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

November 2015

CCSSO
Council of Chief State School Officers

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

Authors:

Robert Linquanti, WestEd

H. Gary Cook, Wisconsin Center for Education Research

NOTE: Although this report highlights and develops several ideas from a national working session of educators with expertise on English learner issues, it reflects the views of the authors and not necessarily those of the working session participants or observers. The authors thank the working session participants for their commitment, engagement, and thoughtful discussion. They also thank Alison Bailey and Rita MacDonald for helping to plan and co-facilitate the working session, and for providing feedback on a draft of this document. Finally, they thank Fen Chou, Kenji Hakuta, Scott Norton, and several outside reviewers for their feedback and helpful comments.

CCSSO
One Massachusetts Avenue, NW, Suite 700 Washington, DC 20001-1431
Phone (202) 336-7000; Fax (202) 408-8072
Copyright © 2015 by the Council of Chief State School Officers, Washington, DC
All rights reserved.

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.[1] 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[2].

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

[1] 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.
[2] See below and Appendix B.

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.[3]

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.

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,

[3] See 34 C.F.R. § 200.20(f)(2).

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.[4]

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.

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

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)[5] has increased the focus on the EL reclassification rate as a key performance outcome in state and local educator accountability systems.[6] 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.[7] Thus a student can meet the "English

[5] For example, students in EL status for more than five or six years have been considered "long-term EL."
[6] 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.
[7] 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

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 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.[8] 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

number of states (29 states and the District of Columbia, up from 14 states in 2006–7) have instituted a single reclassification criterion.

[8] 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.

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.

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).[9] 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.[10] 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:

[9] See AERA/APA/NCME Standard 12.10, p. 198; Standard 12.13, p. 199; and Standard 13.9, p. 213. Discussed further below.
[10] 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.[11] 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

[11] See Appendix C. for Working Session Agenda, and Appendix D. for guiding questions.

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 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

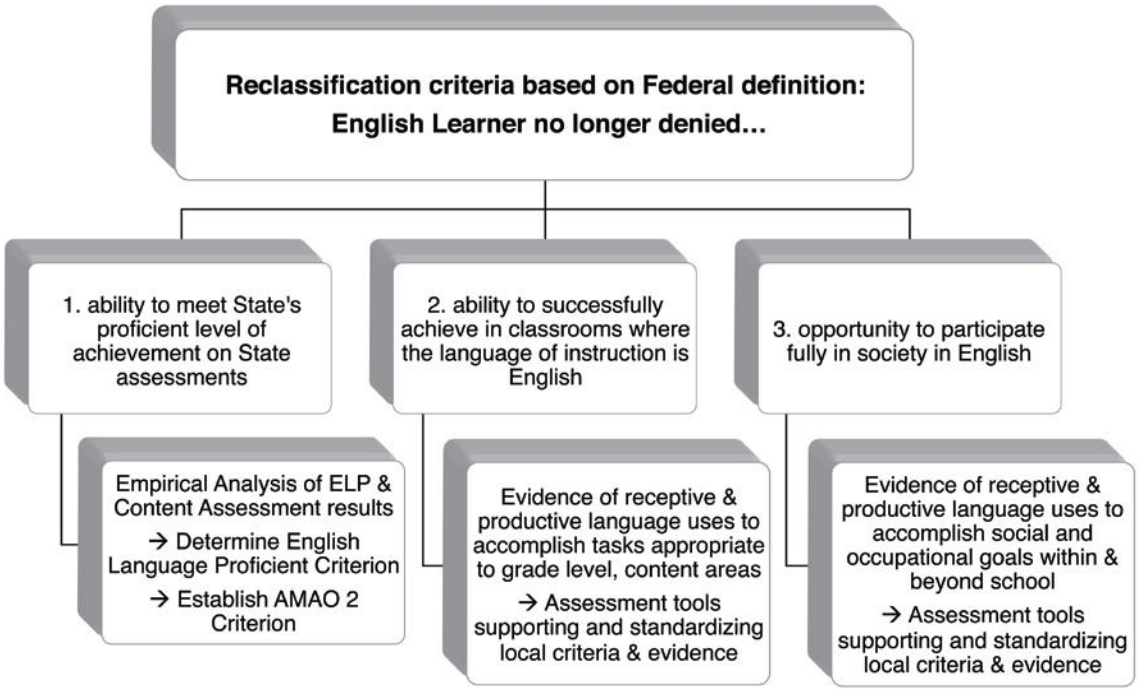
to ensure the rigor and fairness of any process used, with appropriate resources for training and implementation. Consortia, states, and districts should also anticipate, prevent, and address any unintended negative consequences resulting from reclassification policies and procedures. This includes regular 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,[12] 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.

