



Challenges in Assessing California's Diverse Dual Language Learners

AUTHORS: Jodi Jacobson Chernoff, Shannon Keuter, Yuuko Uchikoshi, Heather Quick, and Karen Manship

Introduction

Many people living in the United States speak a language other than English. For years, however, assessments of U.S. children that have provided information to educators to guide instruction and inform program design implementation and accountability have been conducted only in English (Abedi, 2002; Garcia, 1991; García & Torres-Guevara, 2010; Snow & Van Hemel, 2008). This practice has been common even in early childhood education settings (García & Torres-Guevara, 2010). Consequently, the majority of “off-the-shelf” assessments of young children’s academic and social-emotional skills in the United States are available only in English. These assessments have been developed primarily and normed using monolingual English-speaking children in the United States (Hammer et al., 2014). Although efforts have been made to include a diverse sample of children in the norming groups for many instruments, the most established and commonly used child assessments are developed for use with English speakers.

However, given the nation’s language diversity (National Academies of Sciences, Engineering, and Medicine, 2017), it is likely that the full range of knowledge of young, linguistically diverse children growing up in the United States have not been assessed accurately because many do not speak English well enough to demonstrate their abilities on those English assessments. Today nearly one in three young children in the United States live with at least one parent who speaks a language *other than English* at home (Park, O’Toole, & Katsiaficas, 2017), and California is home to more dual language learners (DLLs) than any other state (Federal Interagency Forum on Child and Family Statistics, 2015). Given this growing population, assessing children only in English may not reveal their full set of knowledge and skills (Peets & Bialystok, 2015). Educators, administrators, parents, and researchers need to understand whether low scores on an assessment reflect limited skills in that domain or rather limited proficiency in English. Low assessment scores for these DLLs may also be misinterpreted as learning disabilities. In fact, research shows that English learners are under-represented in special education classrooms in kindergarten and appear to be identified later possibly due to the difficulties in distinguishing between a lack of English proficiency and a learning disability (Samson & Lesaux, 2009). When possible, children should be assessed in both English and their home language so that the sum total of the child’s strengths and areas for development can be

measured (Sugarman & Villegas, 2020; Vitiello, Downer, & Williford, 2011). However, it is often difficult to assess DLL children in their home language because of a lack of reliable and valid assessments normed for children from diverse backgrounds (Snow & Van Hemel, 2008).

To address this need, some assessment developers have begun to develop and validate assessments in non-English languages, primarily in Spanish. These changes have been made in an attempt to ensure that Spanish-speaking children in the United States can be more accurately assessed, and their scores can be examined along with their peers who were assessed in English. However, there are some caveats to consider. For one, many of the Spanish-language assessments are simple translations of the originally developed English assessment (Guzman-Orth, Lopez, & Tolentino, 2017). This could be problematic because sounds and concepts that are frequently used and learned in English may not be as common or of the same difficulty level in Spanish (Guzman-Orth et al., 2017).

Many Spanish-language assessments available for use in the United States were developed, validated, and norm-referenced with monolingual Spanish speakers living *outside the United States*, for example in Mexico or Spain (Guzman-Orth et al., 2017; Hammer et al., 2014). Language variations evolve differently by region, so assessments for Spanish-speaking children living in the United States should be normed and validated on Spanish speakers who also live in the United States.

It has also become apparent that assessments for students who speak languages other than English and Spanish are needed. There are very few assessments available in languages other than English and Spanish. These assessments need to be more than simple translations of their English counterparts; they need to be normed and validated on speakers of those languages who live in the United States. Moreover, all language versions of an assessment need to be placed on the same metric and equated so that the scores they yield can be combined and compared. However, the development of these badly needed assessments will take time and resources to develop (Hammer et al., 2014).

