

Key Issues and Opportunities in the Initial Identification and Classification of English Learners



ETS RR-16-09

Alexis A. Lopez • Emilie Pooler • Robert Linqanti

ETS Research Report Series

EIGNOR EXECUTIVE EDITOR

James Carlson
Principal Psychometrician

ASSOCIATE EDITORS

Beata Beigman Klebanov
Senior Research Scientist

Heather Buzick
Research Scientist

Brent Bridgeman
Distinguished Presidential Appointee

Keelan Evanini
Research Director

Marna Golub-Smith
Principal Psychometrician

Shelby Haberman
Distinguished Presidential Appointee

Anastassia Loukina
Research Scientist

John Mazzeo
Distinguished Presidential Appointee

Donald Powers
Managing Principal Research Scientist

Gautam Puhan
Principal Psychometrician

John Sabatini
Managing Principal Research Scientist

Elizabeth Stone
Research Scientist

Matthias von Davier
Senior Research Director

Rebecca Zwick
Distinguished Presidential Appointee

PRODUCTION EDITORS

Kim Fryer
Manager, Editing Services

Ayleen Gontz
Senior Editor

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

RESEARCH REPORT

Key Issues and Opportunities in the Initial Identification and Classification of English Learners

Alexis A. Lopez,¹ Emilie Pooler,¹ & Robert Linqunti²

¹ Educational Testing Service, Princeton, NJ

² WestEd, San Francisco, CA

English learners (ELs) require access to appropriate instructional services that match their strengths and needs to have an equal opportunity to achieve the same academic standards as other students. If ELs are not properly identified and classified, they may be excluded from services that would better help them meet high academic standards, and they may not receive optimal supports to advance both linguistically and academically. Thus, it is critical to accurately and properly identify newly enrolled students who are in need of supplemental instruction and support. The purpose of this report is to identify issues in current initial English-language proficiency assessment practices used to determine initial EL classification and offer recommendations to strengthen the policies, process, and tools used in the identification and classification of ELs.

Keywords Initial English language proficiency assessments; initial identification; initial classification; English learners; young learners

doi:10.1002/ets2.12090

This paper is the fourth in a series from Educational Testing Service (ETS). The series is intended to provide theory- and evidence-based principles and recommendations for improving next generation English language proficiency (ELP) assessment systems, policies, and practices and to stimulate discussion on better serving K–12 English learners (ELs). The first paper articulated a high-level conceptualization of next-generation ELP assessment systems (Hauck, Wolf, & Mislevy, 2016), the second paper addressed accessibility issues in the context of ELP assessments for ELs and ELs with disabilities (Guzman-Orth, Laitusis, Thurlow, & Christensen, 2016), and the third paper addressed issues related to summative ELP assessments that emerged from the presentations and discussions at the English Language Proficiency Assessment Research working meeting (Wolf, Guzman-Orth, & Hauck, 2016). The present paper focuses on a key concern within such systems—the initial identification and classification of ELs. In this paper, *initial identification* refers to identifying newly enrolled students as potential ELs and *classification* refers to confirming whether these students are ELs.

Current policy, research, and practice all point to the importance of English learner (EL) identification and classification issues. Ensuring that EL students receive the specialized instructional support services to which they are legally entitled is predicated on properly identifying and classifying such students. The United States Department of Justice (DOJ) Civil Rights Division and the United States Department of Education (DE) Office for Civil Rights have recently published consolidated guidelines that highlight the long-standing legal basis for ensuring the adequacy of methods used to identify and classify ELs (U.S. DOJ & U.S. DE, 2015). In conjunction with these guidelines, the U.S. DE Office of English Language Acquisition has published an English Learner Tool Kit, which provides new guidance to state and local education agencies on how to adequately identify ELs (U.S. DE, 2015). This federal guidance was released as a result of prior research that revealed significant variation within and across states in both the instruments and decision-making processes used to identify ELs (Abedi, 2008; Cook & Linqunti, 2015; Linqunti & Bailey, 2014; National Research Council [NRC], 2011; Wolf et al., 2008). The variation issue in EL identification was explicitly pointed out by a panel convened by the NRC (2011) as follows:

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

Corresponding author: A. Lopez, E-mail: alopez@ets.org

In this paper, we focus on issues related to the EL identification process and use of initial English language proficiency (ELP) assessments (often referred to as ELP screeners). Although other instruments are used in the identification process (e.g., home language survey [HLS]), the initial ELP assessment is the key factor in determining initial EL classification. Determining EL status is an important decision, so the stakes associated with initial ELP assessments are significant. If incoming students score below a predetermined cut point on this assessment, they are classified as ELs and placed into specialized educational services. Even if parents decline such services, their students are classified as ELs and required to take an annual ELP assessment to measure their progress in English language development (U.S. DOJ & U.S. DE, 2015).

