THE NUMBER OF YOUNG CHILDREN from homes in which a language other than English is spoken has increased dramatically in the last decade. Currently, more than four million dual language learners (DLLs) are enrolled in early care and education programs (Aud et al., 2012). Nationally, roughly 30 percent of preschoolers in Head Start are from households in which a language other than English is spoken (Aikens et al., 2011), and almost one-third of families enrolled in Early Head Start report speaking a language other than or in addition to English (Vogel et al., 2011). The growing number of young DLLs served increases the need to understand the development of this population, how to monitor that development, and how to interpret research and findings about DLLs’ development.

As research about DLLs and their development grows, it will be important to understand available assessment approaches, the information each provides about children’s development, how approaches are selected, and the implications of those selections. The approach used for assessment can lead to different inferences about children’s development. For example, assessment of DLLs’ skills solely in English may underestimate their true knowledge or ability in areas other than English language development. For young DLLs, information learned in the home language may not be readily accessible in English (and vice versa). In addition, when researchers include or exclude children from study samples and findings based on the assessment approach used, different conclusions may be drawn about the characteristics and skills of the population under study.

This brief examines approaches commonly used in the research literature to assess the language and literacy skills of young DLL children. We describe the different approaches and the information each provides about the children’s skills and development. We also describe key factors that shape the use and selection of assessment approaches, as well as factors that should be considered when interpreting findings in the literature. We draw on a report that examined assessment procedures in seven large-scale studies of early childhood that included DLL children and 73 peer-reviewed research studies of the developmental trajectories of young DLL children.1

The brief addresses three broad questions:

1. What assessment approaches are used in research studies to determine what language to use in assessing DLLs, and what information does each provide?
2. How do factors such as study goals and child characteristics guide the selection and use of assessment approaches with DLLs in these studies?
3. What are the implications of assessment approaches for interpreting data and study findings on DLLs?

What assessment approaches are used in research studies to determine language of assessment for DLLs, and what information does each approach provide?

Researchers use a range of approaches to determine the language (or languages) in which to assess young DLLs’ knowledge. Common approaches include (1) assessment in English or the home language only, based on either performance on a language proficiency measure or parent and/or teacher report of the child’s primary or dominant language; (2) dual-language administration, in which assessments are conducted both in English and the child’s home language; and (3) conceptually scored bilingual assessments that give children credit for responses provided in either English or the home language.
Each approach has a specific purpose and provides unique information about the skills and development of DLLs (Table 1). For example, language-specific assessments provide information about children’s skills in the administered language only. Concepts or vocabulary words that the child does not know in that language will not be represented in the assessment results. In contrast, use of a dual-language approach recognizes the possibility that children know certain concepts or vocabulary words in English that they do not know in their home language, and vice versa. Children respond to the assessment in one language and then the other, providing information on their skills and knowledge in both languages. A conceptually scored approach provides information on children’s knowledge across languages via a single administration.

Table 1. Assessment Approaches with DLLs in Reviewed Studies, Information Provided by Each, and Sample Measures

<table>
<thead>
<tr>
<th>Assessment Approach</th>
<th>Information Provided:</th>
<th>Sample Language/Literacy Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-specific assessment, with the language determined by a language screener and/or teacher/parent report of home language</td>
<td>Children’s knowledge in the home language or in English only</td>
<td>Language Screener: English and Spanish Pre-Language Assessment Scales 2000 (PreLAS; Duncan &amp; DeAvila, 2002) MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 1993) or MacArthur-Bates Inventario del Desarrollo de Habilidades Comunicativas (Inventarios; Jackson-Maldonado, Bates, &amp; Thal, 2003)</td>
</tr>
<tr>
<td>Dual-language assessment</td>
<td>Children’s knowledge in English and in the home language, separately</td>
<td>Woodcock-Johnson Psycho-Educational Battery-Third Edition (WJ III; Woodcock et al., 2001) and Bateria III Woodcock-Muñoz (Woodcock et al., 2004)</td>
</tr>
<tr>
<td>Conceptually scored bilingual assessment</td>
<td>Children’s knowledge across languages, via a single administration</td>
<td>Expressive One-Word Picture Vocabulary Test–Spanish Bilingual Edition (EOWPVT-SBE; Brownell, 2001)</td>
</tr>
</tbody>
</table>

Figure 1 highlights the prevalence of each of these approaches across the reviewed studies. Notably, many studies used more than one approach in assessing DLLs.