This paper begins by discussing some of the issues that arise when conducting research with young DLLs and the challenges that result from the lack of high-quality, validated assessments in multiple languages. The First 5 California DLL Pilot Study (www.californiadllstudy.org) serves as a case study of how these issues and challenges might be addressed. Next, this paper will explain how researchers on the First 5 DLL study have responded to the lack of high-quality assessments in multiple languages by instead using assessments that have either been translated or adapted, again presenting the First 5 California DLL Pilot study as an example. However, these approaches (adaptation and translation) are not perfect and so this paper will go on to help current researchers who use these solutions avoid important pitfalls. Finally, we conclude with a call to action and discussion of possible next steps.

Choosing and Administering Assessments With DLLs

In California, school children speak more than 65 non-English languages (California Department of Education Data Reporting Office, 2019-2020). DLLs make up nearly 60% of the 0- to 5-year-old population; although the large majority speak Spanish, 16.8% of them speak an Asian language such as Chinese or Vietnamese

(Holtby, Lordi, Park, & Ponce, 2017). The remaining DLLs in California speak many other languages, such as Arabic and Russian. Research on young DLLs in California has been working to include more than just English and Spanish. One example of such research is the First 5 DLL Study.

The First 5 DLL study aims to understand the relationship of particular instructional practices in diverse early learning settings in California to young DLLs' development. The study had an explicit goal to include and assess children from homes where languages other than Spanish and English are spoken. Consequently, this study includes children whose families speak Spanish, Mandarin, Cantonese, and Vietnamese.¹

When researching with young DLLs, there are several factors to consider. One must consider which assessments to use based on the constructs to be measured and the age of the children to be assessed. Next, the researcher must consider the language of the assessment. This is often determined by who will complete the survey or which skills need to be assessed in multiple languages. For skills that can be or need to be assessed in only one language, it is recommended to assess in the child's dominant language. Steps need to be taken to determine the child's dominant language. For example, the First 5 DLL study chose to compare scores on the Woodcock-Johnson IV Oral Language battery. In any case, assessments and surveys are needed in the study children's home language beyond just English and Spanish.

Assessment Selection

Because these children live in the United States, they are likely exposed to at least some English in addition to their home language and thus are learning concepts in both languages (Espinosa & Crandell, 2020). Initially, children may only know a concept being assessed in one language. This likely depends on in which language they learned the concept (Hoff et al., 2012). Assessing concepts in only one language could therefore result in an underestimation of what DLLs know; their skills and knowledge may instead be distributed across both or all of their languages. Thus, in the First 5 DLL study it was critical to conduct the direct assessments in both English and the child's home language wherever possible. That is, the First 5 DLL study needed assessments that were available in English and Spanish, as well as in Mandarin, Cantonese, and Vietnamese. This approach of assessing language in both English and the child's home language is not common, although it is becoming more so (Hammer et al., 2014).

MOST PREVALENT NON-ENGLISH LANGUAGES SPOKEN BY KINDERGARTENERS IN CALIFORNIA, 2016-17

1. Spanish
2. Vietnamese
3. Mandarin
4. Arabic
5. Tagalog
6. Cantonese



¹ Mandarin and Cantonese utilize the same Chinese writing system, and its speakers share aspects of history, literature, and culture, but—linguistically speaking—Mandarin and Cantonese are distinct languages.

Moreover, when conducting assessment research with young children, the constructs of interest drive the assessment selection along with the age of the children. For example, the First 5 DLL study included infants, toddlers, and preschoolers, forcing a two-pronged approach. Infants and toddlers are best assessed indirectly through surveys administered to teachers and parents. Preschoolers, however, are old enough to be directly assessed through activities. Infants and toddlers in the First 5 DLL study were assessed for their home and English language skills, cognitive skills, social-emotional development, and motor development (as a proxy for general brain development). For preschool-aged children, the study focused more on school readiness skills, assessing home and English language development, preliteracy skills, early mathematics skills, executive function skills, and social-emotional development.

