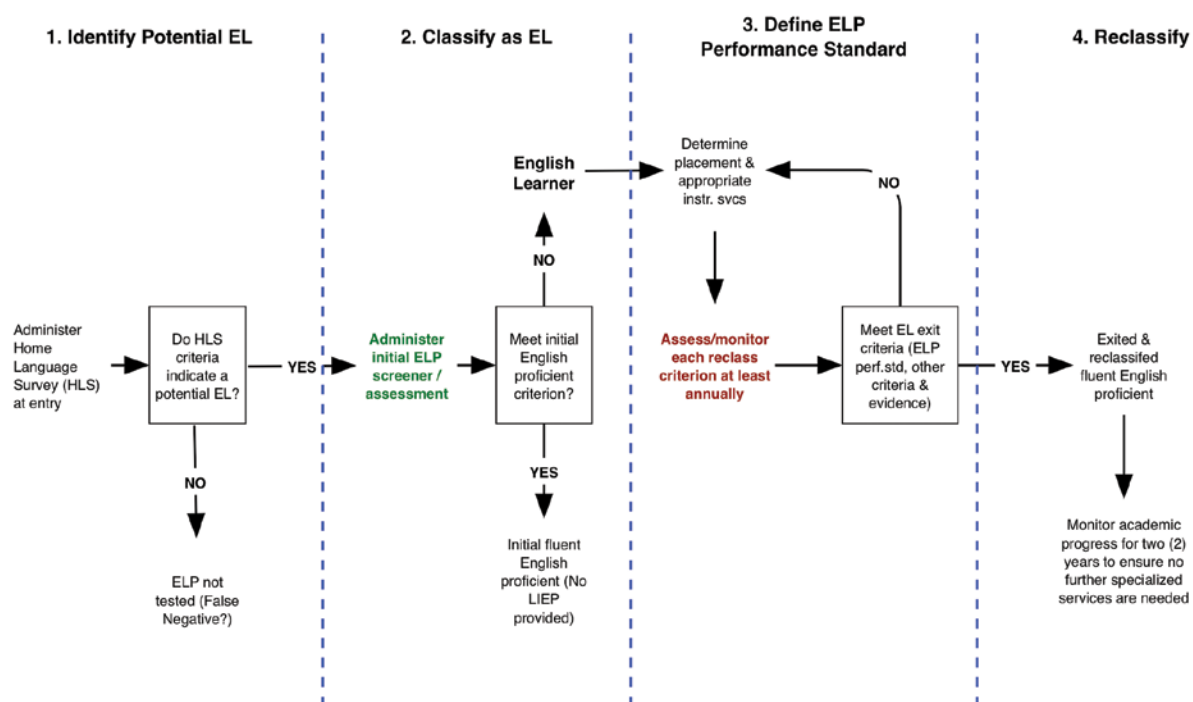
Consortium member states need to explicitly define their approaches to these four key steps, and establish mechanisms to move toward a sufficient degree of "convergence" or alignment over time. We now describe issues and opportunities for consortium member states to consider in each of these four steps.

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 kindergartner 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 in 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[4]), 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.

Issues

According to a recent report by the National Research Council (NRC, 2011), 27 states use a screener/placement test for EL classification.[5] 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,[6] while two states[7] 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

---

[4]   Term "initially fluent English proficient" is defined in the *Establishing Initial English Learner Classification* section below.

[5]   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.

[6]   Alaska, Arizona, California, and Florida

[7]   Connecticut and Nevada

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, 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 non-EL? 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)).[8] 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.

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

---

[8]   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...."

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)[9] 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).[10] 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.

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

---

[9]   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.

[10]   See the 2013 Texas Student Assessment Program Interpreting Assessment Reports for TELPAS, pp. 5.4-5.5, available at http://tea.texas.gov/student.assessment/interpguide/.

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.[11] 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., 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.

---

[11]   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.

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.[12] 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.

Guidance

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

---

[12]   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.

meeting these expectations, then that would constitute evidence for justifying EL reclassification.[13] Such a conceptualization for reclassification criteria is illustrated in Figure 3 below.

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.

Figure 3.
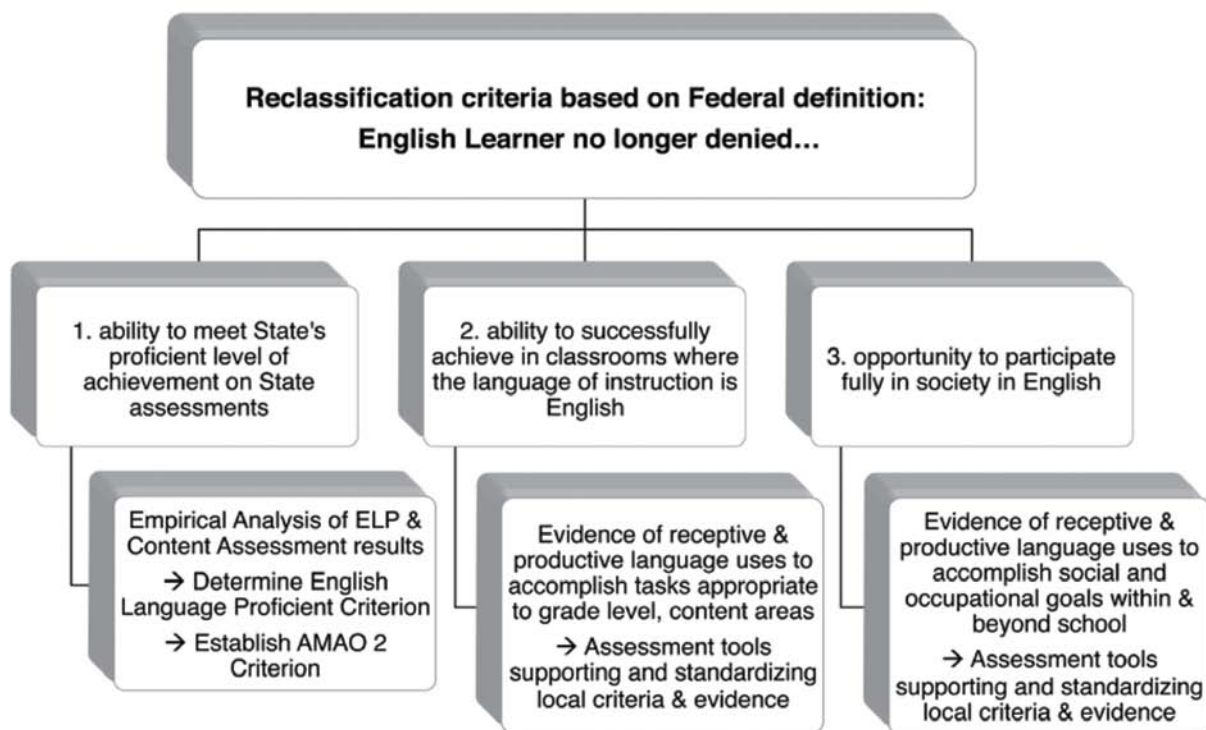
b. **Consider reclassification criteria that address the** linguistic contribution **to academic performance and societal participation** while not requiring a minimum level of performance on an academic content assessment for exit. As noted above, recent methodological advances permit

---

[13] 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.)

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/sbeoverviewpld.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.

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 descriptions 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[14]
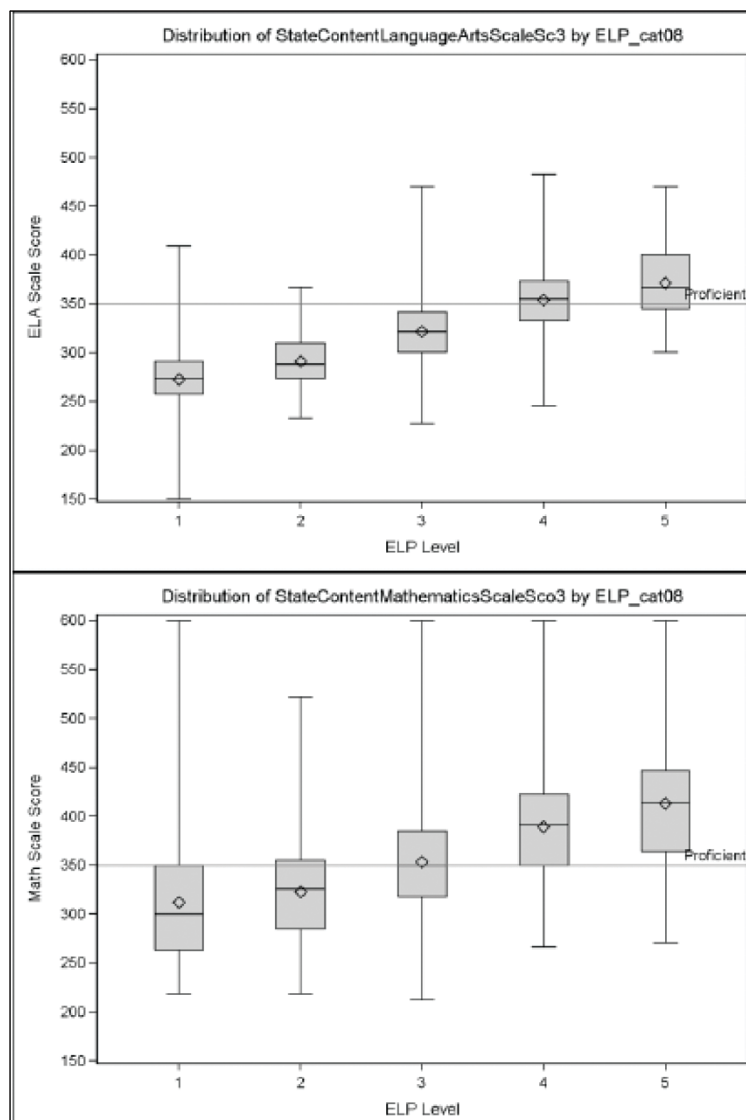
Notice that the lower the language proficiency level, the greater the distribution of students below the academic-proficient line. In the boxplots, the diamond inside each box is the mean score and the line in the center of the box is the median score. When the median score goes above the proficient 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 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.

---

[14]   Graphic adapted from Exhibit 3 in Cook et al., 2012.

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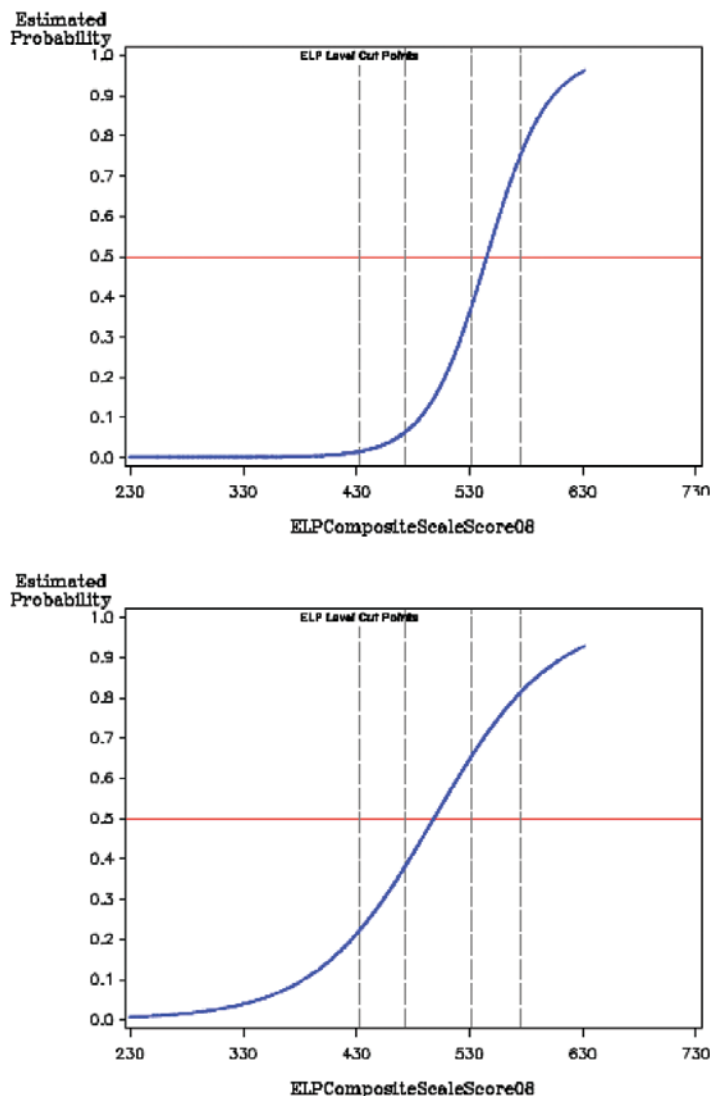
The second method is Decision Consistency. This method examines the decisions made with each assessment. Consider the following matrix in Figure B.3 below.

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
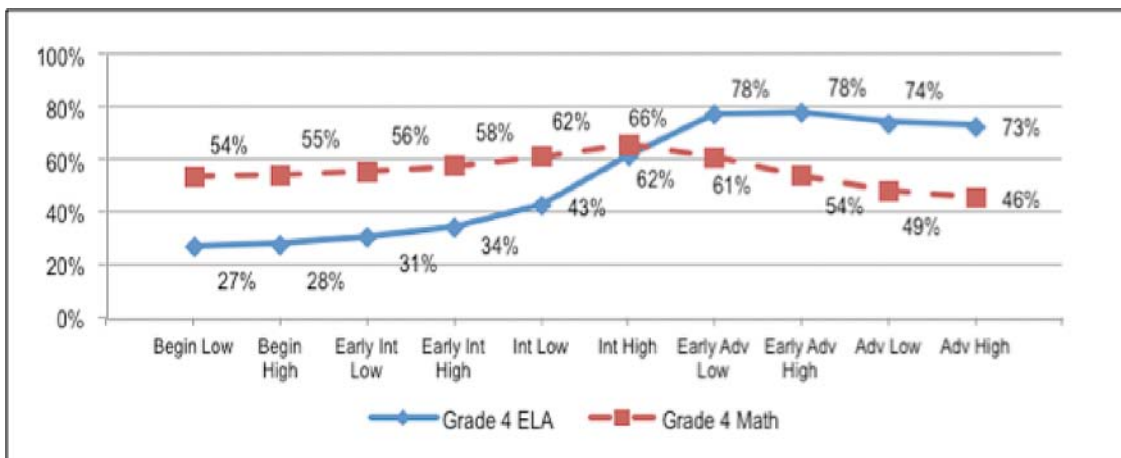
This figure assumes that the state's proficient scores on its content assessments are given and not adjusted. What is examined is the decision consistency when the English language proficiency scores are adjusted. The x-axis shows 10 different categories comprised of 5 proficiency levels (Beginning, 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
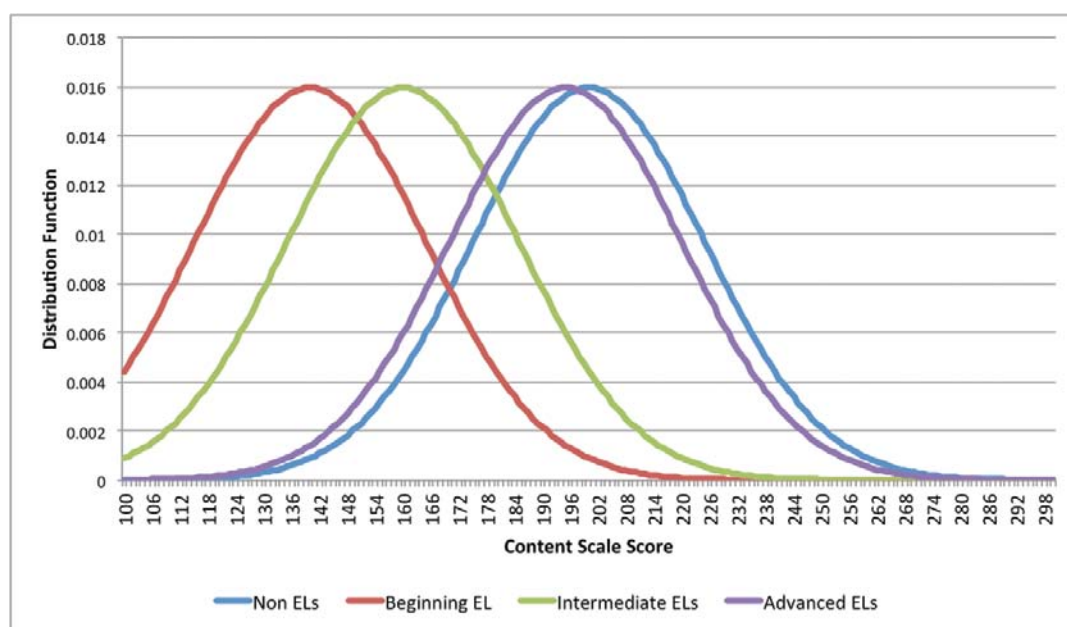
This figure illustrates four theoretical distributions of test/item performances on a hypothetical content assessment. The group performances are of English learners scoring at the beginning (Beginning ELs), intermediate (Intermediate ELs), and advanced (Advanced ELs) levels of English proficiency, as 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.

STAGE 1

Reprising the Home Language Survey: Summary of a National Working Session on Policies, Practices, and Tools for Identifying Potential English Learners

Robert Linquanti, WestEd
Alison L. Bailey, UCLA
2014

Overview:

This document elaborates on the first of four key stages in moving toward a common definition of English learner (EL), as described in the CCSSO publication, *Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options* (Linquanti & Cook, 2013).

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/consortium English language proficiency (ELP) test against which to assess ELs' English proficiency; and

4. **Reclassify** a student to former-EL status through the use of multiple exit criteria.

It summarizes a national working session of state and consortium representatives, experts and stakeholders held on September 18, 2013 at the CCSSO offices in Washington, DC. Participants and the institutions they represented are listed in Appendix A. Specifically, the working session deliberated on the CCSSO guidance document's first stage—identifying a student as a potential English learner—in order to build understanding and consensus among assessment consortia, participating states, and other stakeholders on key issues, and on strengthening related policies, practices, and tools.

Background

English learner (EL) students are legally entitled and expected to receive specialized instructional services and assessments to ensure they develop academic uses of English that enable them to succeed in school, work, and life (LaCelle-Peterson & Rivera, 1994; Hakuta, 2011). Yet the provision of these supports is largely predicated on accurately identifying potential English learners via Home Language Surveys (HLS) (Bailey, 2010). As noted in the CCSSO guidance for states in consortia required to move toward a common EL definition (Linquanti & Cook, 2013), states should first identify from the overall K-12 student population the group of students that is *potentially* EL. Per the definition of an English learner by the Elementary and Secondary Education Act (ESEA)[15], states are required to identify students with a native language other than English, or who come from an environment where a non-English language either is dominant[16] 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. 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 of the guidance's common EL definition framework.

Current HLS Practices in the US

The vast majority of states use home language surveys to identify potential EL students (Bailey & Kelly, 2010; 2013; Kindler, 2002; National Research Council [NRC], 2011; Wolf et al., 2008). However, as Bailey and Kelly (2013) point out, "The U.S. Department of Education Office for Civil Rights *December 3, 1985 Memorandum* and *1991 OCR Policy* address the requirement to have a program in place for adequately identifying students in need of services, but recognize that this may differ widely due to student demographics. No wording in these memoranda obligates states to specifically enforce the use of an HLS in order to initially identify students" (p. 2). In their review, Bailey and Kelly (2010; 2013) identified four HLS practices common in the states: 1) 23 states and Washington, DC have created a single HLS form and mandate its use in schools statewide; 2) 17 states mandate use of an HLS and have created a sample HLS for districts to adopt or substitute with their own version; 3) six states mandate use of an HLS but have created neither a required nor sample HLS, allowing districts to create their own; and 4) four states do not require use of an HLS but may list its use along with additional recommended practices such as use of existing reading scores on state tests and observational scales.

---

[15]   ESEA Sec. 9101(25): "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."

[16]   Although this term is used in the legal definition, session participants had concerns about using it with respect to a student's "language dominance." See Section B. below. See also Thompson (2013).