Figure 1. Prevalence of Approaches in the Assessment of DLLs in the Reviewed Studies

Note: This figure depicts the prevalence of approaches for determining the language of administration. Percentages do not include studies that used parent or teacher report of language proficiency for the sole purpose of describing the characteristics of the study sample. Prevalence rates within peer-reviewed and national studies do not necessarily sum to 100 percent. Some studies used a combination of approaches to determine the language of administration.
**How do study and assessment goals guide selection and use of language and literacy assessment approaches with young DLLs?**

As noted, multiple approaches are taken for the assessment of language skills in young DLLs. Some of the studies we reviewed focused on describing the relationship between children’s proficiency in their home language and English, thus necessitating a dual-language approach to assessment. Others sought to understand children’s development more broadly, with less regard for a specific language. Still others placed less emphasis on linguistic diversity by assessing the skills of all children in either English or in their home language, usually Spanish. Most published assessments are available only in English and Spanish.

Ultimately, the goal of the research drove the assessment approach. For example, if the goal was to document DLLs’ growth in English language proficiency, use of an English language assessment at multiple points in time was needed. However, if the goal was to document DLL children’s abilities in a particular developmental domain, assessment in both languages would yield a more complete (and accurate) picture of their overall competency.

The latter goal could be achieved through either a dual-language assessment or a conceptually scored assessment. However, the use of a dual-language approach compares DLLs to children who are monolingual in each language and often results in a lower score on both measures when compared to scores on a conceptually-scored bilingual assessment. When using a dual language assessment, researchers sometimes also examine the number of unique words that a DLL knows across the languages. Table 2 lists the assessment approaches used in the reviewed studies and examples of study goals related to the use of the approaches.

**Table 2. Assessment Approaches and Associated Goals in the Reviewed Studies**

<table>
<thead>
<tr>
<th>Assessment Approach</th>
<th>Sample Study Goals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language-specific assessment</td>
<td>Track children’s knowledge and skills in English over time</td>
</tr>
<tr>
<td></td>
<td>Track children’s knowledge and skills in home language over time</td>
</tr>
<tr>
<td></td>
<td>Measure literacy (and other school readiness areas) in the language in which the child demonstrated greater proficiency (often for baseline status) and/or in the language of instruction (to assess change over time)</td>
</tr>
<tr>
<td>Dual-language assessment</td>
<td>Understand the contribution of proficiency and skills in the home language to the development of skills in English</td>
</tr>
<tr>
<td></td>
<td>Examine differences in skills in the home language and in English</td>
</tr>
<tr>
<td></td>
<td>Compare DLLs’ skills in the home language and English with those of their monolingual peers</td>
</tr>
<tr>
<td></td>
<td>Provide information about children’s growth in English and Spanish (or other home language) vocabulary</td>
</tr>
<tr>
<td>Conceptually scored bilingual assessment</td>
<td>Describe DLLs’ overall concept development</td>
</tr>
<tr>
<td></td>
<td>Explore the relationship between vocabulary and literacy skills among DLLs</td>
</tr>
</tbody>
</table>

Across the reviewed studies, dual-language administration was the approach most commonly used. In nearly 80 percent of the peer-reviewed studies and over one-half of the large-scale government reports, DLLs were assessed in both English and the home language in at least one area of language or literacy development, irrespective of language proficiency or dominance. This approach reflects the focus of these studies. For example, some studies sought to understand the contribution of abilities in the first language [L1] to the development of skills in the second language [L2; usually English] and differences in skills in L1 and L2. Some studies also examined between-group differences between DLLs’ skills and the skills of their monolingual peers in that language. Notably, when conducting research with samples of monolingual and DLL children, information on factors including socioeconomic background, age ranges represented, and, when available, differences in ethnic and cultural background are critical to contextualizing differences in performance. In the absence of this information, caution should be exercised when making interpretations, since such characteristics may account for observed differences.
Only five of the reviewed studies used at least one conceptually scored bilingual assessment, all of which were Spanish-English bilingual assessments. Conceptually-scored bilingual assessments such as the Bilingual English Spanish Assessment (BESA; Peña, Gutierrez-Cellen, Iglesias, Goldstein, & Bedore, in preparation), the Emergent Literacy Profile (ELP; Dickinson & Chaney, 2004), and the Expressive One-Word Picture Vocabulary Test–Spanish Bilingual Edition (EOWPVT–SBE; Brownell, 2001) were used in studies with various research purposes. These included describing the overall language development of DLLs, examining the influence of children’s home language on the acquisition of English, and exploring the relationship between DLLs’ vocabulary and literacy skills.