It is important to acknowledge that dichotomizing language-minority students as EL or non-EL fails to capture the highly complex reality of language development, a phenomenon that is in fact nonlinear and multidimensional. This is particularly so for students who enter school as emergent bilinguals with varying degrees of proficiency in English and one or more other languages. Nevertheless, current state and federal statutes, regulations, and case law require educators to explicitly identify, classify, and serve those language-minority students who qualify for protected-class status, that is, as officially designated ELs entitled to appropriate services. This paper focuses on ways to improve the practice of accurately and appropriately classifying such students.

Specifically, the goals of this paper are twofold: (a) to identify issues in current initial ELP assessment practices used by states and districts to determine initial EL classification and (b) to offer recommendations to help state and local educators and policymakers strengthen the processes and tools used in the identification and classification of ELs. We start by describing the context of state and local policies and practices used to identify potential ELs and assess them for initial classification as either EL or initially fluent English proficient (IFEP). We then review key issues related to current policies and procedures and offer recommendations for improvement.

Current English Learner Identification and Classification Process

Under federal law, schools are required to identify and classify ELs within 30 days of their enrollment in a K–12 public school. If a student enrolls in a new school after the beginning of a school year, schools must notify parents of EL classification and placement decisions within 2 weeks (No Child Left Behind Act, 2001). The U.S. DE's Office for Civil Rights (OCR) requires that a well-defined process be in place to determine which students are ELs and therefore entitled to receive supplemental instruction and supports (U.S. DOJ & U.S. DE, 2015). However, the OCR does not specify what tools a well-defined identification process must employ.

In virtually all states, the identification of ELs occurs via a two-step process: (a) identify incoming students as potential ELs and (b) determine if the student should be classified as EL through an initial assessment of ELP (Linquanti & Cook, 2013). Although it is not a federal requirement, the most common tool used to identify potential ELs is an HLS. The primary purpose of the HLS is to identify students who either use or are exposed to a language other than English. Currently, virtually all states in the United States use some form of HLS to identify potential ELs (Bailey & Kelly, 2013; Wolf *et al.*, 2008). Although HLS questions vary from state to state (and in some states, from district to district), they are generally designed to gather information about (a) the language(s) a student speaks and understands and (b) the language most frequently used at home by the student, the student's parents or guardians, or other adults in the home.

If the HLS indicates exposure to or use of a language other than English in the student's environment, then the student is identified as a potential EL and is administered an initial ELP assessment. Initial ELP assessments are used to estimate the student's ELP level (e.g., not English fluent, initially English fluent) and thereby classify the student as an EL or non-EL. Given that classification decisions need to be made within 30 days of students' enrollment, the initial ELP assessment is usually brief and relatively simple to administer and score locally.

Initial ELP assessment results may also be used to assist educators in placing confirmed EL students into a language instruction educational program that is appropriate to their assessed ELP level. For example, if an initial ELP assessment provides scores placing EL students into several levels of proficiency, EL students whose scores place them at a beginning level might receive different services than those who are placed at an intermediate level.¹

As defined in federal statute and civil rights case law, all students classified as ELs are entitled to specialized instructional services aimed at developing their academic ELP and ensuring they have meaningful access to grade-level academic content through carefully scaffolded instruction while their English proficiency is developing (Hakuta, 2011; Linquanti & Cook, 2013). Research on the validity and fairness of existing identification and classification policies, processes, and

assessment tools has highlighted several challenges. We turn next to an examination of these key challenges and our recommendations for addressing them.

Key Issues and Recommendations

To date, empirical research and policy review studies (e.g., Abedi, 2008; Bailey & Kelly, 2013; Cook & Linqunti, 2015; Linqunti & Bailey, 2014; Wolf et al., 2008) have identified several issues regarding how students are identified as potential ELs and initially classified as ELs in K–12 public schools in the United States. In this section, we highlight some of these issues and offer recommendations intended to improve current policies and practices.

Inconsistent Use of HLS to Identify Potential English Learners

Variations of the HLS are used for identifying potential EL students in virtually all states. Although schools are required to have procedures in place for identifying potential EL students, no federal guidelines exist regarding what questions should be asked or how the HLS should be used in the initial EL identification process. As a result, current implementations of the HLS vary greatly in purpose (why they are used), targeted constructs (what the questions intend to measure), questions (number and focus), procedures (how they are administered), and interpretation (conclusions and decisions made; Linqunti & Bailey, 2014). Even within many states, standardized processes and tools are lacking. For example, some states provide an HLS template or recommended HLS questions that school districts can consider, but districts are allowed to develop and implement their own language surveys. Several studies have questioned the validity and reliability of current HLS practices, raising issues related to how HLSs are developed, the appropriateness and relevance of the questions in HLSs, and how the information in HLSs is interpreted to determine potential EL status.²