An HLS is typically administered to parents/guardians at a student's initial school enrollment. They include a varied number of differently worded questions that essentially ask parents/guardians for the same information: to indicate whether a language other than English is spoken by the student, and by the parents/guardians to the student, in the home. HLS are not intended to measure English *"proficiency."* However, they sometimes have a secondary purpose of yielding information about students' language backgrounds that can aid schools in making programmatic decisions, such as offering bilingual education in English and high-incidence languages revealed by the HLS. To date, no validity studies have been conducted by states on the accuracy of HLS, so little is known about the quality of the information HLS yield (Bailey, 2010). Where there has been opportunity to evaluate HLS, researchers have found them lacking. For example, one study found to be problematic the use of a single question asking for the student's primary language (Goldenberg & Rutherford Quach, 2010), and HLS being used as the lone indicator of potential EL status has been questioned (Abedi, 2008; Bailey & Kelly, 2013).

CCSSO National Working Session on HLS

In light of the federal requirement for states in the four consortia to move toward a common EL definition, and as a consequence of the current concerns with HLS outlined above, CCSSO convened State and Consortia representatives, researchers, and other key stakeholders to explore how to strengthen and align HLS development and use in consortium member states. The group discussion was structured around five key dimensions of HLS:

A. Purposes and intended uses

B. Key constructs (aligned to purposes and uses)

C. HLS question formulation (aligned to constructs)

D. Administrative procedures

E. Decision rules

Group discussion of each of these dimensions is summarized below.

A. Purposes and Intended Uses

Session participants reached a general consensus that the purposes and intended uses of the HLS are to:

1. Identify potential members of a protected class under Federal Law by helping school officials identify students who may benefit from and be entitled to specialized English language development services, and services that help ensure full access to the academic curriculum;

2. Identify who should be assessed using the English language proficiency (ELP) screener/test to determine EL status; and

3. Better understand students' language environments and use.

Considerations

1. Clearly defining the purposes and uses of the HLS helps guide identification of key constructs, which in turn facilitates the development of more precise HLS questions.

2. The purposes and intended uses of the HLS should be made explicit to those administering and those completing the survey.

3. States may want to clearly specify what the purposes and intended uses of the HLS are **not** (e.g., determining immigration status, pre-determining educational services to be provided).

4. Since language use in the federal definition extends beyond the home, states may want to consider calling it the *Language Use Survey.*

5. Some states integrate the HLS into a larger initial intake form, which can impact the HLS format and context, and may blur purpose and intended uses. In these cases, the HLS should be clearly distinguished from other intake purposes and questions. For example, the following "tiered" intake form has been utilized by the New York City Department of Education and consists of three distinct parts:[17]

   a. Tier (Part) 1 – "Language Assessment Battery-Revised [screener] Eligibility[18]": Essential HLS questions presented here; identifies student as in the protected class that requires further screening

   b. Tier (Part) 2 – "Instructional Planning": Collects other information needed for instruction/placement decisions

   c. Tier (Part) 3 – "Parent Information": Identifies parent's language of choice for receiving school communications

Issues

1. Sec. 9101(25) (A): Age range 3-21 years: Should Pre-Kinder and adult education (to 21 years old) be included? The Federal definition says in "elementary or secondary" school.

2. Sec. 9101(25) (C) "not born in U.S.": Should HLS help determine this? The purpose of HLS is not to determine parent or student immigration status, yet U.S. birth status is part of the protected class definition (see next issue).

3. Current ESEA definition of limited English proficient (LEP) includes within the protected class certain national-origin group members (e.g., Native Americans, Alaskan Natives, and natives of outlying areas) who may be primary (or even monolingual) speakers of a non-standard or contact variety of English (e.g., "Village English")

---

[17]   For further information see The New York City DOE Parent/Guardian Home Language Identification Survey at http://schools.nyc.gov/NR/rdonlyres/2CF11E0F-BDBE-4DCB-9E7B-7452A6AB6CDF/0/HLIS_5_24_07_ENGLISH.pdf
[18]   On February 1, 2014, the New York State Identification Test for English Language Learners (NYSITELL) will replace the LAB–R as the instrument used to confirm (or disconfirm) EL status for initial EL classification.

4. Other speakers of non-standard English varieties are not currently specified within the definition's protected class (e.g., African American Vernacular English [AAVE] speakers) but may also benefit from systematic opportunities to develop the standard-English code or more academic uses of English; since these students fall outside the current Federal LEP definition, a separate task force may be needed to make recommendations on augmenting these students' linguistic repertoire.

B. Key Constructs

The review of state HLS practices (Bailey & Kelly, 2010; 2013) suggests that the constructs being captured on HLS to date have focused primarily on a student's first language learned, language "dominance", and exposure to English and/or other language(s) (i.e., language(s) spoken among adults in the home, and/or language(s) spoken to a student in the home). Other foci in extant HLS included country of origin, prior language services placement, and time in U.S. schools. As noted above, especially when schools and districts use multi-purpose intake approaches, these latter constructs may reflect some HLS being used for purposes beyond collecting information on home language use (e.g., eligibility for immigrant student services).

Considerations

1. Identifying key constructs to be represented on HLS is central to developing more precise HLS questions.

2. Given the purposes of HLS noted in Section A above, the constructs captured by HLS can be considered in terms of degree of relevance for a student's *current* English abilities:

    a. **Essential**: (student's current language(s); frequency of English language use by student; frequency of English language exposure provided by others)

    b. **Associated**: (Languages spoken among adults in the home; history of student's language environment, such as "linguistic history" [a student's first language spoken]; use of other language(s); years in U.S. schooling, etc.)

    c. **Irrelevant**: (country of origin[19])

3. The HLS is not intended as a measure of English *"proficiency"*; rather it is a tool to help gauge the language environments a student has been exposed to and thus the *likelihood* a student has acquired English abilities prior to enrolling in a U.S. school.

4. English language exposure and contexts of English use include these relevant environments:

    a. the home/household

    b. school/child care settings

    c. recreational (non-home/school) settings

---

[19]   However, see section A., Issue 2 above.

5. States participating in consortia can more consistently address essential constructs (see consideration 2a above) by using common HLS questions.

6. **Essential** constructs may be measured by questions that appear as "Tier 1" on a tiered intake form, identifying a student as a member of the protected class requiring follow-up English language proficiency screening/assessment.

7. **Associated** constructs may be measured by questions that appear as Tier 2 on a tiered intake form, collecting additional useful information as needed for instruction/placement.

8. In any multi-purpose intake form, the HLS should be clearly demarcated, with its questions exploring only essential and associated constructs (Tiers 1 and 2).

Issues

1. Participants raised the question of whether to include on HLS (perhaps via Tier 2 questions) the use of non-standard varieties of English (with students in the protected class only).

2. Participants considered the construct of student language *dominance* to be outmoded. Rather, participants focused more on the *contexts of exposure* to English as key to providing relevant information about the likelihood of potential EL status.

3. Reporting fluent bilingualism in English and an additional language should not lead to over-identification of potential EL status *if* the focus is on student English language exposure and use.

4. Participants concluded that parents' use of language(s) (between themselves and among others but not with the student) may be misleading in identifying student's potential EL status.

5. Participants expressed concern that over-identification of potential EL status may occur where parents report the first language as non-English (although not accurate) due to cultural pride; this constitutes another reason why student's first language used (vs. current language use) may not be an essential construct for determining potential EL status.

6. Concern was also expressed that a student's SES may be associated with lack of exposure to academic English.

7. Distinguishing between a student's comprehension (receptive) and production (expressive) abilities in English could be explored in validation studies.

8. Exposure to literacy could be educationally valuable to surveys (Tier 2) for students who have already received schooling in the U.S. or elsewhere.

C. HLS Question Formulation

Participants generated and discussed the following possible HLS questions.

Note: None of these questions has been piloted for ease of comprehension; they are only suggested starting points for further adjusting of wording when tried out with parents/guardians.

Considerations

Tier 1 questions might include:

1. *Which language(s) does your child currently . . .*

    a. understand? [receptive language]

    b. speak? [productive language]

2. *Which language does your child most often use . . .*

    a. at home . . .

        i. with you?

        ii. with siblings?

        iii. with extended family members?

    b. in school/educational settings (i.e., preschool; child-care)?

    c. outside of school (with friends, for recreational activities)?

3. *Which language does your child most often hear . . .*

    a. at home . . .

        i. with you?

        ii. with siblings?

        iii. with extended family members?

    b. in school/educational settings (i.e., preschool; child-care)?

    c. outside of school (with friends, for recreational activities)?

Tier 2 questions might include:

1. *Is your child's language variety of English called by an alternate name (e.g., Indian English, Creole, [adjective as locally appropriate]-English)?*

2. *What language do you (the parent) most frequently use with other adults in the home?*

3. *Which language did your child speak when he/she first began to speak?*

4. *What other languages does your child regularly use/hear?*

5. *Does your child read/write in English?*

6. *Does your child read/write in a language other than or in addition to English?*

7. *How many years has your child attended U.S. schools?*

Tier 3 questions might include:

1. *In what language would you like to receive written information from the school?*

2. *In what language would you prefer to communicate orally with school staff?*

Suggestions

1. Question wording should be straightforward; parent and school personnel focus groups exploring specific question wording/formulation may be valuable.

2. Use of "your child" in the example questions above may inadvertently preclude a wider spectrum of guardians of school-age children. Consider using "the student" as a substitute.

3. Questions should be piloted with a sample of parents/guardians and school personnel across a wide range of different language groups to ensure questions are interpreted as intended and information yielded is meaningful.

4. Questions should be piloted in both print and oral forms.

5. Questions should be piloted using both translation and interpreter formats for additional (locally appropriate) languages.

6. Questions may need refinement based on parent and school personnel focus groups and piloting.

D. Administrative Procedures

Session participants reached a general consensus on the following administrative procedures:

1. Clearly state within the HLS the purposes and intended uses of information being gathered, *particularly if the HLS is embedded within a larger intake form*. In larger intake forms, clearly indicate which questions constitute the HLS.

2. Have administrative procedures reflect whether HLS administration is delivered centrally (district office) or locally (school site).

   a. Training should be standardized and annually conducted for those administering HLS

   b. Standardization should address both oral administration as well as any written forms

   c. Instructions should be provided in simple-to-use scripts (for oral administration), and frequently asked questions (FAQs) should accompany written forms.

   d. Documentation (e.g., handbook) of procedures should be provided.

3. Provide oral administration and written forms in a language and manner recipients understand

   a. All forms should be translated accurately.

    b.   Scripts for oral presentation should be accurately translated into all languages needed and delivered by fluent speakers with sufficient oral proficiency to assist respondents if necessary.

4.   Document all HLS responses and store electronically in a district central data system tied to the individual student. Responses should be part of the student's permanent data record.

    a.   HLS responses documenting potential EL status (Tier 1) should be collected *once.*

    b.   HLS responses documenting Tiers 2 and 3 (e.g., instructional purposes, home/school communication preferences) may be updated.

5.   Administrative procedures should describe exception scenarios (e.g., multiple, contradictory forms completed/provided; contradictory answers within a given form) and delineate how to resolve these contradictions:

    a.   Clarify procedures for further review if HLS data are questionable, especially if other evidence indicates student is possibly EL (e.g., false negative) (see also Bailey, 2011).

    b.   Coordinate with any screening for students with disabilities (SWD) or gifted & talented identification; if possible, this coordination should occur concurrently with or before ELP screener (to avoid false positives).

6.   Administrative procedures should specify next steps in EL identification process.

Considerations

- Conducting the HLS online may allow other options for presenting questions and offering responses (e.g., use of a scale to indicate frequency of language use).

E. Decision Rules

Although time to explore this dimension was limited, session participants identified several important considerations and guidelines related to decision rules, which are summarized below. Further discussion and elaboration of this dimension would be beneficial.

Considerations & Guidelines

1.   Decision rules should specify how survey responses should be interpreted to arrive at a decision.

    a.   Must all responses indicate English, or only some?

    b.   Are questions weighted? If so, how should the weighting of questions be determined?

2.   Question weighting rules should be devised for school personnel to apply straightforwardly (e.g., "English" given as one of the child's languages for Tier 1 sample questions 1a. and 1b., as well as "English" given for minimum of 3 responses to questions 2[a-c] and 3 [a-c]).

3. In larger, multipurpose intake forms, *only those questions reflecting HLS purposes and intended uses should be utilized in decision rules* (see constructs and administrative procedures above).

4. Decision rules should synchronize with current state and/or local laws (e.g., education code, OCR resolution agreements, etc.). Changes to decision rules may require changes to current laws/regulations.

5. Age considerations may influence procedures or decision rules (e.g., Pre-kindergarteners through Grade 2 vs. Grades 3-12).

6. Hypothetical student profiles or scenarios based on HLS response patterns may help to carefully evaluate potential EL status in light of students' range of possible language use:

    a. Preponderance of evidence vs. absolute trigger, using key HLS question responses

    b. One example: (% of HLS responses suggesting English/LOTE[20] exposure/use: Decision)

        i. 100/0 or 95/5: Don't Assess

        ii. 90/10 or 80/20: Watch for possible assessment

        iii. 60/40 or 50/50 and below: Assess

7. Empirical/experimental studies should be conducted on HLS answers in relation to ELP assessment results (to improve quality) and should include students whose parents' HLS responses suggested students were *not* potential EL.

8. Decision rules using each HLS question should be pilot-tested as a "pre-validation" check.

Conclusion

It became clear to participants during the course of this working session that many opportunities exist to strengthen the identification of potential English learners. As detailed above, the session produced insights, considerations, and suggestions for improving the tools and the procedures educators use at this critical first stage of serving language-minority students. These include:

1. Clarifying the key purposes and intended uses of the home language survey;

2. Identifying essential and associated constructs aligned to those purposes and uses;

3. Formulating HLS questions that target those essential and associated constructs;

4. Recommending key administrative procedures to ensure effective survey provision; and

5. Suggesting guidelines for decision rules to strengthen the reliability and validity of inferences and actions derived from survey results.

---

[20] Language Other Than English

Importantly, working session participants came away with a view that the home language survey itself might more properly be called the *language use survey* because it should capture the language exposure and use that students experience in the multiple contexts of household, afterschool/childcare, and recreational settings. Finally, there is clearly a need to empirically test the proposed HLS questions and decision rules in relation to ELP screener/assessment results in order to strengthen survey quality as well as to validate that the right students are being appropriately identified as potential English learners. In effect, enacting the above suggestions and recommendations may optimize the weave of the net that the home language survey casts across all new registrants to our school systems so that it more often captures those students who should move forward to the next stage of confirming potential EL status.

References

Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement: Issues and Practice*, *27*(3), 17-31.

Bailey, A.L. (2010). Implications for assessment and instruction. In M. Shatz & L. Wilkinson (Eds.). *The Education of English Language Learners: Research to Practice.* 222-247. New York, NY: Guilford Press.

Bailey, A.L. (2011). Lessons from Arizona's EL identification struggle: How revisions, guidance could strengthen process. *NCLB Advisor, 6(4),* 5-8.

Bailey, A. L. & Reynolds Kelly, K. (2010). *The Use and Validity of Home Language Surveys in State English Language Proficiency Assessment Systems: A Review and Issues Perspective.* Enhanced Assessment Grant US Department of Education, *Evaluating the Validity of English Assessments (EVEA)* Project deliverable. Available at: http://www.eveaproject.com/doc/HLS%20White%20 Paper%202010.pdf.

Bailey, A. L., & Kelly, K. R. (2013). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy, 27,* 770-804.

Goldenberg, G. & Rutherford Quach, S. (2010). *The Arizona Home Language Survey: The Identification of Students for ELL Services.* The Civil Rights Project/Proyecto Derechos Civiles, UCLA. Retrieved December 20, 2010 from http://civilrightsproject.ucla.edu/research/k-12-education/language- minority-students/the-arizona-home-language-survey-and-the-identification-of-students-for-ell- services/AZ-PHLOTEGO8-17-10revision.pdf.

Hakuta, K. (2011). Educating language minority students and affirming their equal rights: Research and practical perspectives. *Educational Researcher, 40*(4), 163-174.

Kindler, A. (2002). *Survey of the states' limited English proficient students and available educational programs and services: 2000-2001 Summary report*. Washington, DC: National Clearinghouse for English Language Acquisition.

LaCelle-Peterson, M. W., & Rivera, C. (1994). Is it real for all kids? A framework for equitable assessment policies for English language learners. *Harvard Educational Review*, *64*(1), 55-76.

Linquanti, R., & Cook, H.G., (2013). *Toward a "common definition of English learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options.* Washington, DC: The Council of Chief State School Officers.

National Research Council. (2011). *Allocating federal funds for state programs for English language learners.* Panel to Review Alternative Data Sources for the Limited-English Proficiency Allocation Formula under Title III, Part A, Elementary and Secondary Education Act. Committee on National Statistics and Board on Testing and Assessment. Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.

No Child Left Behind Act of 2001 (NCLB) (2001). Conference Report to Accompany H.R., 1, Rep. No. 107-334, House of Representatives, 107th Congress, 1st Session, December 13. Pub. L. No. 107-110, 115 Stat. 1425.

Thompson, K. D. (2013). Is separate always unequal? A philosophical examination of ideas of equality in key cases regarding racial and linguistic minorities in education. *American Educational Research Journal*, *50*(6), 1249-1278.

Wolf, M. K., Kao, J., Griffin, N., Herman, J. L., Bachman, P. L., et al. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses—Practice review* (CSE Tech. Rep. No. 732). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Appendix A.

Participants in the National Working Session on Identifying Potential English learners
(September 18, 2013, CCSSO, Washington, DC)

Participant	Representing the following institutions/organizations*
Patricia Adkisson	State Title III Directors Association (Alaska)
Supreet Anand	US Department of Education (Observer)
Rosa Aronson	Teachers of English to Speakers of Other Languages
Alison Bailey	UCLA (Session Co-Leader)
Tim Boals	WIDA
Karen Cadiero-Kaplan	California Department of Education
Martha Castellon	Understanding Language-Stanford University
Mariana Castro	WIDA
Magda Chia	Smarter Balanced
Fen Chou	CCSSO
Laurene Christensen	National Center and State Collaborative (NCSC)
H. Gary Cook	Wisconsin Center for Education Research (Facilitator)
Elizabeth Cranley	WIDA
Chane Eplin	PARCC and ELPA21 (Florida)
James Ferg-Cadima	Mexican-American Legal Defense and Education Fund
Margo Gottlieb	WIDA
Kenji Hakuta	ELPA21 (Understanding Language-Stanford University)
Margaret Ho	ELPA21 (Washington)
David Holbrook	ASSETS (Wyoming)
Rocio Inclan	National Education Association
Angelica Infante	New York State Education Department
Yvette Jackson	PARCC
Dorry Kenyon	Center for Applied Linguistics
Audrey Lesondak	WIDA (Wisconsin)
Robert Linquanti	WestEd (Session Co-Leader)
Giselle Lundy-Ponce	American Federation of Teachers
Rita MacDonald	Wisconsin Center for Education Research (Facilitator)
Luis-Gustavo Martinez	National Education Association
Martha Martinez	ELPA21 (Oregon)
Robert Measel	PARCC (Rhode Island)
Scott Norton	CCSSO
Anita Pandey	National Association of Bilingual Education

Jen Paul	Smarter Balanced
Margarita Pinkos	Palm Beach County School District
Justin Porter	Texas Education Agency
Tamara Reavis	PARCC
Lily Roberts	California Department of Education
Gail Tiemann	Dynamic Learning Maps Consortium
Gabriela Uro	Council of Great City Schools
Dan Wiener	PARCC (Massachusetts)
Lynn Willner	WestEd (ELPA21)
Carsten Wilmes	WIDA
Santiago Wood	National Association of Bilingual Education

*May differ from participant's institution of employment.

Strengthening Policies and Practices for the Initial Classification of English Learners: Insights from a National Working Session

H. Gary Cook, WCER
Robert Linquanti, WestEd
2015

Overview

This document elaborates on the second of four key stages in moving toward a common definition of English learner (EL), as described in the CCSSO publication, *Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options* (Linquanti & Cook, 2013).

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/consortium English language proficiency (ELP) test against which to assess ELs' English proficiency; and

4. **Reclassify** a student to former-EL status through the use of multiple exit criteria.