In large-scale, national studies in which DLLs were not the primary focus, children were generally assessed in a single language. Some—for example, the Head Start Impact Study and the 1997 and 2000 cohorts of the Head Start Family and Child Experiences Survey (FACES)—used teacher/caregiver report to determine the most appropriate language of assessment. Most often, however, language of assessment was determined by an English language proficiency screener in combination with parent report of home language (Figure 2). Across studies, researchers most often used the Simon Says and Art Show sub-tests from the PreLAS. Children who did not pass the English language screener were either not assessed or were administered the assessment in Spanish.

Although the same assessment tools were often used across these studies, the threshold for determining language of assessment varied. For example, the preschool and kindergarten rounds of the Early Childhood Longitudinal Study–Birth Cohort (ECLS-B) set a low (lenient) criterion that required children to respond correctly to only one item in English (beyond the practice item) to receive the assessment in English. The goal of this approach was twofold: to include as many children as possible in the English assessment and to track progress in their knowledge and skills in English longitudinally. Meanwhile, studies such as (FACES 2006 and 2009) used a higher (more conservative) criterion for English assessment (fewer than five consecutive errors on Simon Says and Art Show) to route children into an English assessment, with more than five consecutive errors on the English assessment routing Spanish-speaking DLLs into a Spanish assessment. As shown in Figure 1, none of the peer-reviewed studies included in the review used English proficiency screening procedures.

**Figure 2. Methods Used to Describe DLLs’ Language Exposure and Proficiency**

![Figure 2: Methods Used to Describe DLLs’ Language Exposure and Proficiency](image)

Note: Other combination approaches include parent report and use of a language proficiency screener such as the PreLAS; combinations of parent, teacher, and assessor reports; and combinations of parent and teacher reports and results from a narrative storytelling task.
FACES is an example of a national study in which DLLs are not the primary focus, but the advantages and disadvantages of existing assessments and approaches were considered with each cohort and the design changed across time and cohorts according to the recommendations of DLL and ECE experts, Head Start encourages support of home language, and 80 percent of DLLs who entered the program in the fall of 2009 were from homes where Spanish was spoken (Aikens et al., 2011). In the more recent FACES 2006 and 2009 cohorts, multiple approaches to assessment were used to describe the development of the growing number of DLLs in Head Start. These included dual-language assessment of children’s receptive vocabulary in English and Spanish, using the Picture Vocabulary Test–Fourth Edition (PPVT-4; Dunn & Dunn, 2006) and Test de Vocabulario en Imágenes Peabody (TVIP; Dunn, Padilla, Lugo, & Dunn, 1986); conceptual scoring of expressive vocabulary (in FACES 2009 only, using the EOWPVT-SBE); and assessment of literacy (and mathematics) tasks in English or in Spanish only, based on the child’s English language proficiency. This combination of approaches was used to describe more fully the language development of DLLs, providing information about their vocabulary development in both Spanish and English and about their literacy (and mathematics knowledge) in the language in which they demonstrated greater proficiency.

**How do child characteristics such as home language and age guide selection and use of assessment approaches with DLLs?**

The selection of approaches in the reviewed studies was also driven by the availability and appropriateness of measures for the population studied. For example, due to the shortage of assessments with evidence of validity for U.S. samples, DLLs in FACES with home languages other than English or Spanish who did not pass the English screener did not receive any additional direct assessments of language and literacy, apart from the English-administered PPVT-4 and EOWPVT. Relatively few fully developed and psychometrically tested assessments are available in Spanish and even fewer in other languages (Espinosa & López, 2007; Peña & Halle, 2011). Except for the parent- or teacher-report measures (for example, ASQ-3 and the CDI/Inventario), the early assessments used in other languages (Table 3) were translated from English or developed by the study researchers (for example, the PPVT-III was translated into Urdu and the PPVT-R into Mandarin).