Recommendation 1: States and ELP Assessment Consortia Should Work to Improve the Design and Implementation of HLS Used to Identify Potential ELs

Accurately identifying potential ELs is the first step in properly identifying the population that may merit services under protected-class status as well as reducing the misclassification of students as ELs or non-ELs. Recent studies and federal guidance have proposed changes to improve the identification of potential ELs through strengthening the design and use of the HLS (Linqunti & Bailey, 2014; U.S. DE, 2015). These documents provide concrete recommendations on how to effectively develop, implement, and interpret an HLS to ensure accurate results. Key examples include the following:

- States should provide clear guidance on HLS purposes and uses, ensure consistent implementation of initial EL identification processes and tools, and establish clear rules for interpreting HLS results.
- HLS questions should focus on the intended target constructs: The language(s) the student currently speaks and understands, the student's degree of current English language use, and the students' degree of current English language exposure.
- The HLS should be given in multiple home languages and in written and oral modes for parents with varying levels of literacy; translations should be standardized and validated.
- States and consortia should conduct studies to monitor the implementation and proper use of the HLS and to validate initial EL identification decisions.

Variation in Conceptualization and Design of Initial ELP Assessments

A clear lack of consensus is evident on what needs to be measured in initial ELP assessments. Currently, different initial ELP assessments are used to determine the classification of potential ELs across and, in some instances, within states. According to recent surveys (e.g., Cook & Linqunti, 2015; NRC, 2011), 23 states allow the use of multiple initial ELP assessment instruments to make EL classification decisions. The NRC panel concluded that there is no evidence of the comparability of conceptualization of ELP on these different instruments, nor is there evidence of the comparability of the performance cut scores used to determine who is considered IFEP. This finding makes it exceedingly difficult to compare results of initial ELP assessments across states—or across districts within states—that use different instruments. That is, a student who is classified as a non-EL on one assessment may be classified as an EL on another. There are also qualitative

differences among the varieties of existing initial ELP assessments. For example, assessments may not define or measure the ELP construct in comparable ways, operationalizing different conceptions of language proficiency (e.g., prioritizing social vs. academic uses of language), weighting language domains (listening, speaking, reading, and writing) differently to arrive at a composite ELP score, or using different groups of ELs as norming populations (Cook & Linqunti, 2015; Faulkner-Bond, Shin, Wang, Zenisky, & Moyer, 2013; Linqunti & Cook, 2013).

Moreover, different initial ELP assessments may not be comparably aligned to a state's ELP standards. Without demonstrated alignment to a state's ELP standards, the constructs of screeners may be poorly defined or may extend to content other than ELP. As a consequence, results on the initial ELP assessment may differ substantially from those on the state's annual ELP assessment.

Additionally, if the initial ELP assessment is to be used for both placement and identification of EL status, careful alignment to ELP standards helps to ensure that the test is able to measure the full range of language proficiency levels described in the standards. Because students can arrive at a school at any grade (from K to 12) and with varied levels of English proficiency (from beginner to advanced), ELP standards typically cover a range of language proficiency levels at each grade. Therefore, different initial ELP assessments need to be developed for different grades or grade bands. For example, an initial ELP assessment for high school students would need to include reading comprehension questions based on texts with age-appropriate topics, including some texts that are simplified and some that represent grade-level complexity.

Variation is also apparent in the way that initial ELP assessments conceptualize the construct of ELP for EL identification purposes. Considerable research has been conducted recently to conceptualize the language used in schools, including both academic and social language (e.g., Bailey, 2007; Schleppegrell, 2004). However, there is no consensus regarding the ways that academic and social language should be operationalized and measured on an initial ELP assessment (Wolf *et al.*, 2008).

Furthermore, test developers find it challenging to develop appropriate content for initial ELP assessments due to the varying characteristics of the target test takers. Students who take initial ELP assessments are diverse in their language background, educational experiences, literacy development, culture, ethnicity, and socioeconomic circumstances (Batalova, Fix, & Murray, 2005). This diversity can be seen throughout the United States, both in states that have long had large numbers of EL students (e.g., California, Texas, New York, Florida, and Illinois) and in states where the presence of significant numbers of EL students is a more recent phenomenon (e.g., North Carolina, Georgia, and Washington; Pandya, Batalova, & McHugh, 2011). Although approximately 71% of EL students across the United States are from Spanish language backgrounds, the languages spoken by EL students are diverse, and this linguistic diversity varies across states (Ruiz Soto, Hooker, & Batalova, 2015). At the beginning of the school year, when most initial ELP assessments are given, many students who have recently arrived in the United States are not yet familiar with US school contexts and culture. Test developers are faced with the challenge of designing and developing initial ELP assessments that allow all students an equal opportunity to demonstrate their ELP, regardless of their cultural and linguistic backgrounds.