Finding “off-the-shelf” assessments of emerging language skills available in languages other than English and Spanish is challenging. Some assessments of Chinese language skills are available, but they are generally for older children who are learning Chinese as a second language and thus not appropriate for assessing the foundational language skills of a young child. Other direct assessments of language skills available in languages other than English and Spanish tend to be time consuming and difficult to administer. For example, the Early Language Listening and Oral Proficiency Assessment (ELLOPA) requires engaging each child in an approximately 20-minute conversation while coding the child’s use of speech (Center for Applied Linguistics, 2012). Even fewer assessments of skills such as literacy, mathematics, and social-emotional development are available in languages other than English and Spanish.

Some parent and caregiver report assessments are available in multiple languages, however. A well-established survey of infant and toddler language development available from the developers in multiple languages is the MacArthur-Bates Communicative Development Inventories (CDI; Fenson et al., 2007). The CDI is a survey in which parents and caregivers report on the language (gestures and words for infants and words and sentences for toddlers) that their child understands or uses. Additionally, two surveys of development for children ages birth to 5 years are available in multiple languages, the Ages and Stages Questionnaire-3 (ASQ-3; Squires & Bricker, 2009) and the Ages and Stages Questionnaire Social-Emotional 2nd Edition (ASQ-SE2; Squires, Bricker, & Twombly, 2015). The ASQ-3 surveys five domains of a young child’s development, including communication and physical development and the ASQ-SE2 thoroughly surveys the social-emotional development of young children. These surveys were incorporated in the First 5 DLL study, but not without challenges. These challenges will be discussed below.

Direct assessments for preschool children (ages 3 to 5 years old) in the content areas of language development and math development, as well as others, are still lacking and research is needed to develop and validate assessments on U.S. children in multiple languages. As these measures are not yet available, researchers are faced with trying to find a quick fix to this problem. One solution is purchasing tests that are already available from the publisher in English and Spanish and utilizing translation. Although the translations may not be equivalent in difficulty level and lack linguistic and cultural equivalence, many researchers find themselves utilizing translation (for example, Chung, Zhou, Anicama, Rivera, & Uchikoshi, 2019; Uchikoshi, 2013). This is what the First 5 DLL study had to do.

The First 5 DLL study decided to administer the Woodcock-Johnson III (Applied Problems) (Woodcock, McGrew, & Mather, 2001) and Woodcock-Johnson IV Oral Language (Picture Vocabulary and Oral Comprehension; Schrank, Mather, & McGrew, 2014) because all of the subtests were available in English and Spanish.² Moreover, the Spanish version of the subtests are not direct translations of the English items but rather psychometrically equivalent measures of their English counterparts. Consequently, a child administered both English and Spanish subtests would be tested on equated but different content, rather than seeing the “same” items twice as they would if the Spanish version were a direct translation of the English subtest. Because the First 5 DLL study sample had many Spanish-speaking children, it made sense to purchase assessments available in English and Spanish, and then translate the Spanish version into Mandarin, Cantonese, and Vietnamese. In this way, each child was assessed with *the same English assessment* and either the Spanish equivalent assessment or a translation of that Spanish equivalent assessment in their home language. In other words, the home language versions of the assessments were similar in that they were all translations of the Spanish form, and, importantly, the English and home language assessments covered equivalent content but were not repetitions of exactly the same content. Additionally, the First 5 DLL study opted to use item response theory (IRT) modeling to ensure equivalence in terms of psychometric properties and drop items that are not equivalent across languages.

Determining Dominant Language

If time and resources require the researcher to only field an assessment in one language to ensure the child’s understanding it is recommended to field the assessment in the child’s dominant language. In some cases, the child’s dominant language is clear, but given that DLLs living in the United States are often learning English alongside their home language, steps are needed to determine a child’s dominant language. One way to do this is to use measures designed to compare performance on the two languages. For example, the developers of the Woodcock-Johnson IV Oral Language battery created what they called the Comparative Language Index (CLI) to compare children’s scores on the Spanish and English versions of two subtests (Schrank et al., 2014). The CLI is calculated by creating a ratio of the child’s Spanish language score to the English score. If the score is less than 1, then the child is determined to be more proficient in English than Spanish. However, when assessing children who speak languages other than English and Spanish, a similar method can be used. For example, the First 5 DLL study compared scores on the two English subtests of the Oral Language battery to the two equivalent subtests in the child’s home language; the language in which the child scored higher was considered the dominant language. For children whose scores were equal across the four subtests, they were randomly assigned English or their home language as their dominant language. Table 1 indicates in which language the assessments were administered in the First 5 DLL study to preschoolers.