[12] 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.[13] 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.[14]

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

[13] 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.
[14] See Linquanti & Cook (2013), pp. 11–14, 25–31, and Cook & MacDonald (2012) for further discussion.

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.[15]

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

[15] See Heritage, Walqui, & Linquanti (2015), and van Lier & Walqui (2012) for further discussion.

for their ELP assessment composite-score-weighting approach and to provide evidence of the efficacy of the identified weighting procedure.[16]

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.[17] 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).

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.

---

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

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.[18]

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.[19] Also, renewed interest in

[18] 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.

[19] 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

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.[20] 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.)

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)).

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).

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

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.[21]

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

[21] 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.

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.

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"	1 R-FEP	2 R-FEP	3 R-FEP*
	"English-Proficient" within CSEM	4 R-FEP	5 R-FEP*	6 EL*
	Not "English-Proficient"	7 EL*	8 EL	9 EL

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.[22] The columns illustrate three similar outcomes on the classroom-based ELP criterion.[23]

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.

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 <u>across</u> 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

[22] 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.
[23] An analogous "borderline" margin of error around the "English-proficient" score would need to be established for the observation protocol outcome.

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.[24] 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 with disabilities that may be unable to meet a specific

[24] See Appendix B. .

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.[25] 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.[26] 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.

[25] See Cook (2013).
[26] See Cook (2014).

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.[27] The variability within and across 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

[27] See pp. 2, 6–7 above.

performance and educational outcomes.[28] 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

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

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 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%20Classification%20Working%20Paper_FINAL_01%2026%202015.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 English 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/**227Report**b.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/epaa/v14n20/.

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 http://www2.ed.gov/policy/speced/guid/idea/memosdclters/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 http://k12center.org/rsc/pdf/elp_policy_report_12_17_14.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.

State	ELP Test Performance Standard	Additional Criteria
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
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

State	ELP Test Performance Standard	Additional Criteria
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
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

State	ELP Test Performance Standard	Additional Criteria
Pennyslvania	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
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

State	ELP Test Performance Standard	Additional Criteria
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); o Writing samples or performance assessments scored with formal, standardized rubrics; o State assessments at applicable grade levels; and o 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**

Appendix C. Working Session Agenda

Working Session Goals:

Provide input for expanded guidance for states and consortia to support LEAs in making decisions about reclassification/exit of EL students

- Explore criteria and methods that examine linguistic contribution to academic performance, classroom achievement, and career/societal participation
- Discuss & provide input on assessment tools to support and help standardize local reclassification criteria that are relevant, reliable, valid, and comparable
- Discuss & provide input on within-state, cross-state, and within- and cross-consortium EL reclassification criteria and procedures

Session Leads: Robert Linquanti and Gary Cook

Facilitators: Alison Bailey and Rita MacDonald

Day 1 (September 23)

Time	Activity
9:30 AM – 10:00 AM	Light Refreshment
10:00 AM – 10:30 AM	Welcome, Introductions, Context (4 stages, prior sessions and outcomes), Review of Working Session Goals & Agenda, & Discussion Framework
10:30 AM – 12:00 PM **Session 1** (Whole Group)	**How might states, consortia, and LEAs move toward more common criteria & methods to reclassify ELs?** (CCSSO Guidance Framework Stage 4) • Define Problem/Opportunity Space • Surface and capture existing and emerging issues & opportunities at state and local levels • Share questions, observations, reflections • Assess consensus on 3 dimensions of 9101(25)
12:00 PM – 12:30 PM	Lunch
12:30 PM – 1:15 PM **Session 2** (Whole Group)	**How might states/consortia establish an "English proficient" performance standard on Title III ELP assessment, related to achievement on Title I academic assessment?** [9101(25)(D)(i)] • Illustrate methods with empirical data from large-scale ELP and academic content assessments • Identify policy & technical issues & opportunities at consortium, state and local levels