Another challenge arises from the use of tests as initial ELP assessments when the tests have not been designed specifically for that purpose. As a result of such practices, educators may make invalid interpretations of scores on initial ELP assessments and students may be misclassified as ELs (false positives) or as IFEP (false negatives; Bailey & Carroll, 2015). Next, we offer three recommendations regarding the construct definition and development of initial ELP assessments.

Recommendation 2–1: Test Developers, States, and Consortia Should Clearly Define the Constructs to Be Measured in Initial ELP Assessments

Properly defining the constructs to be measured in an initial ELP assessment is the fundamental basis for ensuring accurate and valid EL classification decisions. Test developers should therefore define the constructs of initial ELP assessments with a high degree of precision. That is, test designers should clearly conceptualize and operationalize a definition of ELP in school settings, as well as what it is to be an EL. Test designers should also provide clear ELP performance level descriptors, especially regarding the IFEP level. These performance level definitions should enable users to distinguish between students who have sufficient initial ELP from those who do not (and should therefore be classified as ELs).

We also recommend that test developers articulate the language functions and skills that are being measured in initial ELP assessments and provide rationales for including each of them as part of the test construct. Test designers should

also provide evidence that the initial ELP assessment is aligned to state ELP standards. Moreover, test designers should specify which language domains (listening, speaking, reading, and writing) and which target-language-use domains (e.g., social/interpersonal, school navigational, general-academic, and discipline-specific)³ are included in the initial ELP assessment and provide a clear rationale for these decisions. Finally, test designers should articulate and provide adequate justification for decisions regarding how each language domain and target-language-use domain is weighted (i.e., equally or emphasizing certain ones over others).

Recommendation 2–2: States and Consortia Should Work to Ensure Better Interpretation of Scores on Initial ELP Assessments

Initial ELP assessments should be designed according to their intended purposes and targeted constructs. Intended purposes include initially classifying EL students (i.e., confirming or disconfirming the EL status of potential EL students) and, potentially, providing a defensible classification of the student's ELP level for program placement purposes. In order to help prevent misclassification of students, it is important to ensure that initial ELP assessments be developed specifically for initial classification purposes. Furthermore, it is critical to conduct validation studies to gather empirical evidence on the technical qualities and actual uses of initial ELP assessments. The purpose of examining validity evidence is to evaluate the justifiability of the use of any assessment instrument or performance criterion that distinguishes an EL from a fluent English proficient student.

It is also important to validate initial ELP assessment results with external evidence. In particular, external evidence should demonstrate a clear relationship between the initial ELP assessment and the full, annual ELP assessment with respect to the constructs assessed and the classification outcomes (see Cook & Liguanti, 2015). We suggest including both EL students and fluent English-proficient students in prototyping and pilot studies during the assessment development stages to determine which types of assessment tasks provide meaningful information that can be used to distinguish between English proficient and non-English proficient students (Wolf & Lopez, 2014a).

Recommendation 2–3: Test Developers Should Design and Develop Initial ELP Assessments That Are Sensitive to Student Cultural Diversity

To fully accommodate the cultural diversity among the student population, we recommend that test developers be sensitive to cultural references in initial ELP assessments. That is, test developers should not include inappropriate references or content that could have unintended meanings for students (National Association for the Education of Young Children [NAEYC], 2005; Padilla, 2001). Assessments that are sensitive to cultural diversity use assessment materials (e.g., texts, illustrations, contexts, and topics) that acknowledge diversity (e.g., of cultures, gender, ethnicity, race, and so on) and represent nonstereotypical views of different cultural groups (Santos, 2004). Moreover, initial ELP assessments should not include content that may be biased (i.e., that unfairly advantages one group over another), nor should they require that students have prior knowledge of American culture or familiarity with the US school contexts (Padilla, 2001). Thus, we suggest that test developers strive to use contexts that are familiar to EL students with different cultural and linguistic backgrounds.

Inadequacies in the Initial Assessment of Young Learners⁴

The large majority of newly entering students who go through the initial identification and classification process are in kindergarten and Grade 1. According to the Urban Institute, approximately one quarter of children of immigrants ages 5–8 are EL students (Fortuny, Hernandez, & Chaudry, 2010). In California, approximately 37% of students entering kindergarten in 2014–2015 were identified as EL students (California Department of Education, 2015).

Assessing young learners is challenging because young learners are often first-time test takers who may be new both to school and to the procedures of standardized testing, resulting in conditions that increase the likelihood of measurement error. For example, results of measurement of a young student's oral skills may be influenced by the degree of comfort the child feels toward the examiner. Another challenge in assessing young language learners stems from the fact that these students are rapidly developing their language proficiency and preliteracy skills in both their home language and in the second language, and individual students' developmental trajectories can vary widely. This fact makes the initial

EL identification process extremely complicated from a psychometric standpoint. Therefore, clarity in what exactly is being measured, and how, is critically important with young students whose language and literacy development are in their early stages. A related concern is that teachers and other school personnel who often administer ELP assessments to young learners require special training in language development and the assessment of young learners (Wong Fillmore & Snow, 2002).