This report summarizes and further develops ideas discussed at a national working session held on May 23, 2014 to examine issues and options associated with initially classifying ELs. It is the second in a series of guidance papers intended to support states in large-scale assessment consortia that are expected to move toward a common definition of EL as part of their assessment grant requirements.

Introduction

Federal education law entitles ELs to specialized instructional services that support both English language proficiency (ELP) development and content proficiency attainment (NCLB, s. 3102, 3111, 3115). This entitlement flows from federal civil rights statutes and case law (Linquanti & Cook, 2013; Hakuta, 2011). However, methods of identifying and initially classifying ELs have varied among states and even among districts within states. A key challenge facing states that are part of content or ELP assessment consortia is establishing a consistent set of policies and practices to define ELs "in a manner that is uniform across member states and consistent with section 9101 (25)"(USED, 2010, p. 20). Creating a uniform set of policies, practices, and methods for initial EL classification is extremely challenging across states, and even within states that permit local control.

A National Research Council panel convened to examine the comparability of EL definitional processes across states concluded with respect to initial EL classification that

> Because of the differing state policies, practices, and criteria for initially identifying students as linguistic minority and for classifying them as an English language learner (ELL), individuals who are classified as ELL students in one state may not be classified as ELL students in another. In states that permit local control, students classified as ELL in one district may not be classified as ELL in another district in that state. (NRC, 2011, p.86)

As indicated in these NRC report findings and detailed in the above-referenced CCSSO guidance, students are classified as EL in two closely-related stages: Students must first be identified as potential EL ("linguistic minority") and then initially assessed in their English language proficiency to determine if they are to be classified as EL. An earlier CCSSO report in this series (Linquanti & Bailey, 2014) summarizes a national working session addressing issues and options related to identifying potential ELs through home language surveys. Strengthening the design and use of home language surveys can help to more precisely identify the population of potential ELs that require initial ELP assessment. This paper focuses on guidance for a rigorous initial classification process to confirm (or disconfirm) EL status of potential ELs and shares recommendations for strengthening EL classification policies, practices, and tools.

Issues in Current EL Classification Policies

Table 1 lists initial EL classification assessments (by number and type) and authority for defining the EL classification decision process across the 50 states and District of Columbia. As seen in Table 1, there is great variability within and across states in how ELs are initially classified. For example, seven states permit LEAs to choose from among multiple initial ELP assessment instruments (labeled "Multiple Commercial Screeners" in the table).[21] Such instruments often conceptualize the ELP construct differently, are normed on different groups of students, and weight language domains (listening, speaking, reading, and writing) differently in deriving an overall ELP composite score. Even within the WIDA Consortium, a long-established state consortium that shares common ELP standards and uses a common ELP annual

---

[21]   An initial ELP assessment may be relatively brief (often referred to as a "screener"), or may be the state's annual ELP assessment used to fulfill ESEA Title III assessment requirements.

summative assessment, half of its member states allow LEAs to choose between a brief consortium screener assessment and an expanded screener assessment commercially available to the consortium for initial classification assessment purposes.[22] In addition to the variability in screening instruments, local educational agencies (LEAs) in 39 states do not have state initial EL classification guidelines and are solely responsible for determining their initial EL classification process. The remaining 12 states and the District of Columbia provide state-defined parameters for LEAs to follow in their initial EL classification decision-making, typically via guidelines, checklists, or specified procedures. With such variation in instrumentation and process, establishing comparability in how ELs are initially classified across states, and often across districts within states, is very difficult. The scant empirical research that is available on these issues[23] suggests that such variations can substantially affect the EL populations defined, rendering comparisons made under current conditions of questionable meaning and value. There is also emerging evidence that inconsistent classification practices leading to misclassifications can occur even within a given district or school (Okhremtchouk, 2014) and that misclassifying students, especially those very near the cutpoint for initial fluent English proficiency, could lead to long-term negative academic outcomes (Umansky, 2014).

Table 1. State EL Classification Assessments and Locus of Authority (as of May 2014, includes District of Columbia)[24]

| ELP Assessment Used for Initial EL Classification (Type and Number) | Authority for Defining EL Classification Process | | Total |
	LEA Alone	LEA, with SEA Guidelines	
State Summative ELP Test	2	3	5
Single State-developed Screener	1	0	1
Single Commercial Screener	5	1	6
Multiple Commercial Screeners	6	1	7
Single Consortium Screener	14	2	16
Multiple Consortium Screeners	11	5	16
Total	39	12	51

CCSSO National Working Session on Initial Classification of English Learners

In order to foster a more common understanding of key issues and develop guidance for strengthening policies and practices related to initial EL classification, CCSSO convened state and consortia ELP assessment representatives, district EL experts, and EL researchers for structured, facilitated discussions. Group discussions focused on the following three areas:

1. Guidelines for initial EL classification;
2. Strategies to address EL misclassification; and
3. Approaches to support comparability of initial EL classification criteria and procedures both within and across states and consortia.

---

[22]   The W-APT and MODEL, respectively. States with this choice are categorized in the table as having "Multiple Consortium Screeners".
[23]   See, for example, Abedi, 2008; Carroll & Bailey, 2013; and Linquanti, 2008.
[24]   Complete data for 50 states and District of Columbia is displayed in table in Appendix B.

The goal of these discussions was to identify common issues; share promising policies, tools, or practices; and suggest meaningful guidelines for procedures that could be used by schools, districts, and states to more consistently classify ELs. Working session participants described a wide variability in how ELs are classified across schools, districts, and states, and expressed support for implementing policies and practices that foster greater standardization and consistency. The next section distills the group's conversations and suggested guidelines related to the three areas listed above, which the authors have further elaborated and developed for consideration by state and consortium stakeholders and policymakers.

Guidelines for Initial EL Classification

1. *States and districts should provide common guidance on purposes, policies, and practices related to the initial EL classification process.*

Participants agreed that while EL classification is a locally administered process and a local decision, its rigor and consistency within a state is critical and will help to strengthen comparability across states within a consortium. Common guidance could include:

  a. Clear statement of purpose (confirming or disconfirming the status of potential ELs, and providing appropriate specialized language support services to students classified ELs, as entitled under federal law);

  b. Flowchart and checklist of the classification process, including clear procedures for detecting and correcting misclassifications (see below);

  c. Timeframe specifying when students identified as potential ELs (via the home language survey, or HLS) are to be initially assessed ("screened") for English language proficiency and, if needed, placed in a program offering appropriate language instructional services;

  d. Initial classification assessment instrument(s) allowed, including guidelines for scoring and interpreting scores;

  e. Evidence standards and protocols that can be used by educators to detect and evaluate suspected misclassifications and to document and correct confirmed misclassifications; and

  f. Qualifications and training required of those carrying out the initial classification process.

2. *States and districts should ensure the initial EL classification process is appropriately and consistently implemented.*

Appropriate and consistent implementation of the initial classification process involves at least the following:

  a. Identify individuals with sufficient preparation and experience working with language minority students, whose job responsibilities will include conducting the initial EL classification process;

b. Provide training on classification instruments, procedures and practices, that is designed to support a variety of educational contexts (e.g., schools and districts with small vs. large numbers of ELs), delivered by qualified personnel, and available as needed (e.g., annually as well as on-demand);

c. Collect and electronically store HLS responses and determinations (potential EL / not potential EL), classification instrument score results for potential ELs, classification determinations, and any misclassification corrections made; and

d. Audit and/or monitor the initial EL classification process (at school and district levels) to assure integrity of the process and to identify training or resource needs that schools or districts may have.

Strategies to Address Initial EL Misclassification

Throughout the working session, participants identified and discussed strategies to prevent or address EL misclassification. Participants focused particularly on students who receive initial ELP assessment scores at or near the English-proficient cutpoint and on language minority students who have a subsequently recognized language-related learning disability. The following recommendations focus on a provisional classification period and on validation or standardization procedures.

3. *States and districts should consider establishing a "provisional classification" period for students initially classified as EL, in order to allow for correction of misclassification errors.*

Any assessment instrument, no matter how reliable, has measurement error, particularly near an operative cut point. In initially assessing potential ELs, this can lead to classifying students as English learners who are in fact English fluent, or vice versa. One district representative estimated her district's EL misclassification rate to be between one and five percent. Many participants noted that there is no formal procedure in their state or district to correct misclassifications, which contributes to a strong reluctance to acknowledge and address this issue. Participants also expressed concern about language minority students with a language-related learning disability being mistakenly classified as EL rather than as an EL with a disability or even English-fluent student with a disability.[25]

Given the high-stakes consequences of classification decisions made from the initial ELP assessment, participants voiced support for a procedure and a limited time frame to detect and correct erroneous classifications (either as EL or as initially fluent English proficient). This "provisional period" would allow educators to provide services appropriate to the students' initial classification, observe students' language use in the classroom, and detect and address any confirmed misclassifications. The provisional classification time frame should be long enough to detect and correct misclassifications but short enough to avoid being educationally disruptive to the student.

A range of time frames from 45 to 90 days was proposed, and no specific timeline is recommended here. However, participants suggested that the current federally legislated time frame for notifying parents of a student's EL classification and placement (i.e., within 30 days of the beginning of the school year, per NCLB, s.3302.a) was

---

[25] It is important to note that EL students are generally underrepresented in special education in the early elementary grades, but tend to be overrepresented in special education at the upper elementary and secondary levels. See, for example, Artiles et al., (2005), and Samson & Lesaux (2009).

insufficient for detecting and correcting misclassifications. Some noted there may be far fewer than 30 days available for this latter purpose, depending upon when students are initially registered, determined through home language survey results to be potential ELs, and administered the initial classification assessment. They also noted that students initially enrolled *during* the school year must be identified as potential ELs, properly classified through initial assessment, and placed in the appropriate instructional program setting within two weeks of initial school entry (per NCLB, s.3302.d).

Figure 1 depicts one model for enacting a provisional initial EL classification period. Students are identified as potential English learners during the "pre-classification" stage (left side of the figure, analogous to the "identify potential EL" stage in the Linquanti & Cook framework). Once students are identified as potential English learners, they are given the initial ELP assessment. If students score below the proficient cut point on this initial assessment, they are classified as English learners and placed into an appropriate language instruction educational program. Students obtaining scores above the proficient cut point are classified as initially fluent English proficient students and do not receive specialized language development services.

Research has documented that misclassification errors occur in both the pre-classification as well as initial classification stages.[26] In fact, the issue has been recently reported in the popular press.[27] During a designated time frame (e.g., 45 days), classification errors can be detected, changes made to a student's classification status, and instructional services adjusted accordingly.

Figure 1: Flowchart depicting potential EL identification and provisional initial classification

---
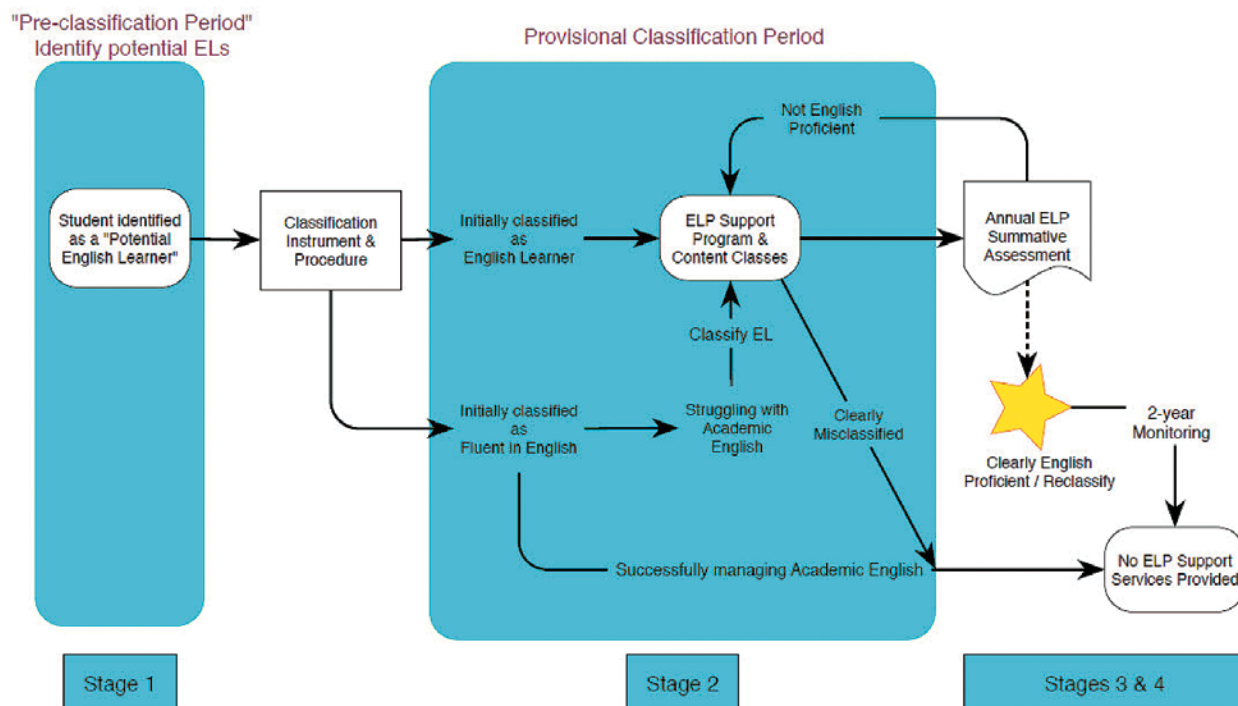
[26]   See Abedi, 2008; Goldenberg & Rutherford-Quach, 2010; Stokes-Guinan & Goldenberg, 2011; Carroll & Bailey, 2013; and California Department of Education, 2011.
[27]   See Taxin, A. (2014).

During this provisional period, educators would be allowed (indeed, expected) to identify and gather evidence about the small number of students they believe have been misclassified.[28] Based on working session discussions, we specify six possible initial classifications that in principle could be confirmed or corrected during the provisional classification period. These classifications, depicted in Table 2, reflect permutations of students' language classification (English learner, Initially English fluent/native bilingual speaker, or monolingual English speaker), and disability status (special-education-identified, or not). The provisional classification period could be used to help determine whether students in any of these categories are more appropriately classified in any of the others.

Table 2. Permutations of language classification and special education status

Language Classification	Not Special Ed-identified	Special Ed-identified
English Learner	I	II
Non-EL linguistic-minority (Initially English fluent / native bilingual)	III	IV
Monolingual English ("English Only")	V	VI

Researchers and practitioners have documented the complexities of, and offered procedures for disentangling, second language development features from language-related learning disabilities in English learners.[29] Clearly, language related learning disabilities could manifest at any age/grade, so it is quite possible that an EL student may be referred for evaluation and identified for special education services long after the provisional period.[30] While the provisional classification period cannot possibly serve to address every potential misclassification, it can provide a recognized time frame and process for detecting and addressing some of the many ways in which students might be misclassified in each of the six classifications designated in Table 2.

For example, working session participants offered three scenarios from their experiences that this process could help to address:

a) A student initially classified as fluent English proficient (Table 2, Cell III) might be found to struggle in comprehending and using English for learning academic content or engaging in content practices in the classroom. Through a specified protocol that includes consultation with administrators, parents, and a teacher or a team of educators, a student might be reassessed and determined from evidence to be more appropriately classified as EL (Table 2, Cell I) and provided with language instructional services and carefully scaffolded content instruction she is entitled to receive.

b) A student initially classified as an EL (Table 2, Cell I) exhibits language use in grade-level content practices in the classroom that clearly suggests she has been misclassified. After formal review of evidence of the student's language use through a protocol involving a teacher, parent, and

---

[28]   A larger proportion of students appearing to be misclassified would trigger a system-level review to detect large-scale problems with initial EL classification criteria, procedures or tools.

[29]   See for example Artiles et al. (2005); Ortiz et al. (2011); and Klingner et al. (2008). For an example of a district-developed, comprehensive evaluation and special-education referral process for English learners, see Gaviria & Tipton, 2012.

[30]   Indeed, determining exactly *when* EL students are identified for special education services is important in exploring why "long-term ELs" are diagnosed with language-related disabilities at substantially higher rates relative to the general population. See, for example, Artiles et al., (2005), and Samson & Lesaux (2009).

administrator, educators may conclude the student should be classified as initially fluent English proficient (Table 2, Cell III) and may discontinue the supplemental services designated for EL students.

c) An initially classified EL student (Table 2, Cell I) is subsequently diagnosed within the 30-45 day provisional classification period as having a language-related learning disability. After discussion with parents, and evaluation by a team with sufficient expertise, the student is reassessed and determined to be a language minority student (e.g., initially English fluent) whose language-related disability was mistaken as insufficient English language proficiency. The educator team designs a program of specialized instructional services that addresses the identified language-related disability but concludes that an English language instructional program is not needed, and the student is not classified EL but rather as an initially English-fluent student with a disability (Table 2, Cell IV). However, another initially classified EL student (Table 2, Cell I) is discovered to also have a language-related learning disability and so qualifies for services related both to their English language proficiency and their learning disability (Table 2, Cell II).

The provisional classification period as described here, which might exceed 30 days, does *not* imply that instructional services and parental notifications would be delayed until after this period concluded. Instructional services based on initial classification would be provided immediately, and parental notifications would occur within federally required time frames. Rather, the provisional classification period provides an opportunity for educators to correct misclassification errors such as those described above. And, as noted, any correction and associated change in instructional services should also involve the parent directly.

4. *States and districts should differentiate initial EL classification procedures and tools for early elementary students (Kindergarten and grade 1) from later-grade students, yet take steps to ensure optimal classification decisions.*

Several states' initial EL classification procedures for students in early grades (i.e., Kindergarten and grade 1) either assess only oral language domains (i.e., speaking and listening) or assess all four language domains (oral language and literacy skills) but significantly underweight literacy skills (e.g., overall proficiency composite score is 90% oral, 10% literacy). This approach reflects a reasonable concern that literacy skills in young children may not reflect English language proficiency so much as prior literacy experiences (e.g., at home or in preschool) and may lead to greater EL misclassifications.

However, language minority students classified as initially fluent English proficient based solely or largely on oral language skills may experience difficulty in developing foundational literacy skills.[31] This in turn may influence these students' later literacy development. Determining the appropriate level of assessment of English literacy skills in young children is a tension point and should be approached carefully to optimize initial classification decisions. This will involve ongoing monitoring and validation of classification procedures and decisions for early elementary language-minority students in particular.

---

[31] This concerns whether high oral-language scores on ELP assessments are sufficiently predictive of foundational literacy skills for students who are potential ELs. The degree to which these scores are not predictive contributes to student misclassification and potentially to not providing language instructional services to students that warrant and are entitled to them.

5. *The initial classification assessment used to screen potential ELs should be aligned to the state's ELP standards and strongly predict performance on the state's annual ELP summative assessment.*

It is important to ensure a strong relationship in both measured constructs and outcomes between the initial classification assessment and the annual ELP assessment. A two-step process is proposed for consideration. First, the state can conduct an alignment study between the state's/district's classification instrument and the state's English language proficiency standards.[32] For states or districts using relatively brief initial "screener" instruments, the alignment will not be as extensive as with the annual summative assessment. This is to be expected given that screeners are shorter tests. Nonetheless, the alignment study's findings on match, depth and breadth of coverage of the initial classification screener to the state ELP standards should be proportionately comparable to that of the annual summative ELP assessment. Alternatively, the state might conduct a study comparing the test blueprints and test maps of the classification screener and annual summative ELP assessment. The goal of this comparison would be analogous to an alignment study (i.e., to determine sufficient proportionality in coverage and design).

Second, a formal study can be undertaken to determine the predictive relationship between classifications based on the initial classification assessment and those of the annual summative assessment. There should be evidence of a strong relationship. This relationship could be examined through correlation and/or regression analyses. For example, the study could administer the initial classification assessment to already confirmed ELs and within a short period of time (e.g., a few days) give the annual summative ELP assessment to the same students.[33] There should be strong relationships between initial and annual test scores and proficiency categorizations for these students. Moderate to low associations (e.g., correlations) would suggest that the assessments are measuring different aspects of English language proficiency and are not yielding comparable classification decisions.