1:15 PM – 3:15 PM **<u>Session 3</u>** (Initial Whole Group) (4 Breakout Groups)	**What criteria and tools might educators use to examine language uses contributing to the "ability to achieve in classrooms where language of instruction is English"?** [9101(25)(D)(ii)] • Examine evidence-gathering tools and practices related to receptive and productive language uses for grade-level content practices • Identify issues & opportunities in supporting and standardizing local criteria and evidence-gathering • Recommend guidelines regarding tools & practices (reliability, validity, comparability) • List opportunities for adaptation, collaboration, piloting, and validation research
3:15 PM – 3:45 PM	**Break**
3:45 PM – 5:00 PM **<u>Session 4</u>** (Whole Group)	**Whole Group Facilitated Discussion** • Takeaways from working session • Policy and practice issues and opportunities • Preview of Day 2

Day 2 (September 24)

Time	Activity
8:00 AM – 8:30 AM	Breakfast
8:30 AM – 9:00 AM **<u>Session 5</u>** (Whole Group)	**What did we learn yesterday? What other questions and issues occur?** • Confirm key learnings and capture new questions and issues
9:00 AM – 10:30 AM **<u>Session 6</u>** (Initial Whole Group) (4 Breakout Groups)	**What criteria and tools might educators use to examine language uses contributing to the "opportunity to participate fully in society" using English?** [9101(25)(D)(iii)] • Highlight language uses for deep learning and 21st century competencies that help students accomplish social and occupational goals within and beyond school • Examine and share national, state, local evidence-gathering tools and practices related to receptive and

Time	Activity
9:00 AM – 10:30 AM **<u>Session 6, cont'd.</u>**	productive language uses for social and occupational goals • Identify issues and opportunities in supporting and standardizing local criteria and evidence-gathering • Recommend guidelines regarding tools and practices (focus on reliability, validity and comparability) • List opportunities for adaptation, collaboration, piloting, and validation research
10:30 AM – 11:00 AM	Break
11:00 AM – 12:00 PM **<u>Session 7</u>** (Initial Whole Group) (4 Breakout Groups)	**What policy and legal issues/opportunities hinder/ support development and use of criteria and tools for EL reclassification?** • Consider current regulations, legal interpretations, and enforcement (e.g., Federal and state program monitoring & compliance, Federal consent decrees, state education codes, etc.) • Examine strategies to strengthen reliability, validity, and comparability while maintaining local control of educational decision-making • Identify consensus guidance points, including "non-negotiables," "optionals," and divergent views
12:00 PM – 12:45 PM	Lunch
12:45 PM – 2:00 PM **<u>Session 8</u>** (Initial Whole Group) (4 Breakout Groups)	**How might consortia, states, and local educators strengthen comparability of EL reclassification within and across ELP assessment consortia and/or standalone states?** • Share innovative EL reclassification policies, tools, and plans • Consider technical challenges and opportunities (e.g., consistent implementation, interrater reliability, online calibration training, etc.) • Review & discuss potential methods to establish comparability of reclassification criteria & evidence • Explore potential for common plans and activities • List opportunities for collaboration, piloting, and validation research
2:00 PM – 2:45 PM **<u>Session 9</u>** (Whole Group)	**Whole Group Facilitated Discussion** • Policy and practice issues and opportunities • Takeaways from entire working session

Time	Activity
2:45 PM – 3:00 PM	Clarify next steps & Wrap-up

Appendix D. Notetaking Guide & Discussion Questions

Day 1 (10:30 AM-12 PM): Whole-Group Session 1

How might states, consortia, and LEAs move toward more common criteria & methods to reclassify ELs?

Notetaking:
Key problems & challenges (local, state, consortium):

- Key opportunities (local, state, consortium):
- Questions I have:
- Observations/reflections:
- Degree of consensus on reclassifying ELs based on 3 dimensions of 9101(25):

Whole-Group Discussion Questions:

1. What are key issues or concerns regarding EL reclassification in your state, district, or schools?
2. To what extent is there consensus on reclassifying ELs based on the three dimensions of 9101(25)?

Day 1 (12:30 PM-1:15 PM): Whole-Group Session 2

How might states/consortia establish an "English proficient" performance standard on the Title III ELP assessment related to achievement on Title I academic assessments?
[9101(25)(D)(i)]

Notetaking:

- Questions about methods using large-scale ELP and academic content assessment data:
- Policy issues in setting "English proficient" performance standard (consortium, state, and local levels):
- Technical issues in setting "English proficient" performance standard (consortium, state, and local levels):

Whole-Group Discussion Questions:

1. What technical or policy questions do you have about these proposed methods?
2. These methods take account of but do not require a set level of academic content test performance. What risks and benefits do you see in this?