Without appropriate initial ELP assessments for young learners, educators are less likely to make valid classification decisions, leading to underidentification or overidentification of students as ELs. Next, we offer a multipart recommendation on strengthening initial ELP assessment practices for young learners.

Recommendation 3: Test Developers and States Should Ensure That Appropriate Initial ELP Assessment Practices Are Utilized with Young Learners

To accurately and validly assess young learners, it is critical to ensure that initial ELP assessments are developmentally appropriate (Espinosa & García, 2012; NAEYC, 2005). Thus, we suggest that test developers give special consideration to developmental and contextual factors that might impact young learners' performance on these assessments (Bailey, 2008; McKay, 2006; NAEYC, 2005). Test developers should collect empirical evidence to support the validity, reliability, and appropriateness of these assessments for the purpose of classifying young learners as EL students.

Equally important is having initial ELP assessments that are age and grade appropriate (Bailey, 2008; NAEYC, 2005). For example, entering kindergartners and first graders—whether ELs or not—are very different from each other in their assessment needs and level of language development. Thus, separate assessments will very likely be required for these two age/grade levels. Given that some newly developed ELP standards (e.g., California Department of Education, 2012; Council of Chief State School Officers, 2012) clearly call for students to complete certain tasks with prompting and support (especially in kindergarten and Grade 1), we suggest including age-appropriate scaffolding and support in initial ELP assessments (Wolf & Lopez, 2014b).

To ensure that assessment practices are appropriate and adequate for young learners, we recommend that all test administrators who will assess young learners be adequately trained to be responsive to these students' unique characteristics (Espinosa & García, 2012; NAEYC, 2005; Santos, 2004). This training will help administrators involved in the assessment of young learners to conduct accurate and fair assessments and to be sensitive and responsive to cultural and linguistic differences.

It is also important for all stakeholders to clearly understand the constructs to be measured in initial ELP assessments for young learners and justify decisions regarding the weighting of language skills or domains. Given that young learners are in the early stages of developing literacy skills in their home language and/or English, we recommend that initial ELP assessments for these learners place greater weight on oral skills (listening and speaking) than on literacy skills (reading and writing). At the same time, test developers should gather sufficient evidence to demonstrate that high oral language scores on initial ELP assessment are sufficiently predictive of foundational literacy skills for students who are potential ELs at these primary grade levels (Cook & Linquanti, 2015).

As an additional point to ensure developmental appropriateness, initial ELP assessments for young learners should be relatively brief or administered in short sessions (Epstein, Schweinhart, DeBruin-Parecki, & Robin, 2004). This will help minimize the risk of young learners becoming disengaged during the assessment process. In order to keep young learners engaged, we recommend designing tasks that include content, topics, and stimuli that are relevant to young learners (NAEYC, 2005; Pooler & Lopez, 2013; Wolf & Lopez, 2014b). For example, assessments for this population should include content and topics likely to be familiar to young learners, such as colors, family, numbers, clothes, and animals. These assessments should also include ample visual stimuli such as pictures, illustrations, and if feasible, videos and animation.

Another promising strategy for engaging young learners during the assessment process is the use of contextualized tasks. Contextualized tasks are comprised of a series of related items that unfold within an appropriate and purposeful context, allowing students to engage multiple language skills (Pooler & Lopez, 2013). Technology-enhanced tasks have the potential to be more engaging for young learners, as they can more easily incorporate contextualized tasks with the use of animation and interactive features rather than static images (Wolf *et al.*, in press). In designing technologically enhanced assessment tasks for young learners, however, special attention should be paid to students' level of experience with the relevant technology. It is critical to conduct empirical studies to ensure that all technology features are intuitive

and easy to use for all students, including those with limited formal schooling or limited previous exposure to technology and assessments.

Finally, empirical studies are needed to validate the effectiveness of other dimensions of practice in assessing young learners' language proficiency. Some of these dimensions include determining adequate administration formats (e.g., individual, small group), appropriate item or task types, and appropriate ways to provide directions to young learners.

Lack of Appropriate Initial ELP Assessments to Classify Students with Disabilities as English Learners

Among the group of potential ELs who will be taking initial ELP assessments are both students with confirmed disabilities and students who may have an undiagnosed disability.⁵ Initial ELP assessments must be fair and valid for these students, just as for other groups of potential ELs. Unfortunately, potential EL students who also have a diagnosed disability experience variable accommodations depending on their state of residence and the ELP assessment(s) that are used in that state (Albus & Thurlow, 2008). Consider a student in need of a particular accommodation (e.g., translation of the test into sign language or use of a magnification device) who takes an initial ELP assessment in a state that offers this type of accommodation. If the student moves to another state that uses a different ELP assessment that does not offer this accommodation, the classification of the student (e.g., EL or not) could change. Variability in the types of accommodations made available to test takers has the potential to affect comparability and reliability of the initial ELP assessment as a whole. Although several research reports have been issued that list a range of types of accommodations that are offered with different ELP assessments (e.g., Christensen, Albus, Liu, Thurlow, & Kincaid, 2013), it is not clear whether these accommodations apply to both initial and summative ELP assessments.