6. *States and districts should examine the validity of instruments and procedures used in the initial EL classification process.*

States and districts should regularly examine the frequency, types, and causes of misclassification errors. It is especially important to monitor the classification accuracy of migrant students, students with interrupted formal education, Native American and Alaska Native students, and students with disabilities since these students may be more prone to experiencing classification errors. Careful review of misclassifications may lead to changes in the initial classification instrument's interpretation and use, and/or in classification procedures.

7. *States and districts should identify and share policies, practices, and tools demonstrated to reduce initial EL misclassifications.*

Virtually all participants expressed concern about the inconsistency in how schools and districts initially classify ELs. This inconsistency has the potential to increase misclassifications, which occur for a variety of reasons. Three commonly discussed reasons for inconsistencies leading to misclassification were lack of guidance, lack of training, and lack of expertise. Participants suggested several possible strategies to

---

[32]   For one example of such an alignment study, see Cook, 2005.
[33]   The timeframe should be short enough that the student's English language proficiency should not be measurably better as a result of instructional services provided during that timeframe.

reduce misclassifications, which could be utilized within the provisional classification period discussed above. Examples include:

a) District-wide experts train school staffs in the initial classification process using a range of scenarios/cases of potential EL students in order to develop understanding and to help resolve challenging classification decisions;

b) Educators engage in informal interviews with parents regarding those students with more complex "linguistic histories" as captured on the HLS prior to and/or after the initial classification instrument has been administered in order to better understand observed performance on the classification assessment;

c) A school or district-based EL classification committee, staffed with appropriately qualified and trained educators, reviews potential misclassifications and appropriately involves parents in deliberations and decisions; and

d) Educators with appropriate EL expertise examine ambiguous EL classification results and support classification decision-making in consultation with parents.

Approaches to support comparability of initial EL classification criteria and procedures within a state and across states within an ELP consortium

Participants overwhelmingly supported efforts to increase comparability of initial EL classification criteria and procedures across districts, states, and consortia, whether they were from a stand-alone or an ELP consortium state. Many participants from ELP consortium states expressed the desire for each consortium to provide leadership in this area by offering recommended policies, practices, and tools that member states might adopt. The following recommendations address these expectations.

8. *Districts within a state, and states within an ELP consortium should use a single ELP assessment to support initial EL classification. If a state allows multiple ELP assessment instruments to be used, it should first provide compelling empirical evidence that these instruments' scores are comparable.*

Participants from states allowing multiple initial classification assessments described challenges in understanding transferring ELs' English proficiency levels from districts not using the same instrument. One way to better understand a student's initial English language proficiency in states using multiple classification instruments is to provide a table of comparative scores. Practically all participants agreed that having a common initial classification assessment was preferred. Multiple classification assessments can lead to multiple understandings of English language proficiency. This is especially relevant given that many commercially available ELP assessments (commonly used for initial EL classification) measure and weight different domains of language proficiency differently.

9. *ELP assessment consortia should provide guidance on recommended instrument(s), data collection and analysis protocols, and administrative policies and procedures used to support initial EL classification across member states.*

While this recommendation is similar to those made above for individual states, it calls for establishing sufficient commonality to allow for aggregation across consortium member states.

10. ELP assessment consortia should conduct studies of instruments, procedures, and practices among member states to assure comparability of initial EL classification outcomes.

At a minimum, studies exploring the comparability of initial EL classification instruments, procedures, and practices across consortium member states should include:

a. a matrix of comparable scores across initial EL classification instruments, and

b. a description of required or recommended initial classification procedures and practices.

Conclusion

Strengthening the stage of confirming (or disconfirming) initial classification of potential ELs as English learner or fluent English proficient is both complex and crucial in moving toward a more common EL definition within and across states. Working session participants stressed the need for both standardization and local flexibility. Many of the above recommendations focus on standardization, yet EL classification is ultimately a local practice. Participants saw the need for more thoughtful, transparent, inclusive, and documented processes to be in place to appropriately classify ELs and to allow for timely correction of misclassifications. They also noted that a strengthened initial EL classification process could be successfully implemented only with proper and sustained training and monitoring. If districts, states, and consortia pursue the above recommendations, the resulting initial EL classification process would very likely lead to more accurate classifications and therefore contribute to a more common definition of English learner within and across states and consortia.

References

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

Artiles, A. J., Rueda,R., Salazar, J. J., & Higareda, I. (2005). Within-group diversity and minority disproportionate representation: English-language learners in urban school districts. *Exceptional Children*, *71*(3), 283-300.

California Department of Education. (2011). *Comparison Study of Kindergarten and Grade 1 English-Fluent Students and English Learners on the 2010–11 Edition of the CELDT.* Sacramento: Author.

Carroll, P.E., & Bailey, A.L. (2013, April). *Combining multiple indicators in classifications of English language proficiency: A descriptive study.* Presented at the National Council on Measurement in Education Annual Meeting, San Francisco, CA.

Cook, H. G. (2005). *Aligning English language proficiency tests to English language learning*

standards (Report 5). Aligning assessment to guide the learning of all students: Six reports on the development, refinement, and dissemination of the web alignment tool. Washington, DC: The Council of Chief State School Officers.

Gaviria, A., & Tipton, T. (2012). *CEP-EL: A comprehensive evaluation process for English learners. A process manual.* San Diego: San Diego Unified School District.

Goldenberg, C., & Rutherford-Quach, S. (2010). *The Arizona home language survey. The identification of students for ELL services.* Los Angeles: University of California Los Angeles Civil Rights Project.

Hakuta, K. (2011). Educating language minority students and affirming their equal rights:

Research and practical perspectives. *Educational Researcher, 40*(4), 163-174.

Klingner, J., Hoover, J., & Baca, L. (2008). *Why do English language learners struggle with reading?: Distinguishing language acquisition from learning disabilities.* Thousand Oaks, CA: Corwin Press.

Linquanti, R. (2008). Assessing language proficiency of California's English learners and what it means for accountability. Invited essay. *University of California Linguistic Minority Research Institute Newsletter, 17* (2), 1-3.

Linquanti, R., & Bailey, A.L. (2014). *Reprising the home language survey: Summary of a national working session on policies, practices, and tools for identifying potential English learners.* Washington DC: Council of Chief State School Officers.

Linquanti, R., & Cook, G. (2013). *Toward a "common definition of English learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options.* Washington DC: Council of Chief State School Officers.

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.

Okhremtchouk, I.S. (2014). Classifying Language-Minority Students: A Closer Look at Individual Student Data. *Bilingual Research Journal: The Journal of the National Association for Bilingual Education*, 37:3, 327-348.

Ortiz, A. A., Robertson, P. M., Wilkinson, C. Y., Liu, Y.-J., McGhee, B. D., & Kushner, M. I. (2011). The Role of Bilingual Education Teachers in Preventing Inappropriate Referrals of ELLs to Special Education: Implications for Response to Intervention. *Bilingual Research Journal*, *34*(3), 316–333.

Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *Journal of Learning Disabilities*, *44*(2), 148-162.

Stokes-Guinan, K., & Goldenberg, C. (2011). Use with caution: What CELDT results can and cannot tell us. *The CATESOL Journal 22*(1), 189-202.

Taxin, A. (2014). Parents lie on survey to identify English learners. Associated Press, November 16, 2014. Retrieved from: http://news.yahoo.com/parents-lie-survey-identify-english-learners-163108335.html.

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

Thompson, K. D. (in press). Questioning the long-term English learner label: How categorization can blind us to students' abilities. *Teachers College Record.*

Umansky, I. (2014). *Peeling back the label: Does English learner classification help or hurt students learning English?* Stanford University Center for Education Policy Analysis Working Paper.

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.

Appendix A.
Working Session Participants/Observers and Organizations Represented

Name	Organization
Fen Chou	CCSSO
H. Gary Cook	WCER, Facilitator
Robert Linquanti	WestEd, Facilitator
Jennifer Timm	Carnegie Corporation of New York
Kenji Hakuta	Stanford/ELPA21 (Facilitator)
Edith Palmberg	ELPA21/Olathe, Kansas
Dorry Kenyon	Center for Applied Linguistics/WIDA
Martha Thurlow	NCEO
Amy Park	California Department of Education
María Trejo	Texas/Cypress-Fairbanks
Angelica Infante	New York State Education Department
Kerri Whipple	WIDA/ South East Education Cooperative, North Dakota
Margaret Ho	ELPA21/Washington OSPI
Martha Martinez	ELPA21/Oregon Department of Education
Audrey Lesondak	WIDA/Wisconsin DPI
Nicole Sayegh	California/Tahoe-Truckee USD
Arlene Costello	Florida/Escambia SD
Maria Angelica Meyer	NYS/Westbury SD
Bill Auty	ELPA21
Julie DeCook	Wisconsin/Janesville SD
Debra Dougherty	California/San Diego USD
Laurie Shaw	Texas/Pflugerville ISD
Chane Eplin	Florida Department of Education
Judy Diaz	NYS/Port Chester SD
Jonathan Gibson	WIDA/Nevada Department of Education

Working Session Observers

Supreet Anand	OESE/USED
Carlos Martinez	OELA/USED

Appendix B.
State EL Classification Assessments and Locus of Authority
(as of May 2014, includes DC)

State	Initial EL Classification Assessment Type	Assessment Name(s)	Initial English proficient performance criterion specified by SEA	Was guidance on the process of initially classifying ELs identified on SEA website?
AL	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 4.8) & district policy of monitoring	No
AK	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	Yes
AZ	Annual Summative	AZELLA	State criteria: "Proficient" on AZELLA	Yes
AR	Commercial Screener	LAS, MACII, ELDA	State criteria (Overall Composite proficiency level 5.0) & district policy	No
CA	Annual Summative	CELDT	Overall performance level is Early Advanced or higher, and Domain scores for Listening and Speaking are Intermediate level or higher	Yes
CO	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
CT	Annual Summative	LAS Links	State criteria: "4" or "5" on LAS LINKS	Yes
DE	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	No
DC	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 4.9)	Yes
FL	Annual Summative	CELLA	State criteria: "Proficient" on CELLA	Yes
GA	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	Yes
HI	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No

State	Initial EL Classification Assessment Type	Assessment Name(s)	Initial English proficient performance criterion specified by SEA	Was guidance on the process of initially classifying ELs identified on SEA website?
ID	Commercial Screener	IELA	State criteria: Score at the Early Fluent (4) or Fluent (5) Level and obtain an (EF+) on each domain	No
IL	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/ Literacy Composite of 4.2 or greater)	Yes
IN	Commercial Screener	LAS Links	State criteria: "Fluent English Proficient" on LAS Links Placement Assessment	No
IA	Commercial Screener	LAS Links	State criteria: "Fluent English Proficient" on LAS Links Placement Assessment	No
KS	Commercial Screener	KELPA-P; LAS LINKS; LPTS; IPT	State criteria of "Fluent"	No
KY	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
LA	Annual Summative	ELDA-P	State criteria of Level "5" in all 4 domains	No
ME	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 6.0)	No
MD	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
MA	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0/4.0 or greater on reading and writing domains)	No
MI	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	Yes
MN	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0/4.0 or greater on all domains	No
MS	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
MO	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
MT	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/ Literacy Composite of 4.0 or greater)	No

State	Initial EL Classification Assessment Type	Assessment Name(s)	Initial English proficient performance criterion specified by SEA	Was guidance on the process of initially classifying ELs identified on SEA website?
NE	Commercial Screener	ELDA-P; LAS LINKS; IPT; TELPA; Woodcock-Munoz	State criteria of "4=Advanced Proficient" or higher	No
NV	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0/Literacy Composite of 5.0 or greater)	No
NH	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	No
NJ	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 4.5)	Yes
NM	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	Yes
NY	Commercial Screener	NYSITELL	State criteria "Proficient"	No
NC	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
ND	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/all domains 3.5 or greater)	No
OH	Commercial Screener	BINL; BOLT; BSM; BVAT; IPT; LAB; LAS LINKS; MAC II; SLEP; Woodcock-Munoz	State criteria (Overall Composite proficiency level 5.0)	No
OK	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/Literacy Composite of 4.5 or greater)	Yes
OR	Commercial Screener	IPT; LAS LINKS; Stanford ELPT; Woodcock-Munoz	State criteria (Overall Composite proficiency level 5.0)	No
PA	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 4.6)	No

State	Initial EL Classification Assessment Type	Assessment Name(s)	Initial English proficient performance criterion specified by SEA	Was guidance on the process of initially classifying ELs identified on SEA website?
RI	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/Literacy Composite of 4.5 or greater)	No
SC	Commercial Screener	IPT; LAS LINKS; ELDA-P; Woodcock-Munoz	State criteria (Overall Composite proficiency level 5.0)	Yes
SD	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	No
TN	Commercial Screener	TELPA	State criteria: "3" or above on TELPA; "5" on English Language Development Assessment (ELDA)	No
TX	Commercial Screener	IPT; LAS LINKS; Stanford ELPT; Woodcock-Munoz; CELLA; TELPAS	State criteria: "4" or above on Woodcock-Munoz and LAS LINKS; "5" on Stanford ELPT; "Advanced High" on TELPAS; Above "E" and "F" on IPT; "Proficient" or above on CELLA	No
UT	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 5.0)	No
VT	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/Literacy Composite of 4.0 or greater)	No
VA	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0/Literacy Composite of 5.0 or greater)	No
WA	Commercial Screener	WEPLA-P	State criteria: "4"	No
WV	Commercial Screener	Woodcock-Munoz-Language Survey	State criteria: "5"	Yes
WI	Consortium Screener	W-APT	State criteria (Overall Composite proficiency level 6.0)	No
WY	Consortium Screener	W-APT or MODEL	State criteria (Overall Composite proficiency level 5.0)	No

STAGE 3

Reference Performance Level Descriptors: Outcome of a National Working Session on Defining an "English Proficient" Performance Standard

H. Gary Cook
Rita MacDonald
Wisconsin Center for Education Research
2014

Overview:

This document elaborates on the third of four key stages in moving toward a common definition of English learner (EL), as described in the CCSSO publication, *Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options* (Linquanti & Cook, 2013).

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/ consortium English language proficiency (ELP) test against which to assess ELs' English proficiency; and

4. **Reclassify** a student to former-EL status through the use of multiple exit criteria.

It explores a critical facet of the third stage of the framework and presents a set of Reference Performance Level Descriptors to support comparison of ELs' performances across different ELP assessments. It was developed from guidance provided at a national working session held on September 17, 2013 at the CCSSO's Washington, DC offices.

Background

The CCSSO guidance posits that a key step needed to define an "English proficient" performance standard is the development of common (or comparable) performance level descriptors (PLDs) of English language proficiency (section 3.a, p. 12). The Reference PLDs presented below were developed to support the enactment of this aspect of the guidance, as the two academic assessment consortia expressed urgent interest in this topic, given the consortium assessment field testing to be held in Spring 2014. An initial draft of the Reference PLDs was shared and discussed extensively at a national working session of state and consortium representatives, experts on English language proficiency (ELP) and ELP assessment, and other stakeholders[34] held on September 17, 2013 at the Washington, D.C. offices of CCSSO. Following this review and feedback on the initial draft, the following set of Reference PLDs was developed.

What the Reference PLDs Are

Rather than attempt to develop a single, common set of ELP performance level descriptors to be used by all states and consortia, the participants advised the development of a reference tool to *enable the comparison of* state/consortium-specific ELP levels when such comparisons are needed. The Reference PLDs were created by carefully examining and amalgamating PLDs available from several states (e.g., Arizona, California, and Texas) and consortia (i.e., ELPA21, WIDA/ASSETS). They represent common elements of ELP performance outlined across all of the aforementioned states and consortia. The Reference PLDs are designed to serve as a common reference point for "translating" English proficiency level categorizations (e.g., the proficiency level information from a state/consortium with 5 levels to that of a state/consortium with 3 levels). This supports a common understanding of the designations of low, moderate, or high proficiency in English across these state educational agencies and consortia. Such comparisons or translations must and will be made by assessment consortia that include states with differing systems of ELP classification; the Reference PLDs can support these comparisons by providing a common lens.

What the Reference PLDs Are Not

1. The Reference PLDs are not intended to replace any state's or consortium's existing English language proficiency levels or PLDs.

2. They are not intended to revise or improve upon any existing English language proficiency PLDs or classifications.

3. They are not standards, nor are they designed to revise or improve upon current English language development standards.

4. The Reference PLDs have not been validated for application to individual English learners; thus, they are not meant to be used to classify the performance of an individual student moving across states or consortia, a purpose for which ELP screeners are intended and better suited.

---

[34]   Participants and the institutions they represented are listed in Appendix A.

Composition of the Reference PLDs

<u>Source of items:</u> All proficiency level descriptors on the Reference PLD chart were drawn directly or adapted from the English language proficiency PLDs of Arizona, California, Texas, the ELPA 21 Consortium, and the WIDA Consortium. Individual elements from these state or consortia PLDs do not appear as originally written; revision was undertaken to facilitate coherence and consistency of the Reference PLDs. The overall goal was to identify common language elements found across these state and consortia PLDs, and incorporate them into the Reference PLDs.

<u>Organization:</u> The Reference PLDs are organized in three ways:

1. *Proficiency levels.* A review of the above-mentioned state and consortium documents revealed a range in the number of proficiency levels from 3 to 6. Working session participants suggested adopting 3 proficiency levels, that number being the "least common denominator" to which classification systems could be reduced. These levels are labeled in the Reference PLDs as "low," "moderate," and "high." Level names were chosen to assure no match to existing state or consortia PLDs.

2. *Language features.* The Reference PLDs identify the quantity and characteristics of *words, sentences, and discourse features* that English learners exhibit at the three proficiency levels.

3. *Modalities and Skills.* The Reference PLDs are arrayed in a 2 x 2 matrix comprised of oral and text-based modalities and productive and receptive skills. Within this matrix, the common domain labels of speaking, listening, reading, and writing can be identified as watermarks.

<u>Functional Language:</u> Several examples of functional language use across the three levels of English language proficiency are provided. These examples are not intended to provide a comprehensive view of functional language use. They merely provide contextualized samples of language used for academic purposes across several example disciplines.

Dimensions of Change in Language Proficiency

Given that performance level descriptor examples and illustrations of functional language use are illustrative, the following descriptions of language proficiency development may be useful in deciding how state or consortium PLDs map onto the Reference PLDs.

As conceptualized by Bailey and Heritage (2013), language proficiency increases along a predictable set of dimensions, each of which represent a student's ability to construct and convey meaning through language:

1. *Increase in amount:* the number or sophistication of words or ways of combining words (phrases, clauses)

2. *Increase in repertoire of use:* the types of relationships students can construct between ideas – e.g., additive, causal, conditional, contrastive – as well as the number of ways students are able to construct those relationships between ideas

3. *Increase in accuracy:* the students' ability to construct precise meanings

4. *Increase in contextualization:* the students' ability to tailor the use of language functions to fit a variety of sociocultural contexts

Additionally, an examination of state and consortium PLDs reveals a fifth dimension:

5. *Increase in autonomy:* the students' autonomy with the language, which is observed by the need for fewer language supports and scaffolds as proficiency increases

Next Steps

States or consortia can map their specific PLDs onto the Reference PLDs. This locally determined comparison will then be available for use as needed.

Reference PLDs

The Reference PLDs are organized by proficiency levels, language features, and modalities and skills. The display on the following page presents these in table format. The columns show skills (productive and receptive) and levels (low, moderate, and high). The rows show modalities (text-based and oral) and language features (discourse, phrase/sentence & word). At the nexus of skills and modalities, we can identify the four language domains: speaking (productive-oral), listening (receptive-oral), writing (productive-text-based), and reading (receptive-text-based), which are indicated by watermarks in each quadrant.