² For 3-year-olds in the sample, instead of fielding the Woodcock-Johnson II Applied Problems, the First 5 DLL study fielded the Woodcock-Johnson IV Early Cognitive and Academic Development Number Sense subtest to avoid floor effects. However, this subtest is available only in English.

Table 1. Study Design for Assessing Preschoolers in the First 5 DLL Study

Assessment Domains	Assessments Administered	Language	Order
Language skills	Woodcock-Johnson IV Oral Language	English and home language (Spanish, Mandarin, Cantonese, or Vietnamese)	Counterbalanced to vary whether English or home language was administered first
Math skills	Woodcock-Johnson III Applied Problems	English and home language	Counterbalanced to vary whether English or home language was administered first
	Woodcock-Johnson IV ECAD Number Sense ^a (3-year-olds only)	Dominant language	Presented just once
Executive function skills	MEFS	Dominant language	Presented just once
Preliteracy skills	Concepts of Story and Print	Dominant language	Presented just once
Social-emotional skills	ASQ-SE	Home language	Parent completed this survey in their language of choice

Note: ECAD is Woodcock-Johnson IV Tests of Early Cognitive and Academic Development. MEFS is Minnesota Executive Function Scale.

^aSchrank, McGrew, & Mather, 2015.

Assigning Language of Assessment

When working with young DLLs, a research consideration is determining which measures to field in which language. For surveys, generally the language of administration is determined by the type of respondent. Infant and toddler surveys administered to *parents* had to be available in the home language of the family and at a literacy level that all parents could understand. In contrast, *teacher*-report assessments could be fielded in English in most cases and literacy level was less of an issue. For example, Table 2 shows how the respondent type drove the language of surveys in the First 5 DLL study.

Table 2. Study Design for the Assessment of Infants and Toddlers in the First 5 DLL Study

Assessments Administered	Respondent Type	Language of Administration
CDI	Parent	Home language: Spanish, Mandarin, Cantonese, or Vietnamese
CDI	Teacher	English
ASQ-3	Teacher	English
ASQ-3	Teacher	English
ASQ-SE	Parent	Parent preference (English or home language)

When directly assessing bilingual children, ideally all measures would be fielded in both English and the child's home language to obtain a complete picture of what the child knows and can do in both languages. However, preschoolers do not have long attention spans and tire easily, so researchers need to think

carefully about what language to administer each measure in their assessment battery and in what order. For example, the First 5 DLL study decided to field assessments of language skills and mathematics skills in both English and the home language to obtain a comprehensive view of preschool DLLs' abilities across both languages in these domains. Although DLLs are developing two languages simultaneously, they are not always learning both at the same rate. Thus, assessing their language skills in both languages provides a better understanding of children's overall language skills than assessing in just one language. With respect to mathematics, assessing in both languages enabled us to understand what children knew in each language. For example, if mathematics was learned in an English-speaking early learning setting, then the chances were good that children would only know how to talk about these skills in English.