Day 1 (1:15 PM-3:15 PM): Initial Whole-Group & 4 Breakout Groups - Session 3

What criteria and tools might educators use to examine language uses contributing to the ability to "achieve in classrooms where the language of instruction is English"?
[9101(25)(D)(ii)]

Focus:

- Examine evidence-gathering tools and practices related to receptive and productive language uses for grade-level content practices
- Identify issues & opportunities in supporting and standardizing local criteria and evidence-gathering
- Recommend guidelines regarding tools & practices (reliability, validity, comparability)
- List opportunities for adaptation, collaboration, piloting, and validation research

DISCUSSION QUESTIONS:

1. What types of tools and practices could be used to gather evidence of students' academic uses of language?
2. What key issues must be addressed in supporting and standardizing local criteria and evidence-gathering? (e.g., feasibility, capacity, multiple/conflicting purposes)
3. What guidelines would you recommend regarding the use of tools and practices for this purpose?
4. What opportunities exist to collaborate on adapting, piloting, and researching the use of tools for this purpose?

Day 2 <u>(9:00 AM-10:30 AM)</u>: <u>Initial Whole-Group & 4 Breakout Groups - Session 6</u>

What criteria and tools might educators use to examine language uses contributing to the "opportunity to participate fully in society" using English? [9101(25)(D)(iii)]

Focus:

- Highlight language uses for deep learning and 21st century competencies that help students accomplish social and occupational goals within and beyond school
- Examine and share national, state, local evidence-gathering tools and practices related to receptive and productive language uses for social and occupational goals
- Identify issues and opportunities in supporting and standardizing local criteria and evidence-gathering
- Recommend guidelines regarding tools and practices (focus on reliability, validity and comparability)
- List opportunities for adaptation, collaboration, piloting, and validation research

DISCUSSION QUESTIONS:

1. What kinds of language uses contribute to "opportunity to participate fully in society" using English? (i.e., for deep learning and 21st century competencies that help students accomplish social and occupational goals within and beyond school)
2. What features in currently available evidence-gathering tools and practices support educators to examine EL students' language uses for social and occupational goals?
3. What new evidence-gathering tools and practices will need to be developed to support educators in examining EL students' language uses for social and occupational goals?
4. What key issues must be addressed in supporting and standardizing local criteria and evidence-gathering? (e.g., feasibility, capacity, multiple/conflicting purposes)

5. What guidelines would you recommend regarding tools and practices?
6. What opportunities exist to collaborate on creating, adapting, piloting, and researching tools and practices?

Day 2 (11:00 AM - 12:00 PM): Initial Whole Group and 4 Breakout Groups - Session 7

What policy and legal issues/opportunities hinder/support development and use of criteria and tools for EL reclassification?

Focus:

- Consider current regulations, legal interpretations and enforcement (e.g., Federal and state program monitoring & compliance, Federal consent decrees, state education codes, etc.)
- Examine strategies to strengthen reliability, validity, and comparability while maintaining local control of educational decision-making
- Identify consensus guidance points, including "non-negotiables," "optionals," and divergent views

DISCUSSION QUESTIONS:

1. What do you see as a particularly important policy and legal challenges to supporting common reclassification criteria and tools in your educational context?
2. What opportunities/leverage points do you see to support common reclassification criteria and tools in your educational context?
3. From your group discussion, what are the agreed-upon "non-negotiables" in producing further guidance for establishing common EL reclassification criteria and tools?

Day 2 (12:45 PM - 2:00 PM): Initial Whole Group and 4 Breakout Groups – Session 8

How might consortia, states, and local educators strengthen comparability of EL reclassification within and across ELP assessment consortia and/or standalone states?

Focus:

- Share innovative EL reclassification policies, tools, and plans
- Consider technical challenges and opportunities (e.g., consistent implementation, interrater reliability, online calibration training, etc.)
- Review & discuss potential methods to establish comparability of reclassification criteria & evidence
- Explore potential for common plans and activities
- List opportunities for collaboration, piloting, and validation research

DISCUSSION QUESTIONS:

1. What are key issues/concerns regarding reclassification comparability within your educational context?
2. How do you currently ensure comparability?
3. What is the minimal level of evidence needed to assure meaningful comparability in your context?
4. What approaches could help districts/states/ consortia establish or improve comparability?