Another challenge is that students with disabilities may have difficulty participating in the assessment of certain ELP domains (listening, reading, speaking, and writing) depending on their disability (Christensen *et al.*, 2013; Guzman-Orth *et al.*, 2016). For example, a student who is deaf or hard of hearing may not be able to meaningfully participate in test items designed to measure listening skill in English. As a result, some students with disabilities may participate in assessment of some ELP domains using an Individualized Education Program–approved accommodation, while others might be excluded entirely from participating in assessment of certain ELP domains. This raises a question with respect to comparability and reliability, because some initial ELP assessments (or portions thereof) might not provide adequate accommodations to certain potential EL students with disabilities.

Recommendation 4: Test Developers and States Should Provide Clear Guidelines to Select and Provide Accommodations on Initial ELP Assessments to Ensure Their Accessibility for Students with Disabilities

To ensure that all students, including students with disabilities, have adequate opportunity to access initial ELP assessments, we recommend implementing evidence-centered design to clarify assessment design decisions. We also recommend that universal design considerations and general test development guidelines can be implemented in designing and producing initial ELP assessments (Thompson, Johnstone, & Thurlow, 2002). It is also advisable that accessibility features and accommodations for students with disabilities be considered from the beginning in the assessment development process to ensure that they are appropriate for the targeted population of students (Guzman-Orth *et al.*, 2016). We suggest that states and test developers list all the accommodations that are offered on the initial ELP assessment and clearly state for which domain (e.g., listening, speaking, reading, and writing) the accommodations are appropriate (Albus & Thurlow, 2008).

Equally important is the inclusion of ELs with disabilities in prototyping and pilot studies during the assessment design stage. These studies should help states and test developers examine (a) the extent to which ELs with disabilities can complete the items or tasks on the initial ELP assessment independently or with appropriate accommodations and scaffolding, (b) the types of difficulties they have completing the items or tasks, and (c) whether these items or tasks elicit the intended target-language uses.

If accommodations are offered, it is also important to examine how ELs with disabilities interact with different types of accommodations, if the accommodations work differently across domains, and what the consequences (validity, effectiveness) of using the accommodations are. It is also important to examine if the accommodations provide an unfair advantage or simply level the playing field for EL students with disabilities. Such evidence can assist in determining the type of accommodations and accessibility features students with disabilities need to meaningfully participate in initial

ELP assessments, and whether justification exists to develop a separate form or forms of the initial ELP assessment for students with disabilities.

Conclusion

In order to ensure that ELs have an equal opportunity to achieve the same academic standards as other students, state and local policymakers and educators must provide these students with appropriate instructional services. This provision can only happen if language-minority students—whether EL or IFEP—are properly identified and classified. Incorrectly classified students are less likely to be provided with instructional services that match their strengths and needs (Bailey & Carroll, 2015). They may be excluded from services that would better help them meet high academic standards, and they may not receive optimal supports to advance both linguistically and academically. The resulting suboptimal learning opportunities could establish a trajectory that places these students at greater risk of losing motivation, disengaging, and falling behind.

In this paper, we have highlighted several issues with current EL initial identification and classification policies, processes, and tools. We have also offered several recommendations to address these issues and improve the initial identification and classification of ELs. Several new K–12 ELP assessment systems (e.g., WIDA, ELPA21, and ELPAC)⁶ have recently been released or are currently being developed, and the fact that each of them includes (or will include) an initial ELP assessment that separate from the summative ELP assessment, is a positive sign. We are hopeful that, by considering these recommendations, the emerging generation of initial ELP assessments will better serve the needs of ELs and their educators.

Acknowledgments

The authors would like to thank the reviewers for their thoughtful comments on the earlier versions of this paper: James E. Carlson (ETS), Vincent Dean (ETS), Phil Everson (ETS), Kenji Hakuta (Stanford University), Maurice Hauck (ETS), Danielle Guzman-Orth (ETS), Scott Paris (ETS), Rose Payan (ETS), Don Powers, (ETS), Catherine E. Snow (Harvard University), and Mikyung Wolf (ETS).