Some skills can be measured in only one language to minimize assessment length and burden. Executive function, for example, refers to higher-order skills involved in planning and executing behavior and so is not dependent on content knowledge; executive function skills can be assessed just once in one language as was done in the First 5 DLL study. Some measures are difficult to field in more than one language in a manner that is comparable, and sometimes modifications are needed to do so. For example, the First 5 DLL study team wished to measure concepts of story and print in which the assessor shares a book with the child and asks them to identify print features, such as the front of the book, the author's name, where on the page a reader should begin reading, and individual letters in the text. Study staff were able to administer the task in English using an English-language book for English-dominant children, and in Spanish using a Spanish-language book for Spanish-dominant children. However, the assessment questions did not apply to the Chinese writing system, which does not have individual letters. Thus, for children whose dominant language was Mandarin or Cantonese, the First 5 DLL study measured their understanding of *English* story and print concepts. For these children, the instructions and assessment items were presented verbally in their home language, but the stimulus book was an English-language book, something they would be familiar with from an English-language classroom. A similar approach was used for Vietnamese speakers because the book was not available in Vietnamese. Therefore, this assessment measures English print literacy.

Addressing the Lack of Assessments: Adaptation or Translation?

In the absence of good quality assessments and surveys in languages other than English and Spanish, researchers find themselves using adaptations and translations. The challenges to assessing young children using adaptations and translations are discussed below.

Adaptation

The developers of the MacArthur-Bates CDI make it clear that the Mandarin, Cantonese, and Vietnamese language versions of the assessment are not translations but rather are adaptations. They are designed to be culturally relevant and to reflect the expected language development milestones of each language. However, as the First 5 DLL study began reaching out to the CDI developers for these language versions, it was discovered that these adaptations were developed for use by people *living in the countries that speak*

these languages. These surveys are often used in the United States but doing so can be problematic as the language and culture of people who come to the United States tend to evolve in different ways from those who remained in the country of origin. For example, the Mandarin CDI was designed for children in Beijing, and the Cantonese CDI was designed for children living in Hong Kong. Given that new words and terms are always being added to local language, people who left China years ago would not necessarily know the more modern (or region-specific) words included in the Cantonese or Mandarin versions of the CDI. It is almost impossible to determine how well these instruments will work for children and families living in the United States, so when adapting assessments to other languages, they must be validated with the *population of interest*.

Translation

As mentioned, translating *from Spanish* (rather than from English) into the Asian languages allowed us to administer an assessment to each child in both English and their home language, but it came with many challenges. The items in Vietnamese and Chinese were merely translations of items that had been written for, and validated on, *Spanish-speaking children*. Translating into Chinese and Vietnamese brought with it both cultural and linguistic challenges.

Cultural Sensitivity

It is imperative to consider cultural differences when translating assessments into other languages. For example, in the Woodcock-Johnson III Applied Problems, there are characters in math story problems with names such as Pablo, Tomás, and Isabel. While these names are common in Spanish-speaking cultures, a Vietnamese or Chinese person is much less likely to be familiar with them. Care must be taken to adapt character names in story problems, as well as the topic of the story problem, to be culturally familiar to children in each language group.

As another example of the need for culturally appropriate translation, in the Spanish version of the Woodcock-Johnson IV Picture Vocabulary assessment, children are asked to name a *cornucopia*. Our Chinese reviewers expressed their concern that Chinese people would not know what this item was, due to their culture, *not* due to their lack of vocabulary. So, they suggested another item that was more relevant to Chinese culture, which the First 5 DLL study substituted for the cornucopia: the treasure pot. However, without testing large samples of both Spanish and Chinese speakers, the First 5 DLL study cannot be completely sure that this assessment item represents the same level of vocabulary difficulty in both languages. This is problematic because the Woodcock-Johnson items are designed to increase in difficulty level as the assessment progresses.

Linguistic Challenges

The structures of English and Spanish differ in fundamental ways from Mandarin, Cantonese, and Vietnamese such that it is impossible to simply translate from English or Spanish to one of these three

Asian languages and maintain the integrity of an assessment. For example, phonological awareness is an important early skill in English and Spanish, but this skill does not translate well to Mandarin or Cantonese. The Chinese logographic writing system does not employ the same relationship between a symbol (e.g., a letter) and a unit of sound (phoneme), as do English and Spanish; individual letters do not exist. Rather, in Chinese (both Mandarin and Cantonese), a *character* represents an entire word or part of a word. A character is not tied to a phonetic unit like a consonant or vowel. Thus, items in the English version of Concepts of Story and Print that ask a child to point to a letter they know and say the name did not translate well to Chinese. Furthermore, the direction of text and the identification of *words* varies across languages, so simple translation is not always feasible.