		RANGE OF PERFORMANCE IN ENGLISH – descriptors reflect performance at exit stage of each level					
		PRODUCTIVE			RECEPTIVE		
		Low	Moderate	High	Low	Moderate	High
Oral	**Discourse**	Engages in basic oral interactions in direct informational exchanges on familiar and routine social and academic topics	Engages comfortably in most social and academic discussions on familiar topics using extended discourse	Produces, initiates, and engages in sustained extended interactions tailored to specific purposes and audiences on a variety of social and academic topics, including new and unfamiliar topics	Understands simple or routine directions and short, simple conversations and discussions on familiar social and academic topics	Comprehends most social and academic discussions on familiar topics and follows discussions related to feelings, needs and opinions in extended discourse	Comprehends longer, elaborated directions, and extended conversations and discussions on familiar and unfamiliar topics in academic and social contexts
	Phrase/Sentence	Uses repetitive phrasal and sentence patterns with formulaic structures common across discipline areas	Uses a variety of sentence structures with varying complexity	Uses a variety of sentence structures with varied levels of complexity tailored to the communicative task	Demonstrates an understanding of repeated phrases and simply-constructed sentences	Comprehends a variety of grammatical constructions and sentence patterns common in spoken language in academic and social contexts	Comprehends a wide variety of complex and sophisticated sentence structures in varied academic and social interactions
	Word	Uses commonly used words and phrases	Uses specific and some technical content-area vocabulary and words or phrases with shades of meaning	Uses a range of abstract, specific and technical content-related vocabulary; uses a range of idiomatic expressions and words or phrases with multiple meanings	Demonstrates an understanding of words and phrases from familiar contexts and previously learned content material	Understands specific and some technical content-related vocabulary; some idiomatic expressions and words or phrases with multiple meanings	Understands a wide range of specific, technical and idiomatic words and phrases; comprehends words and phrases with multiple meanings
Text-Based	**Disclosure**	Produces basic written texts in directed tasks or activities on familiar and routine topics	Produces texts that express ideas to meet most social and academic needs	Produces texts to meet a variety of social needs and academic demands for specific purposes and audiences	Demonstrates an understanding of simple sentences in short, connected texts with visual cues, on familiar topics	Demonstrates comprehension of increasingly complex texts; identifies detailed information on unfamiliar topics with fewer contextual clues	Demonstrates comprehension of a variety of complex texts and identifies general and detailed information in texts on familiar and unfamiliar topics
	Phrase/Sentence	Produces simple sentences	Produces texts that reflect a grasp of basic grammatical structures and sentence patterns with evidence of emerging use of more complex patterns	Produces texts using a variety of grammatical structures and a broad range of sentence patterns matched to purpose	Demonstrates an understanding of basic, routinely used language structures in social and content-area texts	Demonstrates comprehension of a variety of complex grammatical constructions and sentence patterns in social and content-area texts	Demonstrates comprehension of a wide variety of complex and sophisticated sentence structures from varied social and content-area texts
	Word	Uses high frequency and commonly-learned vocabulary and phrases drawn from social contexts and content areas	Uses more varied vocabulary that extends beyond the everyday to include content-specific vocabulary, some idiomatic expressions, and words or phrases with multiple meanings	Uses a broad range of vocabulary, including abstract and technical terms; uses a broader range of idiomatic expressions and words or phrases with multiple meanings appropriate to context	Demonstrates comprehension of frequently occurring content words and phrases in social and content-area texts	Demonstrates comprehension of more varied vocabulary that extends beyond the everyday to include content-specific vocabulary; some idiomatic expressions, and words or phrases with multiple meanings	Demonstrates comprehension of a wide range of vocabulary, including abstract and technical terms; comprehends words and phrases with multiple meanings

Examples of Functional Language Use across English Proficiency Levels

Below are examples of language use in common academic tasks such as explaining, recounting, and arguing from evidence. The enactment of these tasks involves the use of specific language functions such as identify, sequence, compare, and evaluate (underlined in the examples). It is important to note that these are language functions, not cognitive functions. Students at all English proficiency levels can and need to perform all necessary cognitive functions; they simply express their thinking with different, more varied language as their language proficiency increases. In receptive skills, what changes is the amount and/or complexity of the language a student is able to process. In productive skills, students progress from using simple forms (e.g., 'good' or 'bad' to express an evaluation) to more complex and varied forms (e.g., "The second option, which satisfies all but one requirement, is preferable.") The descriptions are written to depict classroom learning contexts and convey both the use of language supports and their removal as students gain autonomy in language.

The following list provides examples of functional language use when students recount, explain, or argue. This list is not exhaustive. It merely provides examples of language used when students engage with content and disciplinary practices.

Recount, as might be observed at low, medium, and high levels of proficiency in the productive domains (speaking and writing)

Low: Students can use sentence starters and graphic organizers to recount the events of a story and to <u>relate predictions</u> about future outcomes, using a series of simple sentences and <u>providing some details</u> with adjectives and common prepositional phrases.

Moderate: Students can use graphic organizers to recount the events of and <u>make predictions</u> related to a story, <u>providing more precise details and more nuanced relationships between ideas</u> through greater variety in word choice and logical connectors.

High: Students can recount the events in a story, <u>provide extensive details</u>, and <u>make predictions of varied degrees of certainty</u> about future outcomes in ways that are more concise, by embedding clauses and phrases within sentences.

Explanation, as might be observed at low, medium, and high levels of proficiency in the receptive domains (listening and reading)

Low: When listening to/reading an illustrated explanation of how to solve two-step equations, students can <u>follow single-step directions</u> provided in short, imperative sentences.

Moderate: When listening to/reading visually supported explanations of how to solve two-step equations, students can <u>follow multi-step directions</u> conveyed in short, simple sentences.

High: When listening to/reading several math partners' explanations of how to solve two-step equations, students can <u>follow detailed multi-step directions</u> conveyed in a variety of simple to complex sentence types.

Explanation, as might be observed at low, medium, and high levels of proficiency in the productive domains (speaking and writing)

Low: When working in a small group to explain why the identity property of multiplication is true, supported by a graphic organizer and a word bank, students can communicate in <u>short</u>, <u>simply-stated explanations of the similarities between</u> multiplication and repeated addition.

Moderate: When working in a small group and supported by a graphic organizer to explain why the identity property of multiplication is true, students can communicate in <u>longer</u>, <u>more detailed explanations of the similarities between</u> multiplication and repeated addition.

High: When working in a small group to explain why the identity property of multiplication is true, students <u>can use a variety of sentence structures</u> to communicate <u>more concise explanations of the similarities between</u> multiplication and repeated addition, and can <u>construct a more authoritative stance</u> through the <u>use of grammatical structures</u> such as <u>nominalizations</u> and the use of <u>passive voice.</u>

Argument, as might be observed at low, medium, and high levels of proficiency in the receptive domains (listening and reading)

Low: When reading a short, highlighted, grade-level paragraph supporting a hypothesis, students can <u>identify the claim and related warrants of</u> an argument when provided a word bank, graphic organizer and scaffolded reading support.

Moderate: When reading multiple, grade-level paragraphs supporting a hypothesis, students can <u>identify the claim and related warrants of</u> an argument when provided a graphic organizer and scaffolded reading support.

High: When reading multiple, grade-level paragraphs supporting a hypothesis, students can <u>identify the claim and related warrants of</u> an argument.

References

Bailey, A. & Heritage, M. (2013). The role of language learning progressions in improved instruction and assessment of English language learners. Unpublished manuscript. UCLA.

Linquanti, R. & Cook, H. G. (2013). *Toward a "Common Definition of English Learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options.* Washington, D.C.: Council of Chief State School Officers.

Appendix A.
Participants in the national working session on defining the "English proficient" performance standard
September 17, 2013, CCSSO, Washington, D.C.

Participant	Representing the following institutions/organizations*
Patricia Adkisson	State Title III Directors Association (Alaska)
Supreet Anand	US Department of Education (Observer)
Rosa Aronson	Teachers of English to Speakers of Other Languages
Alison Bailey	UCLA (Session Leader)
Tim Boals	WIDA
Karen Cadiero-Kaplan	California Department of Education
Martha Castellon	Understanding Language-Stanford University
Mariana Castro	WIDA
Magda Chia	Smarter Balanced
Fen Chou	CCSSO
Laurene Christensen	National Center and State Collaborative (NCSC)
H. Gary Cook	Wisconsin Center for Education Research (Session Facilitator)
Elizabeth Cranley	WIDA
Chane Eplin	PARCC and ELPA21 (Florida)
Thomas Falkinburg	Office of Civil Rights (Observer)
James Ferg-Cadima	Mexican-American Legal Defense and Education Fund
Shannon Glynn	CCSSO
Margo Gottlieb	WIDA
Kenji Hakuta	ELPA21 (Understanding Language-Stanford University)
Margaret Ho	ELPA21 (Washington)
David Holbrook	ASSETS (Wyoming)
Rocio Inclan	National Education Association
Angelica Infante	New York State Education Department
Yvette Jackson	PARCC
Dorry Kenyon	Center for Applied Linguistics
Audrey Lesondak	WIDA (Wisconsin)
Robert Linquanti	WestEd (Session Leader)
Giselle Lundy-Ponce	American Federation of Teachers
Rita MacDonald	Wisconsin Center for Education Research (Session Facilitator)
Emily McCarthy	Office of Civil Rights (Observer)
Luis-Gustavo Martinez	National Education Association
Martha Martinez	ELPA21 (Oregon)
Robert Measel	PARCC (Rhode Island)

Scott Norton	CCSSO
Anita Pandey	National Association of Bilingual Education
Jen Paul	Smarter Balanced
Margarita Pinkos	Palm Beach County School District
Justin Porter	Texas Education Agency
Tamara Reavis	PARCC
Lily Roberts	California Department of Education
Gail Tiemann	Dynamic Learning Maps Consortium
Gabriela Uro	Council of Great City Schools
Dan Wiener	PARCC (Massachusetts)
Lynn Shafer Willner	WestEd (ELPA 21)
Carsten Wilmes	WIDA
Santiago Wood	National Association of Bilingual Education

*May differ from participant's institution of employment.

Re-examining Reclassification: Guidance from a National Working Session on Policies and Practices for Exiting Students from English Learner Status

Robert Linquanti, WestEd
H. Gary Cook, WCER
2015

Overview:

This document elaborates on the final of four key stages in moving toward a common definition of English learner (EL), as described in the CCSSO publication, *Toward a "Common Definition of English Learner": Guidance for States and State Assessment Consortia in Defining and Addressing Policy and Technical Issues and Options* (Linquanti & Cook, 2013).

1. **Identify** a student as a potential EL;

2. **Classify** (confirm/disconfirm) a student as an EL;

3. **Establish an "English proficient" performance standard** on the state/consortium English language proficiency (ELP) test against which to assess ELs' English proficiency; and

4. **Reclassify** a student to former-EL status through the use of multiple exit criteria.

It summarizes and further develops ideas discussed at a national working session held on September 23-24, 2014 at the CCSSO's Washington, DC offices. In particular, it recaps significant issues and tensions surrounding current EL reclassification policies and practices; provides a review of current reclassification criteria for the 50 states and Washington, DC; and offers guidance to districts, states, and multi-state consortia for moving toward more common EL reclassification criteria and processes.

Introduction

This report summarizes and further develops ideas discussed at a national working session held on September 23–24, 2014 to examine issues and options associated with reclassifying English learners (ELs) to fluent English proficient (R-FEP) status.[35] It is the fourth in a series of guidance papers designed to support states in large-scale assessment consortia that are expected to move toward a common definition of English learner as part of their assessment grant requirements. Linquanti & Cook (2013) provide a framework (p.6) for this undertaking, delineated in four stages: 1) identifying potential ELs; 2) establishing initial EL classification; 3) defining an "English proficient" performance standard; and 4) reclassifying ELs. This report focuses specifically on Stage 4, although it also necessarily touches on Stage 3, as the English-proficient performance standard on the state English language proficiency (ELP) assessment is one criterion that is very often used in determining readiness to exit specialized support services designated for English learners. Indeed, among 29 states and the District of Columbia at present, it is the *only* criterion used to determine a student's reclassification to former English learner status[36].

Federal civil rights statutes and case law entitle ELs to specialized instructional services that support both English language development and content proficiency attainment (Linquanti & Cook, 2013; Hakuta, 2011) . In addition, the Elementary and Secondary Education Act (ESEA) contains requirements for the annual ELP assessment of ELs in Title I (ESEA, s. 1111(b)(7)), and many school districts receive funds under Title III of the ESEA to support ELs' linguistic and academic growth (ESEA, s. 3102, 3111, 3115). 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 requires specialized linguistic and academic support services she is legally entitled to receive in order to meaningfully participate in classroom learning where the language of instruction is English. Under federal law, once a student is exited from EL status, local educators are expected to report on 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 this two-year monitoring period but are not permitted to do so beyond that timeframe.[37]

This paper first provides a recap of the significant issues and tensions surrounding current EL reclassification policies and practices within and across states. It next offers guidance to districts, states, and multi-state consortia, addressing key challenges and concerns identified by working session participants for moving toward more common EL reclassification criteria and methods. In doing so, it clearly signals where there was consensus among working session participants regarding reclassification criteria and processes that can be feasibly operationalized in state and local contexts. Finally, it offers some suggested approaches and strategies for moving forward.

---

[35]   The meeting was sponsored by the Council of Chief State School Officers (CCSSO) and funded, in part, by the Carnegie Corporation of New York. Working session participants and observers, and the organizations they represented, are found in Appendix A.

[36]   See below and Appendix B.

[37]   See 34 C.F.R. § 200.20(f)(2).

Issues and Tensions in Current EL Reclassification Policies and Practices

EL reclassification is complex from technical, policy, and practice perspectives. For example, researchers have documented issues in using nonlinguistic 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); in determining appropriate cutpoints of assessment-related criteria and timing of service removal (Robinson, 2011; Robinson-Cimpian & Thompson, in press); and in reporting reclassification rate outcomes and holding educators accountable given an unstable EL cohort and temporary monitoring of former ELs (Linquanti, 2001; Saunders & Marcelletti, 2013; Hopkins et al., 2013; Abedi, 2008). In effect, exit from EL status is a high-stakes decision because a premature exit may place a student who still requires specialized support related to her linguistic needs at risk of academic failure, while unnecessary prolonging of EL status (particularly at the secondary level) can limit educational opportunities, lead to stigmatization, lower teacher expectations, and demoralize students (Linquanti, 2001; Callahan, 2005; Callahan et al., 2010; Robinson, 2011; Thompson, 2015; Thompson, in press; and Estrada, 2014). These issues were of concern to participants in the working session, and the next section explores some of them in greater detail.

Number and Type of Reclassification Criteria: Review of 50 States and District of Columbia

In order to ground the discussion of EL reclassification issues nationally, we undertook a review of current reclassification criteria in the 50 states and District of Columbia. Figure 1, below, summarizes the number and type of criteria in use as of September 2015.[38]

---

[38] All information on reclassification criteria was obtained from SEA websites. See Appendix B. for summary of each state's reclassification criteria.

Figure 1. Reclassification criteria (number and type) by state.

Reclass. Criteria	States (50 + DC)		Type of Reclassification Criteria				
			State ELP Test		Additional criteria		
# of criteria	# of States	Names	Overall composite score only	Overall composite + Domain score(s)	Academic content test	Teacher input/ evaluation	Other (e.g., parent notification)
1	11	AR, DC, IN, LA, ME, NE, NM, OH*, OR*, SC, WA	X				
	19	AK, AZ, CT, GA, HI, IL, KS, KY, MD, NV, NH, NC, ND, OK, SD, TN, VT, VA, WY		X			
2	4	AL, MS, NY*, WV	X		X		
	1	UT	X			X	
	1	IA		X	X		
	2	CO, MT		X		X	
	1	DE		X			X
3	2	NJ, PA*	X		X	X	
	5	ID, MA, MI, RI*, TX*^		X	X	X	
4	1	MO*	X		X	X	X
	4	CA, FL*, MN, WI*		X	X	X	X

* = specifies possible alternate reclassification pathway/procedure. ^ = permits use of approved commercial ELP tests.

As noted above, 29 states and the District of Columbia rely solely on the state ELP assessment for reclassifying ELs, with 10 states and DC using the overall composite score, and 19 states considering both the overall composite and one or more domain scores. Ten states use the state ELP assessment and an additional criterion, which may include academic content test results, one or more forms of local educator input or evaluation (e.g., course grades, GPA, observations, scored writing samples, etc.), and in one case, other criteria. Seven states use three criteria, and five states use four criteria. These patterns represent a notable consolidation of the number and kind of reclassification criteria used compared to past national reviews. For example, Wolf et al. (2008) found that only 12 states used a state ELP assessment as the exclusive criterion for exit, compared to 29 states currently. Note also that 16 states use teacher input/ evaluation and one or more other criteria, which are all locally determined. Further information on each state's reclassification criteria is available in Appendix B.

Construct-Relevant Reclassification Criteria

Researchers have long expressed concerns about maintaining a student in EL status based on nonlinguistic performance criteria that could relate more to aptitude in mathematics or language arts, which monolingual English speakers may also have difficulty demonstrating (Linquanti, 2001; Abedi, 2008). They have also noted serious threats to validity in using academic achievement assessments that are neither designed

nor intended to support inferences about EL students' English language proficiency, as well as in using course grades or grade point average (GPA), which regularly include nonlinguistic factors (e.g., attendance and homework submissions) and are rarely standardized (Linquanti, 2001). In the lack of standardization, researchers have also documented instances of vaguely defined and inconsistently applied subjective criteria in teacher judgments that can counter or overrule more objective, standardized measures of English proficiency (Estrada & Wang, 2015).

English Learner Status as Gatekeeper and Reclassification as Performance Metric

As noted above, ample research evidence suggests that prolonged EL status based on questionable criteria and poorly implemented procedures may limit EL students' opportunity to learn. Particularly at the secondary level, EL status can act as a gatekeeper to rigorous curriculum and college- and career-readiness. Growing awareness of the population of ELs who are not well served by the education system and become "long-term ELs" (Olsen, 2010)[39] has increased the focus on the EL reclassification rate as a key performance outcome in state and local educator accountability systems.[40] While intended to focus educator attention on the quality of educational services provided to ELs, and to strengthen interventions for ELs not making sufficient progress, such policies have raised concerns about a "rush to reclassification" (Gándara, in press). In particular, chasing the indicator could motivate an unreasonable lowering of reclassification performance criteria or foreclose effective EL instructional program options. For example, there is research evidence that high-quality instructional programs developing EL students' academic proficiency in two languages may lengthen the time to reclassification yet yield better long-term academic outcomes *in English* for EL students relative to those receiving non-bilingual instructional services (Umansky & Reardon, 2014). Moreover, the reclassification rate can be a problematic performance metric; rates are often not comparable across districts in states that allow for locally-defined reclassification criteria because the criteria can vary significantly from district to district (Linquanti, 2001; Parrish et al., 2006; Ragan & Lesaux, 2006). Criteria variability and standardization are discussed further below.

English-Proficient Versus Reclassified

There currently exists a disjunction between meeting the "English proficient" requirement for ESEA Title III and exiting from EL status for accountability purposes under ESEA Title I. Specifically, the former is based solely on the state ELP assessment while the latter may involve multiple criteria that include the Title III ELP assessment result, academic achievement assessment results, and other locally identified criteria.[41] 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 per Title I. This very issue

---

[39]   For example, students in EL status for more than five or six years have been considered "long-term EL."

[40]   For example, California has included districts' EL reclassification rate as a key indicator in its *Local Control Accountability Plan* system, and New York includes EL reclassification as a key indicator in its teacher evaluation policy, which therefore disallows local teacher judgment to inform reclassification decisions.

[41]   As seen in Figure 1 and Appendix B, 21 states use multiple (two or more) criteria in reclassification decisions. Many of these states permit locally established criteria that vary *within* a state, thus leading to non-uniform, within-state definitions of EL (National Research Council [NRC], 2011). However, as also noted, an increasing number of states (29 states and the District of Columbia, up from 14 states in 2006–7) have instituted a single reclassification criterion.

generated significant contention between the federal government and states as noted in the Federal Notice of Final Interpretations (NOFI) of Title III with respect to annually assessing EL students.[42] Related to this issue, the US Department of Education's Office for Civil Rights (OCR) and the US Department of Justice's Civil Rights Division (DOJ) recently issued a 40-page "Dear Colleague" letter updating and consolidating their interpretation of statute and case law with regard to serving English learners (US ED & US DOJ, 2015). Concerning EL reclassification, OCR and DOJ maintain that EL students must meet the English-proficient performance standard on the state ELP test in order to exit EL status, and that any additional criteria used statewide or locally "may not serve as a substitute for a proficient conjunctive or composite score on a valid and reliable ELP assessment" (p. 33).