The few reviewed studies that included samples of children younger than two years of age used different assessment approaches than those examining older children. For example, although the nine-month and two-year data collections of the ECLS-B included English and Spanish versions of the direct assessments, the procedure for determining the language in which to assess children relied on parent report rather than the English PreLAS used in the preschool and kindergarten rounds. In the descriptive study of Early Head Start (Baby FACES), assessments for the one-year-old cohort included parent- and teacher-report measures (Ages and Stages Questionnaires–Third Edition [ASQ3], MacArthur-Bates Communicative Development Inventories [CDI/Inventario]) and videotaped parent–child interactions. Other research studies of very young children collected language samples (including audio and video recordings of natural conversations) or used parent reports of vocabulary. In most cases, the parents completed word lists for the child’s predominant language (as selected by the parent) or completed both the English and home language forms of the CDI/Inventario. While few measures of early development are available in languages other than English for preschool-aged children, the availability of such assessments for infants and toddlers is even more limited. An adaptation of the Bayley Scales of Infant Development, administered in Spanish in ECLS-B at ages 9 and 24 months, was the only assessment translated and used with DLLs under the age of two years (personal communication, Jerry West, July 16, 2012).
Table 3. Assessments in Languages Other Than English and Spanish

<table>
<thead>
<tr>
<th>Study Citation</th>
<th>Name of Assessment</th>
<th>Languages Other Than English and Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anderson, 2004</td>
<td>Researcher-developed word lists</td>
<td>Korean, Russian, French</td>
</tr>
<tr>
<td>Buckwalter &amp; Gloria Lo, 2002</td>
<td>Emergent literacy task</td>
<td>Chinese</td>
</tr>
<tr>
<td>Davidson, Raschke, &amp; Perez, 2010</td>
<td>PPVT-III translation</td>
<td>Urdu</td>
</tr>
<tr>
<td>Gildersleeve-Neumann, &amp; Wright, 2010.</td>
<td>Researcher-developed: Phonological and Articulatory Bilingual Assessment</td>
<td>Russian</td>
</tr>
<tr>
<td>Kan &amp; Kohnert, 2005</td>
<td>Researcher-developed picture naming (derived from the CDI) and picture identification tasks</td>
<td>Hmong</td>
</tr>
<tr>
<td>Kan &amp; Kohnert, 2008</td>
<td>Researcher-developed picture naming (derived from the CDI) , picture identification, and fast mapping tasks</td>
<td>Hmong</td>
</tr>
<tr>
<td>Kim, 2009</td>
<td>Researcher-developed phonological test, and Ready-to-Read Word Test (sight vocabulary)</td>
<td>Korean</td>
</tr>
<tr>
<td>Marinova-Todd, Zhao, &amp; Bernhardt, 2010</td>
<td>PPVT-R; phonological tasks</td>
<td>Mandarin</td>
</tr>
<tr>
<td>Sheng, , McGregor, &amp; Marian, 2006</td>
<td>Repeated Word Association Test</td>
<td>Mandarin</td>
</tr>
</tbody>
</table>

What are the implications of assessment approaches for interpreting data and study findings on DLLs?

Available evidence suggests that different methods for assigning children to a single language of assessment will affect conclusions drawn about their development. Methods other than direct measurement of English language skills can place children into a language of assessment with different levels of accuracy. Researchers found stronger relations of direct assessments of child outcomes with parent report than with teacher report (Vagh, Pan, & Mancilla-Martinez, 2009). For young DLLs, parents appeared to be better reporters of vocabulary as measured by the CDI than teachers or other caregivers outside of the home, but it may depend on the amount of time the child spends in each setting.