Recommendations for Researchers

In addition to challenges discussed above, the research team also became aware of many other nuances of language important to take into account in translation. Thus, if faced with the task of administering child assessments in Asian languages before adapted and normed assessments become available, researchers should consider the following recommendations before translating assessments. This can help avoid some of the pitfalls that the First 5 DLL study encountered.

1. Communicate who will administer the assessment and who will receive the assessment.

Translators need to know if the project has hired younger versus older assessors, or assessors familiar to the children versus strangers, because these assessor characteristics may require the use of different pronouns in certain languages (e.g., Vietnamese). Translators also need to know if the assessment is to be administered to each child individually or to an entire classroom full of children, which can also affect pronouns and grammatical particles in languages such as Vietnamese.

2. Clarify modality and format of assessment. Let translators know whether the items will be read *out loud* during the assessment. This can have implications for phrasing, for example, in Cantonese. Converting Standard Written Chinese into spoken Cantonese requires both grammatical and lexical changes.

3. Establish literacy and education levels of the study population. Professional translations are often at a level too high to be appropriate for the population, so care must be taken to avoid this problem.

4. Hire translators fluent in both source language (the language of the original assessment) and the target language (the language into which the assessment is being translated) if possible. Whenever possible, avoid translating into an intermediary language. Hire translators who can translate directly from the source language (e.g., Spanish) into the target language (e.g., Vietnamese), rather than from Spanish into English and then again from English into Vietnamese.

5. Understand target language writing systems. The Chinese writing system has evolved to include simplified versions of many traditional characters, which require fewer brush strokes. Mandarin and

Cantonese speakers may know and use either system, or both. For assessments written in either Mandarin or Cantonese, make sure it is clear whether the assessors read traditional or simplified Chinese characters, or pinyin (a Romanization system for Chinese characters). Both simplified and traditional versions of Mandarin *and* Cantonese instruments may be needed.

“Door” (traditional characters)	“Door” (simplified characters)
門	门

6. Insist on back-translation of assessments into the original source language, by a native speaker of the source language. This will help to ensure the integrity of the intention of each item, as well as to check for errors in details such as the numbers used in math items (a mistake in translation of a number could lead a child to answer the item incorrectly).

7. Consult with a regional language expert and conduct pilot testing. Until assessments are developed locally in many different languages for bilingual American children, it is imperative that different language versions of each assessment are reviewed before entering the field to ensure that they represent the way children in a study sample are likely to speak. A regional language expert would recognize terms that are unique to particular geographic regions and can confirm whether the language of the assessment is likely to be in the dialect that the children in your sample speak (e.g., northern Vietnamese dialect versus southern Vietnamese dialect). A regional language expert can recommend wording that is common to the people from a particular area or recommend allowing for multiple correct responses. For example, the Spanish word for “corn” is different in different Latin American countries (*maíz, elote, choclo*), so children living in the United States may come from varying regional language backgrounds and use different words for “corn,” all of which are correct. Pilot test the assessments with children *from the same regions as the children in your sample* to ensure understanding. Any confusing or unclear language can be identified and revised through pilot testing.

Future Directions

Evidence is growing that using direct translations and adaptations from other countries as well as tools normed in other countries may not yield accurate assessment scores for all U.S. children. The research community needs to create assessments that are culturally sensitive and shown to be valid for each language group living in the United States. These assessments need scores that are equated in all available languages so they can be combined and compared. The assessments will allow all children entering early education and care to be more accurately assessed; *their language will no longer prevent them from giving their best answer*. Precise measures of what children know and can do are essential for determining proper support and interventions, and ensuring a more level playing field for all children to succeed. Researchers need the help of assessment developers and translators to remove the language barrier and allow all U.S. children to be properly evaluated upon entering early education and care.