Monitoring Performance after Exit

With regard to the obligations of states toward former English learners, Title III requires states to monitor the subsequent academic performance of reclassified former ELs for two academic years following their exit from the status. OCR and DOJ have specified that during this monitoring period, school districts should ensure that former ELs "have not been prematurely exited; any academic deficits they incurred as a result of participation in the EL program have been remedied; and they are meaningfully participating in the standard instructional program comparable to their never-EL peers" (p. 34). Further, they specify that those former-ELs exhibiting academic difficulties attributable to "a persistent language barrier" should be retested on the state ELP assessment, and that those students scoring below the English-proficient standard must be reentered into EL status and offered EL services (p. 34). While these stipulations are meant to incentivize stronger EL program services and discourage—or rectify the effects of—premature EL reclassification, they pose significant implementation challenges. For example, how are educators to distinguish "deficits" incurred specifically as a result of the EL language instruction educational program? How are they to distinguish subsequent academic difficulties attributable to a persistent language barrier? And if they are to use the state's ELP assessment to do so, should that be done outside the annual test window?

Importantly, very few states examine former ELs' long-term outcomes *beyond* the two-year monitoring period. The language demands reflected in the new content standards increase substantially at higher grade levels. Some English learners reclassified in early elementary grades may face challenges that go unrecognized and unaddressed because they manifest after the two-year period. This absence of meaningful accountability for the long-term academic performance of students entering US schools as ELs has been thoroughly documented (ELL Policy Working Group, 2011; Saunders & Marcelletti, 2013, Hopkins et al., 2013). In particular, it creates a "Catch-22" with respect to reporting outcomes of a continually changing population of current ELs and ignores the long-term performance of those students who began as EL.

---

[42]   ED "strongly encourage[d]" states to use the same definition of "English proficient" for purposes of Title III (AMAO 2) and for purposes of exiting the "LEP" subgroup under Title I. See Notice of Final Interpretations, 73 Fed. Reg. 61828 (Oct. 17, 2008), pp.61837–61838.

Local Control, Standardization, and Comparability

Methods of reclassifying ELs vary among states and even among districts within states permitting local control of EL reclassification. States that are part of content or ELP assessment consortia are working to establish a more consistent set of policies and practices to define ELs "in a manner that is uniform across member states and consistent with section 9101 (25)" (USED, 2010, p. 20). Yet creating a uniform set of policies, practices, and methods for EL reclassification is extremely challenging across states and within states that permit local control. A National Research Council panel convened to examine the comparability of EL definitional processes across states concluded, with respect to EL reclassification, that

> Because of the wide variety of state policies, practices, and criteria for reclassifying students as former English language learners, and thereby exiting them from Title III services, a given English language learner student may remain in the classification longer in one state than in another state. In local control states, similar variation may exist among districts within the same state. (NRC, 2011, p.90)

This also means that a student who qualifies for EL services in one district or state may not in a bordering district or state. Thus, a tension exists in states with a tradition of local control in educational decision-making. On the one hand, moving towards more standardized reclassification criteria and processes within a state is made more challenging by state laws allowing (or requiring) local educators to exercise judgment in a reclassification decision. And, as mentioned above, there is research evidence suggesting that local educator judgments in reclassification can be inconsistent and construct-irrelevant. This may be contributing to states' movement towards standardized, single-criterion policies. On the other hand, since reclassification is a high-stakes decision with potentially substantial educational consequences, professional standards of educational and psychological testing suggest that such a decision should *not* be made using a single test score (AERA/APA/NCME, 2014).[43] Moreover, the federal definition of what constitutes "limited English proficiency" (ESEA s.9101(25))—which applies to ESEA Title I and Title III and which the US Department of Education (ED) requires consortia to utilize in moving toward a more common definition of English Learner—highlights multiple dimensions of English language proficiency.[44] This suggests the use of complementary sources of ELP evidence that could strengthen the validity of inferences about English language proficiency and of reclassification decisions. Yet such complementary sources of evidence are difficult to implement in states that permit local reclassification criteria, let alone across states with varying reclassification policies.

CCSSO National Working Session on Reclassification of English Learners

In order to foster a more common understanding of these key issues and develop guidance for strengthening policies and practices related to EL reclassification, CCSSO convened state and consortia ELP and alternate assessment representatives; district EL experts; EL advocates; and EL researchers in a working session of structured, facilitated discussions. The working session goals included the following:

---

[43]   See AERA/APA/NCME Standard 12.10, p. 198; Standard 12.13, p. 199; and Standard 13.9, p. 213. Discussed further below.
[44]   Discussed further below, and in Linquanti & Cook, 2013, pp. 4–5, 16–17.

1. Provide input for guidance for states and consortia to support local educational agencies (LEAs) in making decisions about reclassification/exit of EL students;

2. Explore criteria and methods that examine the linguistic contribution to academic performance, classroom achievement, and career/societal participation;

3. Discuss and provide input on assessment tools to support and help standardize local reclassification criteria that are relevant, reliable, valid, and comparable; and

4. Discuss and provide input on within-state, cross-state, and within- and cross-consortium EL reclassification criteria and procedures.

Participants engaged in whole-group and small-group discussion sessions, using guiding questions provided for each session.[45] Small groups met independently to work through guiding questions, then reconvened to share ideas and work toward consensus. Participants had the opportunity to ask questions or seek clarification in the whole group. The next section distills the group's conversations and suggested guidelines related to the areas listed above, which the authors have further elaborated and developed for consideration by state and consortium stakeholders and policymakers.

Guidelines for EL Reclassification

1. ***In strengthening reclassification policies and practices, states and districts should clearly define intended purposes and outcomes—and anticipate and address unintended negative consequences—for English learners.***

Throughout the working session, participants noted an underlying tension regarding EL reclassification. On the one hand, English learners are a protected class under federal and state law. They have the right to receive supplemental English language development (ELD) and specialized academic instruction to ensure their development of English proficiency and meaningful access to grade-level academic curricula and content learning. On the other hand, EL status itself can function as a gatekeeper to more rigorous curriculum and instruction, particularly as ELs enter upper elementary and secondary levels. Prolonged EL classification, especially when based on construct-irrelevant criteria, may itself lead to (or reflect) lower expectations, a deficit orientation, tracking, marginalization, stigmatization, and inferior learning opportunities. These concerns are particularly relevant in the current period of implementation of challenging college- and career-ready standards and corresponding ELP standards that reflect more rigorous language uses needed to carry out content-area practices.

Session participants emphasized that *reclassification criteria and processes should never result in denying ELs access to rigorous standards-based content instruction and appropriate opportunities to learn*. They also advised that any criterion used for reclassification should demonstrably address the construct of interest (language proficiency and use); be targeted, relevant, and meaningful to teachers and students; and be fairly applied. Reclassification criteria should not establish unnecessary hurdles that similarly-situated never-EL students would not be able to meet. Additionally, they noted that in

---

45   See Appendix C. for Working Session Agenda, and Appendix D. for guiding questions.

alignment with federal guidance flowing from *Castañeda v. Pickard*, local districts should regularly evaluate and act on evidence of the effectiveness of instructional services provided to ELs, particularly "long-term" ELs.

Participants also noted that policies and processes should be in place to detect and correct initial classification errors so that those initially misclassified as EL have their classification corrected and are not required to meet reclassification criteria (see Cook & Linquanti, 2015 for discussion of these ideas). Finally, participants also warned that making reclassification a high-stakes performance indicator for educators could incentivize lowering of performance standards or have a chilling effect on instructional program options that pursue more ambitious goals (e.g., biliteracy and academic achievement in two languages) and therefore may require longer time frames for reclassification. They emphasized that any change in reclassification policy should be carefully analyzed and designed to prevent unintentionally dismantling civil rights protections for ELs, reducing the rigor of instruction, or foreclosing instructional program options.

In sum, participants advised consortia, states, and districts to explicitly define the intended purposes of EL reclassification. This includes a rationale for any criterion selected and guidelines to ensure the rigor and fairness of any process used, with appropriate resources for training and implementation. Consortia, states, and districts should also anticipate, prevent, and address any unintended negative consequences resulting from reclassification policies and procedures. This includes regular evaluation of the implementation and the effects of reclassification policies and practices on former English learners, as well as current EL students not meeting reclassification criteria within a reasonable timeframe. The intended purposes and potential unintended negative consequences of other guidelines (presented below) should also be identified and discussed.

2. ***States and districts should select reclassification criteria that directly relate to students' uses of language needed to carry out grade-level practices in academic content areas and to meet grade-level content standards.***

In this working session and in previous guidance,[46] EL reclassification criteria were conceptualized using the federal definition of an English learner ("limited English proficient" or LEP in ESEA s. 9101(25)) that states in consortia are required to use. This definition calls out students' linguistic-minority background, non-English-speaking environment and language use, and associated difficulties in English reading, writing, speaking, and comprehension that *may* deny them: 1) the ability to meet the state's proficient level of achievement on state assessments; 2) the ability to successfully achieve in classrooms where the language of instruction is English; or 3) the opportunity to participate fully in society (ESEA, s.9101(25)). A decision to reclassify could entail providing evidence that the linguistic barriers to EL students' *possibility* of meeting these criteria have been removed. Each of these criteria, illustrated in Figure 2 below, along with potential sources of evidence for addressing them, was explored and discussed in depth.

---

[46]   See Linquanti & Cook, 2013 (pp. 13-14, 16-17).

Figure 2. Reclassification criteria operationalized from federal ESEA "LEP" student definition.
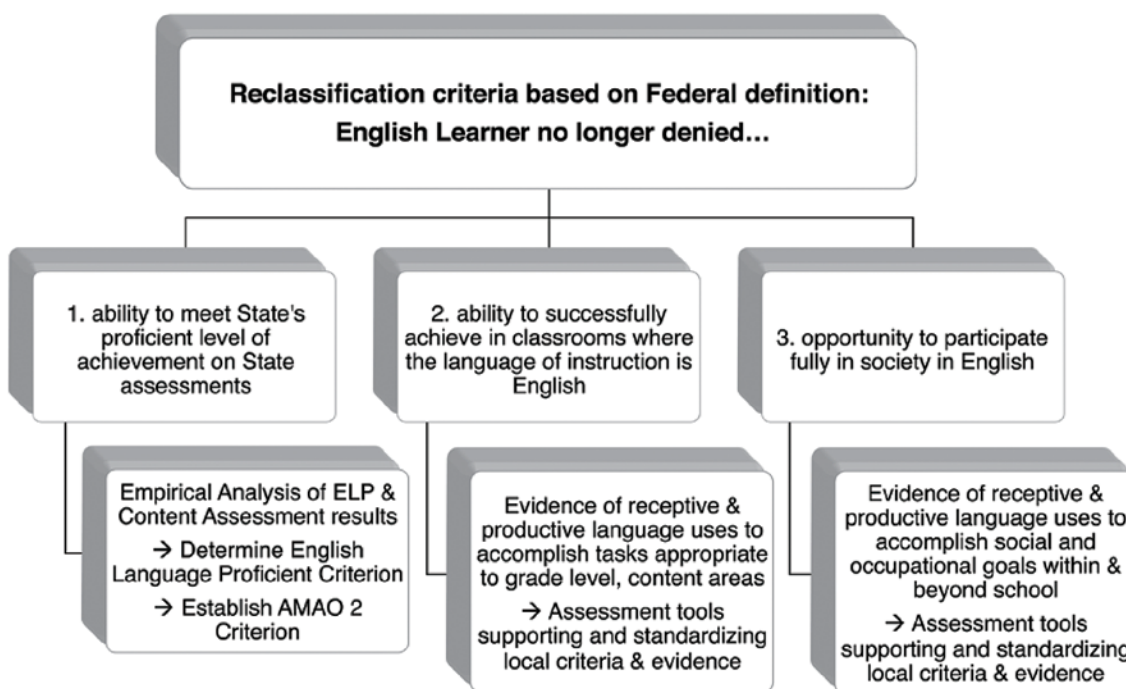
From Linquanti & Cook, 2013 (p. 16)

Working session participants expressed strong support for using the state or consortium ELP assessment to address the first criterion. They concluded that a state should utilize empirical techniques that examine ELP and content assessment relationships to help determine an English-proficient performance standard on the ELP assessment. These empirical techniques take account of the student's performance on content tests when establishing an English-proficient performance standard. Specifically, such methods attempt to account for the *linguistic contribution* to academic performance *while not requiring a minimum level of performance on the state's academic content tests for exit*. (See guideline 3 below for further elaboration.)

Participants also reached consensus on the second criterion and endorsed gathering evidence of EL students' uses of language in academic content area classrooms. They did so for several reasons. First, language-intensive practices (e.g., constructing arguments from evidence and critiquing others' reasoning; providing detailed explanations and communicating information; seeking clarification and building on what others say in oral exchanges, etc.) are explicitly defined and critically important in new college- and career-ready content standards and corresponding ELP standards. Second, as these practices entail more interactive and strategic uses of language, large-scale standardized testing approaches are less able to appropriately sample such uses of language in a single, annual administration. Although gathering and evaluating such evidence locally in a standardized, comparable way is challenging, participants concluded that doing so could capture *complementary* evidence that examines EL students' language uses while engaging in content learning and demonstrating learning in the classroom. They also saw its potential to help educators better recognize and foster students' discipline-specific uses of language during content instruction (e.g., in science, math, social studies, etc.), *if they are appropriately supported to do so*. (See guidelines 4 and 5 below for further elaboration.)

Finally, participants did *not* reach consensus on the third criterion, examining language uses that contribute to EL students' opportunity to participate fully in society using English. Participants concluded that this goal was too distant and related to the previous two criteria to adequately distinguish it in most K–12 grades. They concluded that many of the language uses associated with this dimension are already reflected in content and corresponding ELP standards, and expressed concern that such a criterion could create an additional, unnecessary hurdle to reclassification that never-EL peers would not face. Importantly, participants recognized this dimension as more a *capstone program outcome* than individual EL student exit criterion. As a result, they recommended that it be addressed through evaluation of more long-term outcomes of EL program effectiveness (e.g., high school graduation rates; college-entry and completion rates; attainment rates of biliteracy and academic achievement in two languages, etc.). Such indicators can contribute to a broader validity argument for reclassification criteria and processes, as well as strengthen system accountability for students entering school as English learners. (See guideline 9 below for further elaboration.)

3. ***States should establish the "English proficient" performance standard on the state ELP assessment using methods that take account of EL students' academic proficiency on content assessments.***

Participants strongly endorsed states utilizing empirical methods that examine the relationship between EL students' performance on English language proficiency and academic content assessments in order to establish an "English proficient" performance standard on the state ELP assessment. In recent years, several methods have been introduced and employed to determine the range of performance on an ELP assessment where EL students' academic content achievement assessed using English becomes *less* related to their ELP level.[47] That is, their level of English language proficiency no longer appears to inhibit meaningful participation on state academic assessments. Previous guidance in this series illustrated how such methods can be used by policymakers to establish an English-proficient performance standard on state ELP tests and how such performance standards might be examined for comparability across different ELP assessments used by consortia and standalone states.[48]

Participants noted strong benefits in these proposed methods. In particular, such methods take into account EL students' academic performance but do *not require a minimum level of academic performance on a content test in order for EL students to reclassify*. This removes the risk of holding ELs to a higher standard than non-ELs, while still highlighting the relationship between language proficiency and academic content performance. Participants also noted risks in using such methods and suggested ways to mitigate them. For example, they noted: a) the methods are contingent upon states setting the academic proficient level appropriately on the state content assessments; b) content assessment performance standards may be raised over time, as students initially perform lower on new, college- and career-ready academic assessments, and then improve as curriculum and instructional practices aligned to the new standards are implemented; and c) the relationship between ELP and academic performance could be misunderstood as straightforwardly causal and lead educators to take a sequential approach to teaching ELs (language first, then academics) rather than recognize the simultaneous nature of developing language, content knowledge, and content area practices.[49]

---

[47]   See Cook, Linquanti, Chinen, & Jung (2012), pp. 7–28, for a thorough discussion and demonstration of three empirical methods for determining a range of ELP performance where policy deliberations could begin.
[48]   See Linquanti & Cook (2013), pp. 11–14, 25–31, and Cook & MacDonald (2012) for further discussion.
[49]   See Heritage, Walqui, & Linquanti (2015), and van Lier & Walqui (2012) for further discussion.

Participants believed that each of these three issues could be adequately managed and were outweighed by the benefits of empirically establishing an English-proficient performance standard on the state ELP assessment. First, states in Smarter Balanced and PARCC, the two academic content assessment consortia, have set performance criteria through a transparent performance standard-setting process and have committed to examine the comparability of their college- and career-ready performance standards across these assessments. Second, regarding rising content assessment performance standards, states and ELP assessment consortia can correspondingly re-evaluate the English-proficient performance standard (e.g., every three years) as needed. Moreover, consortia, states, and local educators are leveraging the new standards and assessments to strengthen pedagogical practice with ELs (see guideline 5 below). To the last concern, there is always a risk that test users will equate correlation with causation. These empirical approaches help determine when English language proficiency is no longer a dominant factor in determining academic content proficiency. States can clearly communicate that such methods are not designed for and cannot speak to predicting academic performance based on an EL student's ELP assessment score.

While not directly related to the above concerns, the construction of ELP assessment composite scores also influences their relationship to academic content assessments. Specifically, the weighting of domain scores in creating the composite score matters when examining ELP to academic content assessment relationships because domain scores (listening, speaking, reading, and writing) have been shown to have different relationships with content performance (Parker, Louie and O'Dwyer, 2009; Linquanti & Cook, 2013; Cook, 2014; Wolf, Guzman-Orth, Hauk, 2014). Thus, establishing an English-proficient performance standard requires stakeholders, assessment developers, and policymakers to articulate a clear rationale for their ELP assessment composite-score-weighting approach and to provide evidence of the efficacy of the identified weighting procedure.[50]

4. *States and districts should make EL reclassification decisions using more than an annual summative ELP assessment result; they should also examine EL students' classroom language uses as an additional reclassification criterion.*

As noted above, the AERA/APA/NCME professional standards of educational and psychological testing (2014) suggest that a highly consequential decision such as EL reclassification should *not* be made solely on the basis of a single test score.[51] For example, consider the following professional standards:

Standard 12.10: In educational settings, a decision or characterization that will have major impact on the student should take into consideration not just scores from a single test but other relevant information (p. 198).

Standard 12.13: When test scores are intended to be used as part of the process for making decisions about educational placement, promotion, implementation of individualized educational programs, or provision of services for English language learners, then empirical evidence documenting the relationship among particular test scores, the instructional programs, and desired student outcomes should be provided. When adequate empirical evidence is not available, users should be cautioned to weigh the test results accordingly in light of other relevant information about the students (p. 199).

---

[50] See Linquanti & Cook, 2013, p. 13.
[51] See pp.7–8 above, and footnote 9.

In line with addressing the federal definition of "limited English proficiency" discussed above, participants expressed strong consensus on gathering evidence of ELs' language uses in the classroom context to support judgments about students' "ability to achieve in classrooms where the language of instruction is English." Specifically, they saw a clear value to EL students' teachers examining the more collaborative, interactive language uses posited in new ELP standards that are not adequately captured by large-scale ELP assessments and which could yield complementary evidence useful for reclassification decisions. In effect, the participants considered this to be "other relevant information" about the student, as expressed in these professional standards.

In discussions, the group suggested several guidelines related to the development and implementation of evidence-gathering strategies and tools on classroom language uses. First, evidence gathered should be complementary to, and not duplicative of, language uses targeted on the state ELP assessment. Any observational protocol or evaluative rubric used to gather evidence of classroom language uses should be student-focused, assets-based (i.e., describing what EL students *can do* with English), seen as relevant and pedagogically useful by classroom teachers, and developed for use both by *ESL and academic content area* teachers.