Longitudinal measurement can be particularly challenging for DLLs. FACES allows change in the language of assessment across data collection rounds (from a non-English to an English assessment) based on children’s performance on a language screener. Reports from the study typically discuss results on literacy assessments based on children who remain in the same language of assessment across the fall and spring. Thus, children who change from a Spanish to an English assessment are not represented in the mean change over time. With most instruction in the United States occurring in English, the results on a Spanish measure of literacy may underestimate the advances Spanish-speaking children are making in this area. In addition, a shift in the language of administration across the fall-to-spring assessments makes it challenging to describe growth meaningfully across the program year for this subgroup of children.
When reviewing published studies or findings on DLLs, it is important to consider the method or approach used for assessment and what it may mean for inclusion or exclusion from the study sample, and for interpretation of the results for DLLs. The method and goals of the assessment, as well as appropriateness of selected measures, should guide the interpretation of findings. For example, if the question of interest is how children progress in programs in which instruction occurs only in English, an assessment of Spanish literacy may underestimate the knowledge of a young DLL. Though early literacy skills such as print concepts apply to both English and Spanish, DLLs would need to generalize these skills across languages in order to do well on literacy assessments in the opposite language, and they may have had no exposure at all to the names of letters in Spanish. On the other hand, a study by Anthony and colleagues (2009) indicated a stronger correlation between vocabulary and phonological awareness in the same language than across languages, so expected performance of DLLs on an English assessment of phonological awareness would differ from that of children who speak only English. It is easier to hear the individual sounds of a word when the word is familiar (Metsala, 2009). Researchers need to consider how the language of assessment influences the meaning of the results.

Data from large-scale national studies focused on young children (e.g., ECLS-B, FACES, Head Start Impact Study) are made publicly available to interested researchers for secondary data analysis. When using these data sets to examine the development of DLLs, researchers need to consider how the language of assessment was determined and what that means for the research questions of interest and for the sample of children included in the analysis. Did the approach exclude children, and, if so, whom will the findings represent? Were children assessed in only one language or in both separately? How do the differences in approach affect what can be said about children’s development? Are selected measures appropriate for use with low-income and culturally or linguistically diverse children?

Summary

Multiple approaches to the assessment of DLLs are used in the research literature, each providing different perspectives on children’s development. Approaches are guided by the goals of the assessment and of the study, as well as the availability and appropriateness of measures for the population under study. Each of these factors should be considered when interpreting findings in the literature and analyzing data that include DLLs.

(Endnotes)

1 The full report (Bandel et al., 2012) includes information from 73 peer-reviewed journal research articles published between 2000 and 2011 that studied U.S. or Canadian samples with at least one direct child assessment or standardized rating of the development of DLL children prior to kindergarten entry and English as one of the two languages being learned. In addition, we reviewed seven government reports of large-scale studies of early childhood published between 2000 and 2011 that included DLL children and at least one direct assessment of children’s development prior to kindergarten entry.

2 One study administered the EOWPVT-SBE in English and Spanish separately, rather than using the conceptual scoring.
References


---

**About CECER-DLL**

CECER-DLL is a national center that is building capacity for research with dual language learners (DLLs) ages birth through five years. CECER-DLL aims to improve the state of knowledge and measurement in early childhood research on DLLs, identify and advance research on best practices for early care and education programming, and develop and disseminate products to improve research on DLLs. CECER-DLL is a cooperative agreement between the Frank Porter Graham (FPG) Child Development Institute at The University of North Carolina at Chapel Hill and the Office of Planning, Research, & Evaluation (OPRE) in the Administration for Children & Families (ACF), in collaboration with the Office of Head Start and the Office of Child Care.

**Suggested citation**


This brief summarizes results from a review of the literature sponsored by CECER-DLL conducted by a research team consisting of Eileen Bandel, Sally Atkins-Burnett, Dina C. Castro, Claire Smither Wulsin and Marisa Putnam, with Margaret Burchinal, Lisa López, Vera Gutiérrez-Clellen, and Ellen Peisner-Feinberg as research partners. The work was supported by a cooperative agreement funded by the Office of Planning, Research, and Evaluation (OPRE), U.S. Department of Health and Human Services. Permission to copy, disseminate, or otherwise use information from this document for educational purposes is granted, provided that appropriate credit is given.

**Additional Resources:** For additional information regarding this research brief, see [http://cecerdll.fpg.unc.edu](http://cecerdll.fpg.unc.edu)