Thus, the authors of this paper submit a call to action for more widely available assessments in English, Spanish, and other prevalent languages spoken in the United States that meet the following criteria:

- **Assessments must be normed and validated on populations of children who speak the target language.** To ensure that the norms represent the population, it is important to adapt the assessments to different languages and then test and norm them *with DLLs living in the United States*. Then, the results will better generalize to these children. As mentioned, simply using a translation of each item on an English language assessment runs the risk that each item is not necessarily at the exact same difficulty level in English and in the translation. For example, on a picture vocabulary assessment, a particular word in English may be considered to be at a lower or higher difficulty level than the translation of that word into Vietnamese, due to phonetic, phonological, syntactic, or semantic differences between languages.

- **Assessments must be normed and validated on children who live in the same country as the study children.** Assessments conducted on children in the United States need to be tested and validated on children living in the United States, so that the questions and scenarios put forth make sense to children living in the U.S. context. For example, one of the vocabulary words on the Vietnamese CDI is “bim bim,” which is the brand name of a candy available only in northern Vietnam and may not be known to Vietnamese-speaking children growing up in the United States. It does not make sense to include a vocabulary item such as this on an assessment intended for children in the United States.

ASSESSMENTS MUST:

- be normed and validated on populations of children who speak the target language,
- be normed and validated on children who live in the United States,
- reflect the cultures being represented, and
- be equated to ensure equivalent difficulty levels and comparability of scores.

- **Assessments must reflect the cultures being represented.** All aspects of culture must be considered so that unfamiliarity with a certain cultural aspect will not be conflated with lack of knowledge of the subject matter being measured. For example, appropriate character names should be used in story problems, and relevant objects should be used in vocabulary assessments.

- **Assessments must be equated to ensure equivalent difficulty levels and comparability of scores.** All versions of an assessment in different languages must be equated so that researchers, educators, and practitioners can combine scores and look at outcomes for children from different language backgrounds on one scale. Equated scores will make it easier to accurately identify children who need additional learning support.

This area of research is growing, and those who engage in this work need to collaborate with translators and linguists to improve the measures available to us as a field. Researchers and practitioners need to learn from one another so that mistakes are not repeated. Together, the field can work to provide fellow

researchers, educators, and practitioners with the tools they need to accurately measure young children's knowledge and skills.

References

- Abedi, J. (2002). Standardized achievement tests and English language learners: Psychometrics issues. *Educational Assessment*, 8(3), 231–257. https://doi.org/10.1207/S15326977EA0803_02
- California Department of Education Data Reporting Office. (2019-2020). *English learner students by language by grade*. <https://data1.cde.ca.gov/dataquest/SpringData/StudentsByLanguage.aspx?Level=State&TheYear=2019-20&SubGroup=All&ShortYear=1920&GenderGroup=B&CDSCode=00000000000000&RecordType=EL>
- Center for Applied Linguistics. (2012). *Early Language Listening and Oral Proficiency Assessment for Grades PreK-2 (ELLOPA)* [Paper presentation]. American Council on the Teaching of Foreign Languages (ACTFL), Philadelphia, PA, United States.
- Chung, S., Zhou, Q., Anicama, C., Rivera, C., & Uchikoshi, Y. (2019). Language proficiency, parenting styles, and socioemotional adjustment of young dual language learners. *Journal of Cross-Cultural Psychology*, 50(7), 896–914. <https://doi.org/10.1177/0022022119867394>
- Espinosa, L., & Crandell, J. (2020). Early learning and care for multilingual and dual language learners ages zero to five. In California Department of Education (Ed.), *Improving education for multilingual and English learner students: Research to practice* (Ch. 4). California Department of Education. <https://www.cde.ca.gov/sp/el/er/improvingmleeducation.asp>
- Federal Interagency Forum on Child and Family Statistics. (2015). *America's children: Key national indicators of well-being 2015*. U.S. Government Printing Office. https://www.childstats.gov/pdf/ac2015/ac_15.pdf
- Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories: User's guide and technical manual* (2nd ed.). Brookes.
- Garcia, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly*, 26, 371–391.
- García, O., & Torres-Guevara, R. (2010). Monoglossic ideologies and language policies in the education of U.S. Latinas/os. In E. Murillo, S. Villenas, R. Trinidad Galván, J. Sánchez Muñoz, C. Martínez, & M. Machado-Casas (Eds.), *Handbook of Latinos and education: Research, theory and practice* (pp. 182–194). Lawrence Erlbaum.
- Guzman-Orth, D., Lopez, A. A., & Tolentino, F. (2017). A framework for the dual language assessment of Young dual language learners in the United States. *ETS Research Report Series*, 2017(1), 1–19.
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29, 715–733. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4394382/>
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39, 1–27.