Participants also stressed that substantial professional development and sustained administrative support would be critical to successfully implement a locally-administered observation protocol statewide. This includes a mechanism for calibrating judgments among teachers. (See guideline 5 below for further elaboration.) The evidence-gathering methods should help educators to regularly examine and recognize a *range of proficiencies* in target language uses and not just focus at the level of performance judged English-proficient for reclassification purposes. Participants argued that such observational rubric/protocols should be useful throughout the year for formative purposes (i.e., to gather evidence of strengths and growth areas, guide instructional moves, provide feedback to students, and scaffold students' further language use and disciplinary learning). They also proposed that the observational rubric/protocol be used within a specific assessment window for summative purposes related to reclassification decision-making, particularly as a more standardized method to inform "teacher judgment/recommendation" criteria, while acknowledging the potential conflict in using the same protocol for formative and summative purposes. They also recognized that such an observational tool/protocol might be difficult to implement as a reclassification criterion in states that currently use only the ELP test for reclassification decisions, or that use English proficiency attainment or reclassification rates as part of teacher evaluation.

Given the complementary nature of the two reclassification criteria specified in this guideline, participants noted the need for clear rules to define how these criteria should be combined to make a reclassification decision. For example, performance results on these two criteria could be set conjunctively, or could allow a judiciously compensatory approach with some conjunctive minimum performance on each. See guideline 6 for further discussion.

5. ***States and districts should ensure that local educators have training, tools, and ongoing support to effectively and consistently apply the classroom language-use criterion for reclassification decisions and are held appropriately accountable for doing so.***

As noted in Guideline 4 above, participants predicated their support for a common statewide, locally-administered language use observation protocol/rubric on the condition that local educators have high-quality, validated tools and processes, and adequate training and ongoing support to effectively and consistently apply the classroom language-use criterion for reclassification decisions. They acknowledged that such tools and practices will need to be collaboratively developed, perhaps through funded R&D efforts that involve iterative prototyping and field-testing of particular language-intensive practices within content areas. They also recommended the use of video and audio samples of EL language use for online calibration training and certification for summative purposes. At least one large EL-enrolling state has employed student video samples and other digitized artifacts for online calibration and certification of teachers for summative ELP assessment purposes, and one of the ELP assessment consortia is also doing so for locally scoring its speaking domain.[52]

While acknowledging the ambitiousness of this undertaking, participants noted states' growing recognition that new college-and career-ready standards strongly emphasize these more collaborative, interactive uses of language to carry out science, math, and English language arts practices. For example, state instructional frameworks are being implemented that support teachers' capacity to note and foster all students' language uses in content classrooms, and pre-service and in-service professional learning initiatives are also developing new tools and methodologies for developing educator capacity to do so.[53] Also, renewed interest in performance assessments that explore more complex, transferable capacities—including linguistic capacities and growing interest in teacher formative assessment practices that support ELs' language uses—offers positive supports for language observation protocols.[54] These efforts were also seen as significant in order to remove instruments that assess decontextualized, atomistic language use, or allow subjective, unanchored judgments of what EL students can do with English; as well as for replacing such nonstandardized and possibly construct-irrelevant indicators as grades, GPA, etc.

6. States and districts should collaborate to establish common reclassification criteria and processes <u>within</u> states, with a goal of strengthening the validity of inferences made from local educator input and the accuracy of decisions based on multiple sources of evidence.

As noted above, there is growing awareness of the need to establish common classification and reclassification criteria and processes within a given state in order to ensure educational equity for a state's English learners. When criteria and processes vary from one local educational agency to the next, students who are designated English learners in one district may not be so in a bordering district and therefore cannot be assured comparable or coherent services. This also undermines EL performance comparisons of districts within the state. (See guideline 9 below.)

---

[52]   Texas employs web-based calibration and certification for teachers to make classroom-based summative judgments of EL students' listening, speaking and writing. See http://tea.texas.gov/student.assessment/ell/telpas/. The WIDA Consortium also does so for the scoring of speaking and writing.

[53]   See for example, California's new ELA/ELD curriculum framework, especially vignettes emphasizing integrated and designated ELD (http://www.cde.ca.gov/ci/rl/cf/elaeldfrmwrksbeadopted.asp); the Stanford University *Understanding Language* initiative's online and MOOC-related training and resources on classroom discourse (http://ell.stanford.edu/content/moocs); and WIDA's Dynamic Language Learning Progressions (http://www.dllp.org).

[54]   See Darling-Hammond & Adamson (2014); Duckor (2014); Heritage, Walqui, & Linquanti (2015); and MacDonald, Boals, Castro, Cook, Lundberg, & White (2015).

Importantly, the cumulative weight of research evidence, advocacy, and policy discussions has registered this issue with federal legislators. As of this writing, Senate Bill 1177 reauthorizing the Elementary and Secondary Education Act, as approved by the full Senate, contains a provision that require states to "establish and implement, with timely and meaningful consultation with local educational agencies representing the geographic diversity of the state, standardized statewide [EL] entrance and exit procedures" (Senate Bill 1177, s.3111(b)(2)(A); and s.3113(b)(2)).

As reflected in this draft legislative language, the state/local balance is particularly important for states that permit or require local educator judgment. Such policies value and privilege the judgment of educators closest to the student, who presumably understand best what ELs can and cannot do with English. Participants also noted that local decision-making may be specified in state statute and in federal civil rights consent decrees negotiated with local districts that may specify local reclassification criteria and processes to be used. Therefore, they urged that guidelines strike a balance between standardized statewide criteria and informed, standardized local input in order to yield optimal reclassification decisions. This is possible to the extent that sources of evidence can be standardized, standards-based, complementary, and validated. In particular, participants agreed that locally-administered, language use observation protocols/rubrics can help to strengthen the validity of inferences of a greater range of language uses EL students need for success in school and also help to develop educator capacity to note and support these more complex interactive language uses specified in new content and ELP standards.[55]

Session participants also noted that an English learner could score "not English-proficient" on the state ELP test, yet score proficient on the academic content test of English language arts. To the extent that ELP standards correspond to the academic content standards, and ELP and content tests are appropriately aligned to their respective standards, this should occur only in a very small percentage of cases. Indeed, empirical studies have found these cases to occur infrequently and at lower grade levels (see, e.g., Haas, 2010). Clearly, such occurrences beyond a very small number should trigger a systematic review of ELP-to-ELA standards correspondence and test-to-standards alignment to ensure the assessments allow for sufficiently valid inferences. Nevertheless, given that the ELA assessment points to outcomes of central interest related to EL students' protected class status, participants argued that a proficient ELA test result might be considered as a "corrective criterion" in those limited instances where the evidence available from the state ELP assessment and classroom language use observations is ambiguous.

How might evidence from statewide standardized ELP test results and locally gathered, standardized language observations be combined to judge student's English language proficiency and readiness for reclassification? How and when might an EL's proficient ELA test result be used appropriately as a "corrective criterion?" Figure 3, below, illustrates in matrix format one approach that combines complementary assessment evidence in ways that are compensatory, yet with conjunctive minimum levels, and that also account for potential measurement error in either or both sources of evidence.

---

[55]   See Resnik, Asterhan, & Clark (2015) for an up-to-date review of how carefully structured academic discussions and dialogue can increase students' cognitive demand, intellectual engagement, language use, and learning opportunities.

Figure 3. Sample reclassification decision matrix combining multiple sources of ELP evidence.

		Additional ELP Criterion (classroom language observation protocol)		
		"English-Proficient"	Borderline "English-Proficient"	Not "English-Proficient"
ELP Test Result	"English-Proficient"	R-FEP [1]	R-FEP [2]	R-FEP* [3]
	"English-Proficient" within CSEM	R-FEP [4]	R-FEP* [5]	EL* [6]
	Not "English-Proficient"	EL* [7]	EL [8]	EL [9]

R-FEP: Reclassified Fluent English Proficient. CSEM: Conditional standard error of measurement.
***Indicates ambiguous ELP results where proficient ELA test result might be utilized as a "corrective criterion."**

In Figure 3, the rows illustrate three possible outcomes on the state ELP assessment: a score that is clearly "English-proficient"; a score that is clearly "Not English-proficient"; and a score that falls within a conditional standard error of measurement of the English-proficient cut score.[56] The columns illustrate three similar outcomes on the classroom-based ELP criterion.[57]

The matrix illustrates *hypothetical* reclassification decision rules for ELP assessment results and also indicates unusual cases where use of a proficient ELA test result as a "corrective criterion" might be considered. For example, cells 1 and 9 provide clear signals (respectively, reclassify and remain EL). Cells 2, 4, and 8 fall within defensible ranges for decision-making (reclassify for the first two, remain EL for the third). The remaining cells yield ambiguous results and would call for consideration of a proficient ELA test result as a potential "corrective criterion": Cells 3 and 7 portray divergent outcomes, cells 5 and 6 portray borderline results, and all four illustrate a slightly greater weighting toward the ELP assessment result. Such approaches can support use of multiple forms of evidence to construct a more complete picture of EL students' English language proficiency while still allowing consistent reclassification decisions to be made. Ultimately, the goal is to establish and validate state ELP assessment cutpoints and local decision-making criteria that *maximize* EL students' classification in cells 1 and 9 and *minimize* their classification in cells 3 and 7 of the matrix.

---

[56] The conditional standard error of measurement (CSEM) can be used to generate confidence intervals around a specified test score level. Since reclassification decisions are concentrated in the area of the score scale where the "English proficient" cut score is set, the CSEM can help to define the margin of error around that score point.
[57] An analogous "borderline" margin of error around the "English-proficient" score would need to be established for the observation protocol outcome.

7. States in consortia should move toward a common English-proficient performance standard on any shared ELP assessment and acknowledge variability of other EL reclassification criteria and processes across states. They should ensure complete transparency and examine cross-state comparability as new criteria and processes are implemented.

Working session participants acknowledged the challenges of multiple states arriving at a common English-proficient performance standard even on a shared ELP assessment. This guidance series has illustrated methods for how an individual state can set a common English-proficient performance standard using ELP and academic content assessments. Nevertheless, in practice, states may justifiably have varying perspectives about what English-proficient means. For example, they may have different populations of English learners, and these students may face different language-development contexts, needs, and expectations that influence where they choose to set the "finish line." Also, states within a given ELP assessment consortium may be using different academic content assessments which reflect different operationalizations of similar or even the same academic standards and thus influence where the ELP "finish line" is set. It will therefore be critically important for states in consortia to acknowledge where variability exists and to be transparent in explaining the reasons for it.

One alternative in addressing this variability challenge is to identify a minimally "English proficient" criterion. Such an approach identifies the point (using empirical methods described earlier) above which all "finish lines" would be established. This provides for a commonly recognized minimum English-proficient performance standard while allowing states the flexibility to determine their unique "English-proficient" criterion on the shared ELP assessment, using data from methods that consider ELP assessment outcomes and outcomes from relevant academic content assessments.

Establishing a common English-proficient performance standard is still more challenging when comparing assessment results *across* ELP assessement consortia or non-consortia states. While linking studies could be conducted to establish comparability across ELP assessments, doing so may not be feasible. How then might a common understanding of an English learner, as it relates to ELP assessment scores, be pursued across states and consortia? One approach is to create a reference level of English language proficiency to which all state and consortia assessments' proficiency levels might be related. This concept, used to create the *Common European Framework of Reference for Languages* (Council of Europe, 2001), was applied in this guidance series to create *reference performance level descriptors* (R-PLDs) relating ELP assessment levels of some large stand-alone states and one ELP assessment consortium via a common frame (see Cook & MacDonald, 2014). States and consortia could compare their ELP assessments' proficiency levels using the R-PLDs. Such an alignment process could have stakeholders from states and consortia examine their current ELP levels and map them onto the R-PLDs. This approach can provide a level of transparency regarding comparability across state and consortia assessments' ELP performance standards.

Several studies could be undertaken to examine this comparability. For example, one study might examine how state and/or consortia ELP assessment proficiency levels compare for each domain (i.e., speaking, listening, reading, and writing). Oracy and literacy composite proficiency levels could be compared as well. Another study might examine the relationship between R-PLDs, linked across states and consortia, and college- and career-ready academic assessment proficiency levels. Such a study could specifically

explore language proficiency expectations required by different academic assessments. Mechanisms such as these can relate states' ELP levels and support answering important questions about English learners' English language proficiency and ELP assessment performance standards across states.

8. *Consortia, states, and districts should carefully examine the application of reclassification criteria and processes for primary-grade EL students, and EL students with disabilities, in order to maximize validity, reliability, and fairness.*

Working session participants expressed a variety of concerns regarding the reclassification of English learners in Kindergarten, 1st and 2nd grades. A key concern relates to early elementary EL students' literacy development. How can educators reliably ascertain a young EL's reading and writing levels when she is either pre-literate or emerging in literacy skills? In an effort to avoid initially misclassifying students as EL, many states assess (and classify as EL) K and 1st grade students using ELP assessments with composite scores that *under-weight* literacy in these grades (Cook & Linquanti, 2015). In some states, this under-weighting of literacy may carry over to judgments of English proficiency for reclassification in these grades. This in turn can lead to the risk of *de facto* premature EL reclassification.

Conversely, measurements of literacy on ELP assessments in these early grades may be more associated with "literacy potential" than with requisite academic literacy skills needed to participate in grade-level content practices and may not be predictive of literacy in later grades. That is, the types of academic, college- and career-ready literacy skills needed to successfully participate in the classroom using English are not exhibited (or substantially taught) until 2nd or 3rd grade and beyond. This in turn has led some states to set early-grade reclassification criteria extremely high. For example, Wisconsin sets the English-proficient score on its ELP assessment in grades K to 3 at the highest possible ELP level, unlike that at grades 4 through 12.[58] Reclassifying early-grade ELs is difficult in part because of the challenge in meaningfully assessing literacy. States should therefore examine carefully reclassification criteria and processes for these students. Several questions might be examined to better understand and address concerns about early-grade EL reclassification, especially as it relates to academic literacy. For example, researchers could identify students who exited EL status in 1st and 2nd grade and examine their literacy scores on the ELP assessment in relation to their subsequent performance on state reading/language arts assessments. Do these students exhibit acceptable grade-level literacy practices in the classroom? Do these students need additional literacy support in later grades? If so, what type of support is needed? Findings from such studies can inform the validity of inferences from the state ELP assessment's reading and writing subtests, which may necessitate revision of these assessments. Similarly, states could explore how well the ELP assessment speaking and listening scores associate with reclassified ELs' ELA or academic reading assessment performance. They might also examine the reasonableness of weighting oral skills more heavily for early-grade EL students to help determine the optimal composite score weighting strategy for these students.

Working session participants also expressed concerns about reclassification criteria and practices for ELLs with disabilities. In particular, they expressed concerns about students who: 1) may have been wrongly initially classified as EL, when in fact they instead had a language-related learning disability; or 2) are ELLs

---

[58]   See Appendix B.

with disabilities that may be unable to meet a specific reclassification criterion *due to the specific nature of their disability*. These issues have challenged the field for some time. In a previous document in this guidance series, Cook & Linquanti (2015) suggest a procedure to detect and correct within a reasonable timeframe initial misclassifications so that the student misclassified as EL—whether of SWD status or not— would *not* need to be reclassified. Nevertheless, some disabilities may emerge long after initial EL classification.

The US Department of Education clearly states in its notice of final interpretations (NOFI, 2008, p. 61831) that "Title III does not provide exemptions from annual ELP assessments for any Title III-served LEP student" and suggests that accommodations be used to ensure all four language domains are assessed. ED's subsequent guidance (USED, 2014) emphasizes that ELLs with disabilities must participate in the annual state ELP assessment with or without appropriate accommodations or by taking an alternate assessment if necessary, consistent with the student's individualized education program (IEP). It also emphasizes that ELLs with disabilities cannot be reclassified unless they meet the state's definition of English-proficient.

The technical and legal complexities of this topic are beyond the scope of this guidance paper. However, we note that recent research has examined the validity and comparability of alternate approaches to establishing English proficiency composite scores without one or two domains for those EL students whose disability prevents them from being validly assessed in those domains.[59] Recent research has also examined the relationship between alternate ELP assessments and alternate academic assessments to explore different conceptualizations of English language proficiency for EL students with severe cognitive disabilities.[60] There is clearly a need for ongoing dialogue to ensure valid inferences of English language proficiency for EL students with disabilities, fairness and reasonableness in practice, and the protection of the civil rights of students under EL and disability statuses.

9. ***Consortia, states, and districts should, as part of ensuring the consequential validity of reclassification criteria and processes, carefully examine the subsequent academic performance of reclassified English learners, for as long as these students remain in the district or state.***

As noted above, working session participants concluded that EL students' opportunity to participate fully in society was more an *outcome of*—rather than an input to—appropriate reclassification policies and practices. Ensuring the consequential validity of EL reclassification criteria and processes therefore necessitates examining the short-term *and* long-term outcomes of exit decisions. Federal law currently requires states to monitor the subsequent academic performance of former ELs for two years following exit and allows states to include these former ELs in the EL subgroup for Title I accountability for the same time period. However, session participants expressed concern that this timeframe does not allow for an appropriate examination of the long-term outcomes of students who entered school as ELs, nor for properly evaluating the consequences of EL reclassification policies and practices. Much research has highlighted the need to stabilize the cohort of students who began schooling in the state as EL when evaluating long-term consequences, and that EL reclassification, while important, is neither the whole story nor the end of story in ensuring educational equity for ELs.[61] The variability within and across

---

[59]   See Cook (2013).
[60]   See Cook (2014).
[61]   See pp. 2, 6–7 above.

states of former EL performance over the long term calls for a careful examination of these students' outcomes on such longer-term outcome measures as Advanced Placement (AP) and other college-ready course participation and outcomes; graduation rates; Seal of Biliteracy attainment; and college- and skilled career-application, acceptance, and completion rates. Such evidence gathering also reflects the expectation articulated in federal case law (e.g., *Castañeda v. Pickard*) that educators evaluate the effects of EL instructional programs and make any required improvements necessary.

Promising analytical methods have been recently published that support states and consortia in evaluating the effects of ELP assessment cut-score selection on ELs' subsequent academic performance and educational outcomes.[62] Reporting methods that properly stabilize the cohort of students who began state schooling as ELs have also been illustrated (e.g., Saunders & Marcelletti, 2013) and are being used by some states (e.g., Washington, Oregon). As more states incorporate longitudinal student data into their decision-making, examining the long-term outcomes of former ELs can and should become a part of ensuring the consequential validity of reclassification policies and practices. Doing so will not only strengthen system accountability, it will also ensure defensible comparisons of former ELs' long-term outcomes across districts within states, and across states within consortia.

Conclusion

As noted throughout this guidance, EL reclassification is a very complex issue, and it will require consortia, state, and local policymakers and leaders to proceed carefully and deliberately. Despite this complexity, there was a notable consensus among working session participants representing many different constituencies. They agreed that EL reclassification policies and practices can and should be strengthened, made more coherent, and standardized within states in ways that enable local educators— those closest to EL students—to meaningfully participate in reclassification decision-making. Developing and implementing a statewide classroom language use observation protocol was considered critical to ensuring this participation and to providing needed complementary evidence of more interactive language uses that are not captured by large-scale, summative ELP assessments. As a result of the working session and subsequent discussions, this guidance series will produce an additional document providing guidance on observational tool development that consortia, state, and local educators may find valuable as they proceed with such efforts.

Moreover, there was a strong belief that new college- and career-ready standards, corresponding ELP standards, and their respective aligned assessments provide real opportunities for educators to better understand and support the development of ELs' academic uses of language and rigorous subject matter practices and learning needed for academic success. They clearly agreed individual states can effectively work on internal coherence and collaborate with local educators to move toward consistent reclassification policies. They also believed states can continue cross-state dialogue and collaboration, within and across consortia and with standalone states, to learn from one another and to increase transparency and ensure greater comparability of EL reclassification policies and practices over time. Indeed, several states as well as the two ELP assessment consortia have been working on potential

---

[62]   See Robinson-Cimpian & Thompson, 2015.

EL identification and EL classification stages previously discussed in this guidance series. States and consortia are also preparing to use multiple analytical methods described earlier to identify a "sweet-spot" range for the English-proficient performance standard on new ELP assessments in light of new ELP standards and new content standards and assessments. While the technical demands of these methods are substantial, there is also the opportunity to tap the analytical infrastructure of applied research and technical assistance support available through the federal comprehensive centers, content centers, and regional educational laboratories.