Notes

- 1 If an initial ELP assessment is intended to provide information to inform both identification/classification decisions and placement of ELs into appropriate levels of language instruction, the assessment should be designed as such (e.g., by including test items that provide measurement at a range of EL levels rather than concentrated only on the distinction between EL and IFEP status).
- 2 See Abedi (2008), Bailey and Kelly (2013), and Linquanti and Bailey (2014) for further discussion of these issues.
- 3 See Bailey and Heritage (2008) and Wolf, Everson, et al. (2014) for a discussion on these language domains.
- 4 In this paper, we use the term *young learners* to refer those who are in early elementary grades (Grades K–2, between the ages of 4 and 8).
- 5 It is important to note that students who may have an undiagnosed disability need to be identified both as ELs and as students with disabilities using different instruments to ensure that they also have access to appropriate services.
- 6 For more information on WIDA, see <http://www.wida.us/> For more information on ELPA21, see <http://www.elpa21.org/> For more information on ELPAC, see <http://www.cde.ca.gov/ta/tg/ep/>

References

- Abedi, J. (2008). Classification system for English language learners: Issues and recommendations. *Educational Measurement Issues and Practice*, 27(3), 17–22.
- Albus, D., & Thurlow, M. L. (2008). Accommodating students with disabilities on state English language proficiency assessments. *Assessment for Effective Intervention*, 33(3), 156–166.
- Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.
- Bailey, A. L. (2008). Assessing the language of young learners. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 7. Language testing and assessment* (2nd ed., pp. 379–398). Berlin, Germany: Springer Science+Business.

- Bailey, A. L., & Carroll, P. E. (2015). Assessment of English language learners in the era of new academic content standards. *Review of Research in Education*, 39, 253–294.
- Bailey, A. L., & Heritage, M. (2008). *Formative assessment for literacy: Building reading and academic language skills across the curriculum*. Thousand Oaks, CA: Corwin.
- Bailey, A. L., & Kelly, K. R. (2013). Home language survey practices in the initial identification of English learners in the United States. *Educational Policy*, 27, 770–804.
- Batalova, J., Fix, M., & Murray, J. (2005). *English language learner adolescents: Demographics and literacy achievements. Report to the Center for Applied Linguistics*. Washington, DC: Migration Policy Institute.
- California Department of Education. (2012). *California English language development standards: Kindergarten through grade 12*. Retrieved from <http://www.cde.ca.gov/sp/el/er/documents/eldstndpublication14.pdf>
- California Department of Education. (2015). *DataQuest*. Retrieved from <http://data1.cde.ca.gov/dataquest/>
- Christensen, L. L., Albus, D. A., Liu, K. K., Thurlow, M. L., & Kincaid, A. (2013). *Accommodations for students with disabilities on state English language proficiency assessments: A review of 2011 state policies*. Minneapolis: University of Minnesota, Improving the Validity of Assessment Results for English Language Learners with Disabilities (IVARED).
- 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.
- Council of Chief State School Officers. (2012). *Framework for English Language Proficiency Development Standards corresponding to the Common Core State Standards and the Next Generation Science Standards*. Washington, DC: Author.
- Epstein, A. S., Schweinhart, L. J., DeBruin-Parecki, A., & Robin, K. B. (2004). *Preschool assessment: A guide to developing a balanced approach*. Retrieved from <http://nieer.org/resources/policybriefs/7.pdf>
- Espinosa, L. M., & García, E. (2012). *Developmental assessment of young dual language learners with a focus on kindergarten entry assessment: Implications for state policies* (Working paper #1). Chapel Hill: University of North Carolina, Frank Porter Graham Child Development Institute.
- Faulkner-Bond, M., Shin, M., Wang, X., Zenisky, A. L., & Moyer, E. (2013, April). *Score reports for English proficiency assessments: Current practices and future directions*. Paper presented at the National Council on Measurement in Education, San Francisco, CA.
- Fortuny, K., Hernandez, D. J., & Chaudry, A. (2010). *Young children of immigrants: The leading edge of America's future* (Brief No. 3). Retrieved from <http://www.urban.org/UploadedPDF/412203-young-children.pdf>
- Guzman-Orth, D., Laitusis, C., Thurlow, M., & Christensen, L. (2016). *Conceptualizing accessibility for English language proficiency assessments* (Research Report No. RR-16-07). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/ets2.12093>
- Hakuta, K. (2011). Educating language minority students and affirming their equal rights: Research and practical perspectives. *Educational Researcher*, 40, 163–174.
- Hauck, M. C., Wolf, M. K., & Mislevy, R. (2016). *Creating a next-generation system of K-12 English learner (EL) language proficiency assessments* (Research Report No. RR-16-06). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/ets2.12092>
- Linquanti, R., & Bailey, A. L. (2014). *Reprising the home language survey: Summary of a national working session on policies, practices, and tools for identifying potential English learners*. Washington, DC: Council of Chief State School Officers.
- Linquanti, R., & Cook, H. G. (2013). *Toward a "common definition of English learner": Guidance for states and state assessment consortia in defining and addressing policy and technical issues and options*. Washington, DC: Council of Chief State School Officers.
- McKay, P. (2006). *Assessing young language learners*. Cambridge, England: Cambridge University Press.
- National Association for the Education of Young Children. (2005). *Screening and assessment of young English-language learners*. Washington, DC: Author.
- National Research Council. (2011). *Allocating federal funds for state programs for English language learners*. Washington, DC: National Academic Press.
- No Child Left Behind Act, 20 U. S. C. §§ 3302(d) (2001).
- Padilla, A. M. (2001). Issues in culturally appropriate assessment. In L. A. Suzuki, J. P. Ponterotto, & P. J. Meller (Eds.), *Handbook of multicultural assessment: Clinical, psychological, and educational applications* (2nd ed., pp. 5–27). San Francisco, CA: Jossey-Bass.
- Pandya, C., Batalova, J., & McHugh, M. (2011). *Limited English proficient individuals in the United States: Number, share, growth, and linguistic diversity*. Washington, DC: Migration Policy Institute.
- Pooler, E., & Lopez, A. A. (2013, November). *Exploring innovative ways to better measure language proficiency of English language learners (ELLs)*. Paper presented at the educational summit of the Association of Latino Administrators and Superintendents, Denver, CO.
- Ruiz Soto, A. G., Hooker, S., & Batalova, J. (2015). *Top languages spoken by English language learners nationally and by state*. Washington, DC: Migration Policy Institute.
- Santos, R. M. (2004). Ensuring culturally and linguistically appropriate assessment of young children. *Young Children*, 59, 48–50.
- Schleppegrell, M. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Erlbaum.