Moving toward greater consistency and comparability of EL reclassification criteria, policies, and practices requires a willingness to collaborate—districts with their states, states within consortia, and consortia with each other and with standalone states. It will also require greater transparency in outcomes and in discussions of key challenges over time. Such efforts clearly require taking the long view and, if sustained and supported, can systematically improve policy and practice along the way to ensure greater educational equity and success for current and former English learners.

References

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

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (AERA/APA/NCME). (2014). *The standards for educational and psychological testing.* Washington DC: Author .

Callahan, R. M. (2005). Tracking and high school English learners: Limiting opportunity to learn. *American Educational Research Journal 42*(2): 305–328.

Callahan, R., Wilkinson, L., & Muller, C. (2010). Academic achievement and course taking among language minority youth in U.S. schools: Effects of ESL placement. *Educational Evaluation and Policy Analysis, 32,* 84–117.

Cook, H. G. (2014). *Examining relationships between Alternate ACCESS and state alternate assessments: Exploring notions of English language proficiency.* Madison, WI: Wisconsin Center for Education Research – WIDA Research Report. Retrieved from www.wida.us/get.aspx?id=750.

Cook, H. G. (2013). *Examining how to establish English proficiency using alternate composite scores in a WIDA state: A logistic regression approach.* Madison, WI: Wisconsin Center for Education Research – WIDA Research Report. Retrieved from www.wida.us/get.aspx?id=634.

Cook, H. G., & Linquanti, R. (2015). *Strengthening policies and practices for the initial classification of English learners: Insights from a national working session. Washington DC: Council of Chief State School Officers.* Retrieved from http://www.ccsso.org/Documents/EL Classification Working Paper_FINAL_01 26 2015.pdf.

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: US Department of Education, Office of Planning, Evaluation and Policy Development. Retrieved from http://www2.ed.gov/rschstat/eval/title-iii/implementation-supplemental-report.pdf.

Cook, H. G., & MacDonald, R. (2014). *Reference performance level descriptors: Outcome of a national working session on defining an "English proficient" performance standard.* Washington DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Documents/2014/CCSSO%20Common%20EL%20Definition%20Reference%20Performance%20Level%20Descriptors%2001272014.pdf.

Council of Europe. (2001). *Common European Framework of reference for language learning, teaching and assessment.* Strasbourg: Language Policy Division: Author. Retrieved from http://www.coe.int/t/dg4/linguistic/Source/Framework_EN.pdf.

Darling-Hammond, L., & Adamson, F. (2014). *Beyond the Bubble Test: How Performance Assessments Improve Teaching and Learning.* San Francisco: Jossey-Bass.

Duckor, B. (2014). Formative assessment in seven good moves. *Educational Leadership, 71* (6), 28–32.

Estrada, P. (2014). *English learner reclassification to fluent English proficient: Meeting criteria, roadblocks, opportunities, and consequences: Years 1 and 2 findin*gs. Practitioner Research Brief 1. Project on Reclassifying English Learners to Fluent English Proficient: Access and Achievement. Santa Cruz: University of California.

Estrada, P. & Wang, H. (2015). *Making English learner reclassification to fluent English proficient attainable or elusive: When meeting criteria is and is not enough.* Manuscript submitted for publication.

Gandara, P. (in press). Charting the Relationship of English Learners and the ESEA: One Step Forward, Two Steps Back. *The Russell Sage Foundation Journal of Social Sciences, 1*(3).

Haas, E. (2010). *Assessing achievement of English language learners: Pass-fail status on Arizona's language and content tests.* Paper presented at the annual meeting of the American Educational Research Association, Denver Colorado.

Hakuta, K. (2011). Educating language minority students and affirming their equal rights: Research and practical perspectives. *Educational Researcher, 40*(4), 163–174.

Heritage, M., Walqui, A., & Linquanti, R. (2015). *English language learners and the new standards: Developing language, content knowledge, and analytical practices in the classroom. Cambridge: Harvard Education Press.*

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.

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., & Bailey, A.L. (2014). *Reprising the home language survey: Summary of a national working session on policies, practices, and tools for identifying potential Engli*sh learners. Washington DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/Resources/Publications/Reprising_the_Home_Language_Survey.html

Linquanti, R., & Cook, G. (2013). *Toward a "common definition of English learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options. Washington DC: Council of Chief State School Officers. Retrieved from* http://www.ccsso.org/Resources/Publications/Toward_a_Common_Definition_English_Learner_.html.

MacDonald, R., Boals, T., Castro, M., Cook, H.G., Lundberg, T., and White, P. (2015). *Formative Language Assessment for English Learners: A four step process*. Portsmouth, N.H.: Heinemann.

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. Notice of Final Interpretations, 73 Fed. Reg. 61828 (Oct. 17, 2008), pp.61837–61838.

Olsen, L. (2010). *Reparable harm: Fulfilling the unkept promise of educational opportunity for California's long term English learners*. Long Beach, CA: Californians Together.

Parker, C. E., Louie, J., and O'Dwyer, L. (2009). *New measures of English language proficiency and their relationship to per- formance on large-scale content assessments* (Issues & Answers Report, REL 2009–No. 066). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands. Retrieved from http://ies.ed.gov/ncee/edlabs.

Parrish, T., Perez, M., Merickel, A., & Linquanti, R. (2006). *Effects of the implementation of Proposition 227 on the education of English learners, K–12: Findings from a five-year evaluation* (Final Report). Palo Alto, CA and San Francisco, CA: AIR and WestEd. Retrieved from https://www.wested.org/online_pubs/227Reportb.pdf.

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 from: http://epaa.asu.edu/ojs/article/view/91.

Resnik, L.B., Asterhan, C.S.C., & Clark, S.N. (Eds.) (2015). *Socializing intelligence through academic talk and dialogue.* Washington DC: American Educational Research Association.

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.

Robinson-Cimpian, J.P., & Thompson, K.D. (In press). The effects of changing test-based policies for reclassifying English Learners. *Journal of Policy Analysis and Management.*

Saunders, W. M., & Marcelletti, D. J. (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.

The Elementary and Secondary Education Act (ESEA). Public Law 107-110, Jan 8, 2002. Stat.115.1425–2094.

Thompson, K. D. (2015). English learners' time to reclassification: An analysis. *Educational Policy*. Advance online publication. DOI: 10.1177/0895904815598394

Thompson, K. D. (in press). Questioning the Long-Term English Learner label: How classification and categorization can blind us to students' abilities. *Teachers College Record*.

Umansky, I. M. & Reardon, S. F. (2014). Reclassification patterns among Latino English learner students in bilingual, dual immersion, and English immersion classrooms. *American Educational Research Journal, 51*(5), 879–912.

U.S. Department of Education (USED). (2014). *Questions and answers regarding inclusion of English learners with disabilities and English language proficiency assessments and Title III annual measurable achievement objectives.* Washington DC: Author. Retrieved from https://www2.ed.gov/policy/speced/guid/idea/memosdcltrs/q-and-a-on-elp-swd.pdf.

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

U.S. Department of Education (USED) and U.S. Department of Justice (USDOJ). (2015). *Dear colleague letter*. Retrieved from http://www2.ed.gov/about/offices/list/ocr/letters/colleague-201111.html.

Van Lier, L., & Walqui, A. (2012). Language and the common core state standards in K. Hakuta & M. Santos (Eds.), *Understanding Language: Commissioned Papers on Language and Literacy Issues in the Common Core State Standards and Next Generation Science Standards* (pp.75–85). Stanford, CA: Stanford University Understanding Language Initiative.

Wolf, M.K., Guzman-Orth, D., and Hauck, M.C. (2014). *Next-generation summative English language proficiency assessments for English learners: Priorities for policy and research.* Princeton, NJ: Educational Testing Service. Retrieved from www.p12.nysed.gov/biling/bilinged/documents/elp-policy-report.pdf.

Wolf, M.K., Kao, J., 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 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 from http://ellpolicy.org/wp-content/uploads/QA.pdf.

Appendix A.
Working Session Participants/Observers and Organizations Represented

Participant	Institution Represented
Elena Fajardo	California Department of Education
Lily Roberts	California Department of Education
Debra Dougherty	San Diego USD (CA)
Maria Larios-Horton	Santa Maria Joint Union High School District (CA)
Maria A. Meyer	Westbury UFSD (NY)
Judy Diaz	Port Chester UFSD (NY)
Cristina Vazquez	Texas Education Agency
Susie Coultress	Texas Education Agency
María Trejo	Cypress-Fairbanks ISD (TX)
Laurie Shaw	Pflugerville ISD (TX)
Tim Boals	WIDA Executive Director
Audrey Lesondak	Wisconsin Department of Public Instruction
Jonathan Gibson	Nevada Department of Education
Antonieta Bolomey	Boston Public Schools (MA)
Julie DeCook	Janesville School District (WI)
Kerri Whipple	South East Education Cooperative, Fargo (ND)
Kenji Hakuta	ELPA21 Principal Investigator
Margaret Ho	ELPA21 Sustainability Director
Jobi Laurence	Iowa Department of Education
Martha Martinez	Oregon Department of Education
Veronica Gallardo	Seattle Public Schools (WA)
Jennifer Reid	Millard Public Schools (NE)
Emily Dossett	Manhattan USD 383 (KS)
Martha Thurlow	National Center and State Collaborative (NCSC)
Meagan Karvonen	Dynamic Learning Maps (DLM)
Delia Pompa	National Council of La Raza (NCLR)
Gabriela Uro	Council of the Great City Schools (CGCS)
Robert Linquanti	WestEd (Session Co-Lead)
H. Gary Cook	Wisconsin Center for Education Research (Session Co-Lead)
Rita MacDonald	Wisconsin Center for Education Research (Session Facilitator)
Alison Bailey	University of California - Los Angeles (Session Facilitator)
Paola Uccelli	Harvard Graduate School of Education
Okhee Lee	New York University
Fen Chou	CCSSO (Session Organizer)

Working Session Observers

Participant	Institution Represented
Supreet Anand	US Department of Education, OESE (Observer)
Marianna Sanchez Vinson	US Department of Education, OELA (Observer)
Carlos Martinez	US Department of Education, OELA (Observer)
Emily McCarthy	US Department of Justice (Observer)
Carolyn Seugling	US Department of Education, Office for Civil Rights (Observer)
Scott Norton	CCSSO (Observer)
Cassius Johnson	Carnegie Corporation of New York (Observer)
Jennifer Timm	Carnegie Corporation of New York (Observer)

Appendix B. State Reclassification Criteria

State	ELP Test Performance Standard	Additional Criteria
Alabama	ACCESS 4.8 Composite and proficiency on the reading portions of the statewide assessments	N/A
Alaska	ACCESS 5.0 Composite or higher on Tier B or C and 4.0 or higher in each domain (L,S,R,W)	N/A
Arizona	AZELLA: "Fluent English Proficient" on Overall Composite Score, Reading domain score, and Writing domain score	N/A
Arkansas	ELDA: Level 5 "Fully English Proficient" (Based on composite score (average of four domain tests).	N/A
California	CELDT: "Early Advanced" or "Advanced" Level (Domains of listening, speaking, reading, and writing are at "Intermediate" level or above)	Multiple measures be used to reclassify ELs but must include at least all four of the following: – Assessment of English language proficiency – Teacher evaluation – Parental opinion and consultation – Comparison of student performance in basic skills against an empirically established range of performance in basic skills based on the performance of English proficient students of the same age
Colorado	ACCESS 5.0 Overall and 5.0 Literacy on Tier B or C	Additional evidence to confirm fluent English proficiency aligned with the CELP Standards. At least one piece of local data that confirms grade level proficiency in reading. At least one piece of local data that confirms grade level proficiency in writing
Connecticut	LAS LINKS: Composite Level 4 or 5; and Reading Score of 4 or higher and Writing Score of 4 or higher.	N/A
Delaware	ACCESS 5.0 CPL on Tier C and a Reading PL no less than 4.7 and a Writing PL no less than 4.5	Individual domain scores should be reviewed by district/charter ELL coordinator before student is exited
District of Columbia	ACCESS 5.0 or higher composite on Tier B or C	N/A

State	ELP Test Performance Standard	Additional Criteria
Florida	Scores of "Proficient" at the applicable grade level on each subtest of statewide English Language Proficiency Assessment	Scores on applicable FSA in ELA, as follows: a. For students in grades K-2, the statewide English Language Proficiency Assessment is the only assessment required; b. For students in grades 3-11, earning scores at or above the 50th percentile on the grade level FSA in ELA administered in the 2014-2015 school year; or c. For students in grades 11-12, a score on the 10th grade FCAT in Reading sufficient to meet applicable graduation requirements, or an equivalent concordant score pursuant to Section 1008.22, F.S.
Georgia	ACCESS Kindergarten: 5.0 CPL and no domain less than 5.0; Grades 1-12: 5.0 CPL and Literacy PL no less than 4.8	N/A
Hawaii	ACCESS 4.8 or higher Composite and minimum 4.2 Literacy	N/A
Idaho	IELA: Score at Early Fluent (4) or Fluent (5) Level and score EF or + on each domain	One of the following: a) Receive an Idaho Reading Indicator (IRI) score of at least a 3; b) Receive an Idaho Standards Achievement Test (ISAT) score that meets the "Basic" level; c) Demonstrate access to core content with a student portfolio using work samples from at least two (2) core content areas that demonstrate a Level 4 "Expanding" as defined by WIDA's Performance Definition rubrics and Can Do Descriptors.
Illinois	ACCESS 5.0 CPL and a Reading PL no less than 4.2 and a Writing PL no less than 4.2	N/A
Indiana	ACCESS 5.0 CPL	N/A
Iowa	I-ELDA: Level 6 (Full English Proficiency) Level 6 is at least two subskills at Level 5 and the other two at Level 4 or 5.	Students in grades 3-8, 10, and 11 must show proficiency on the state Title I AYP assessment.
Kansas	KELPA-P: Score "fluent" on Composite and in listening, speaking, reading, and writing for 2 CONSECUTIVE years. (KELPA-Placement proficient composite score used to determine year 1 and/or year 2 "fluent" score.)	N/A
Kentucky	ACCESS 5.0 Composite on Tier B or C AND 4.0 or higher in Literacy for grades 1-12	Kindergarten students cannot exit until after taking Grade 1 ACCESS.
Louisiana	ELDA: Level 5 (Full English Proficiency)	N/A

State	ELP Test Performance Standard	Additional Criteria
Maine	ACCESS 6.0 CPL	N/A
Maryland	ACCESS 5.0 CPL and Literacy PL no less than 4.0	N/A
Massachusetts	ACCESS 5.0 CPL and Literacy PL no less than 4.0	Student performance on MCAS, other academic assessments; student's academic grades; written observations and recommendations from classroom teachers
Michigan	ACCESS 5.0 CPL and no domain PL less than 4.5	K-2: Scores at or above grade level on state-approved reading assessment; grade-level proficiency on local writing assessment. 3-12: Scores Proficient or Advanced Proficient on the State reading assessment (M-STEP or ACT/SAT), or as defined by a state-approved reading assessment. Grade-level proficiency on local writing assessment.
Minnesota	ACCESS 5.0 CPL and no domain PL less than 4.0	Teacher recommendation, parental input, district may include state assessments
Mississippi	ACCESS 5.0 CPL on Tier B or C	Proficient or advanced on MCT2-Language Arts or passing English II MC
Missouri	ACCESS 6.0 CPL on Tier C **or** 5.0 CPL and a score of Basic on state content assessment, and some additional criteria, **or** 4.7 CPL on Tier C and a score of Basic on state content assessment, and some additional criteria	District benchmark examinations, writing performance assessments scored with the Missouri standardized rubric, writing samples, academic grades, agreement between ESL teacher, content teachers, other relevant staff and parents/guardians
Montana	ACCESS 5.0 CPL and Literacy PL no less than 4.0	Input from additional measures of reading, writing, or language development available from school assessments
Nebraska	ELDA: Composite Performance Level of 4: Advanced or Level 5: Full English Proficiency.	N/A
Nevada	ACCESS 5.0 CPL and Literacy PL no less than 5.0	N/A
New Hampshire	ACCESS 5.0 CPL and no domain PL less than 4.0	N/A
New Jersey	ACCESS 4.5 CPL on any tier	Multiple indicators that shall include, at a minimum: classroom performance; the student's reading level in English; the judgment of the teaching staff member or members responsible for the educational program of the student; and performance on achievement tests in English.
New Mexico	ACCESS 5.0 CPL on Tier B or C	N/A

State	ELP Test Performance Standard	Additional Criteria
New York	NYSESLAT: Grades K-12: Score "Commanding/Proficient" level; Grades 3-8: Score "Expanding/Advanced" level and 3 or above on the NYS ELA assessment within the same school year; Grades 9-12: Score "Expanding/Advanced" level and 65 or above on the Regents Exam in English within the same school year.	Alternate exit criteria require use of NYS ELA assessment or Regions English Exam.
North Carolina	ACCESS 4.8 CPL and Literacy PL no less than 4.0 on Tiers B or C for grades 1-12	N/A
North Dakota	ACCESS 5.0 CPL and no domain PL less than 3.5	N/A
Ohio	OTELA: Composite score of 5; or Composite score of 4, and subsequently completes a trial period of mainstream instruction and obtains a composite score of 4 or above during trial period of mainstream instruction.	N/A
Oklahoma	ACCESS 5.0 Overall Proficiency Level and 4.5 Literacy. Scores from Tiers B and C will be accepted as well as the accountability score from the Kindergarten ACCESS for ELLs® Test.	N/A
Oregon	ELPA: Level 5 (Advanced)	N/A
Pennsylvania	ACCESS 5.0 Composite on Kindergarten or Tier C.	BASIC on reading and math academic assessments and EITHER grades of C or better in core subjects OR scores comparable to BASIC on district-wide academic assessments.
Rhode Island	ACCESS 4.5 or higher Literacy score AND 5.0 or higher Comprehension AND Speaking proficiency score above a district established minimum	Any three of the following: Passing grades in all core content classes (as reflected on mid-year or end-of-year report card), or • ESL/bilingual Education teacher recommendation, or • At least two general education core content teacher recommendations, or • At least three writing samples demonstrating skill not more than one year below grade level, or • Score on a district reading assessment not more than one year below grade level as defined by the publisher or the district NOTE: Grades 1-12 (K not eligible for exit)
South Carolina	ELDA: Composite score of 5	N/A
South Dakota	ACCESS 4.7 CPL and Reading PL no less than 4.5 and Writing PL no less than 4.1	N/A

State	ELP Test Performance Standard	Additional Criteria
Tennessee	ACCESS 5.0 CPL and no domain PL less than 5.0	N/A
Texas	TELPAS or other state-approved ELP tests: Score "Fluent" on listening, speaking, writing.	Reading: 40th percentile or higher on Norm-Referenced Standardized Achievement Test (or STAAR, and an Agency-Approved Writing Test). For State of Texas Assessments of Academic Readiness (STAAR) English reading and English writing, the performance level for program exit is Level II (Satisfactory Academic Performance) or above.
Utah	ACCESS Level 5—Bridging	Teacher recommendation, sample writing, student grades
Vermont	ACCESS 5.0 CPL on Tier B or C and Literacy PL no less than 4.0	N/A
Virginia	ACCESS 5.0 CPL and Literacy PL no less than 5.0 on Tier C.	N/A
Washington	WELPA: Level 4: Transitional	N/A
West Virginia	WESTELL: Score "5" for two consecutive years	Scores at "Mastery" level or above on the Reading Language Arts (RLA) section of the WESTEST 2.
Wisconsin	ACCESS Composite 6.0 (K-12) or 5.0 or higher Composite and 5.0 or higher Literacy score is Automatic exit for students in grades 4-12.	Manual reclassification requires two pieces of evidence from academic work showing grade level language ability with parent, teacher and staff guidance. Evidence sources include:o District benchmark examinations (in multiple content areas);

• Writing samples or performance assessments scored with formal, standardized rubrics;

• State assessments at applicable grade levels; and

• Academic records such as semester or end-of-course grades. |
| Wyoming | ACCESS 5.0 CPL and no domain lower than 4.0 | N/A |

*Current as of 9/27/2015