- Thompson, S. J., Johnstone, C. J., & Thurlow, M. L. (2002). *Universal design applied to large scale assessments* (NCEO Synthesis Report No. 44). Minneapolis: University of Minnesota, National Center on Educational Outcomes.
- U.S. Department of Justice & U.S. Department of Education. (2015). *Dear colleague letter: English learner students and limited English proficient parents*. Retrieved from: <https://www2.ed.gov/about/offices/list/ocr/letters/colleague-el-201501.pdf>
- United States Department of Education. (2015). *English learner tool kit*. Retrieved from: <http://www2.ed.gov/about/offices/list/oela/eltoolkitchap1.pdf>
- Wolf, M. K., Everson, P., Lopez, A., Hauck, M., Pooler, E., & Wang, J. (2014). *Building a framework for a next-generation English language proficiency assessment system* (Research Report No. RR-14-34). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/ets2.12034>
- Wolf, M. K., Guzman-Orth, D., & Hauck, M. C. (2016). *Next-generation summative English language proficiency assessments for English learners: Priorities for policy and research* (Research Report No. RR-16-08). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/ets2.12091>
- Wolf, M. K., Guzman-Orth, D., Lopez, A., Castellano, K., Himelfarb I., & Tsutagawa, F. (in press). Integrating scaffolding strategies into technology-enhanced assessments of English learners: Task types and measurement models. *Educational Assessment*.
- Wolf, M. K., Kao, J. C., Griffin, N., Herman, J. L., Bachman, P. L., Chang, S. M., & Farnsworth, T. (2008). *Issues in assessing English language learners: English language proficiency measures and accommodation uses—Practice review* (CSE Tech. Rep. No. 732). Los Angeles: University of California, National Center for Research on Evaluation Standards and Student Testing (CRESST).
- Wolf, M. K., & Lopez, A. A. (2014a). *Comparing the performance of K-12 English language learners and native speakers of English on assessment tasks*. Paper presented at the meeting of the Language Testing Research Colloquium, Amsterdam, The Netherlands.
- Wolf, M. K., & Lopez, A. A. (2014b). *The use of scaffolding strategies in the assessment of English learners*. Invited paper presented at the Teachers College Roundtable in Second Language Studies, New York, NY.
- Wong Fillmore, L., & Snow, C. E. (2002). What teachers need to know about language. In C. T. Adger, C. E. Snow, & D. Christian (Eds.), *What teachers need to know about language* (pp. 7–53). McHenry, IL: Delta Systems & Center for Applied Linguistics.

Suggested citation:

Lopez, A. A., Pooler, E., & Linquanti, R. (2016). *Key issues and opportunities in the initial identification and classification of English learners* (Research Report No. RR-16-09). Princeton, NJ: Educational Testing Service. <http://dx.doi.org/10.1002/ets2.12090>

Action Editor: Donald Powers

Reviewers: Danielle Guzman-Orth and Philip Everson

ETS and the ETS logo are registered trademarks of Educational Testing Service (ETS). MEASURING THE POWER OF LEARNING is a trademark of ETS. All other trademarks are property of their respective owners.

Find other ETS-published reports by searching the ETS ReSEARCHER database at <http://search.ets.org/researcher/>