- Holtby, S., Lordi, N., Park, R., & Ponce, N. A. (2017). Families with young children in California: Findings from the California Health Interview Survey, 2011–2014. *American Journal of Medical Research*, 4(2), 168–178.
- National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the educational success of children and youth learning English: Promising futures*. The National Academies Press. <https://doi.org/10.17226/24677>
- Park, M., O’Toole, A., & Katsiaticas, C. (2017). *Dual language learners: A national demographic and policy profile*. Migration Policy Institute. <https://www.migrationpolicy.org/research/dual-language-learners-national-demographic-and-policy-profile>
- Peets, K., & Bialystok, E. (2015). Academic discourse: Dissociating standardized and conversational measures of language proficiency in bilingual kindergarteners. *Applied Psycholinguistics*, 36(2), 437–461. <https://doi.org/10.1017/S0142716413000301>
- Samson, J. F., & Lesaux, N. K. (2009). Language-minority learners in special education: Rates and predictors of identification for services. *Journal of Learning Disabilities*, 42(2), 148–162.
- Schrank, F. A., Mather, N., & McGrew, K. S. (2014). *Woodcock-Johnson IV tests of oral language*. Riverside Publishing Company.
- Schrank, F. A., McGrew, K. S., & Mather, N. (2015). *Woodcock- Johnson IV tests of early cognitive and academic development*. Riverside Publishing Company.
- Snow, C. E., & Van Hemel, S. B. (2008). *Early childhood assessment: Why, what, and how*. The National Academies Press.
- Squires, J., & Bricker, D. (2009). *Ages & Stages Questionnaires®, Third Edition (ASQ®-3): A parent-completed child monitoring system*. Paul H. Brookes Publishing Co., Inc.
- Squires, J., Bricker, D., & Twombly, E. (2015). *Ages & Stages Questionnaires®: Social-Emotional, Second Edition (ASQ®:SE-2): A parent-completed child monitoring system for social-emotional behaviors*. Paul H. Brookes Publishing Co., Inc.
- Sugarman, J., & Villegas, L. (2020). *Native language assessments for K–12 English learners: Policy considerations and state practices* [Policy brief]. Migration Policy Institute. https://www.immigrationresearch.org/system/files/MPI-native-lang-assessments_FINAL.pdf
- Uchikoshi, Y. (2013). Predictors of English reading comprehension: Cantonese-speaking English language learners in the U.S. *Reading and Writing*, 26, 913–939. <https://doi.org/10.1007/s11145-012-9398-z>
- Vitiello, V., Downer, J., & Williford, A. (2011). Preschool classroom experiences of dual language learners: Summary of findings from publicly funded programs in 11 states. In C. Howes, J. T. Downer, & R. C. Pianta (Eds.), *Dual language learners in the early childhood classroom* (pp. 69–91). Brookes.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson® III Test*. Riverside Publishing Company.

About the First 5 California DLL Pilot Study

In 2015, First 5 California committed \$20 million for the DLL Pilot Study to support effective and scalable strategies in early learning and care programs to promote learning and development for DLLs and their families. A key component of this overall initiative seeks to describe and evaluate the range of strategies to support DLLs, including three strategies of particular interest: instructional practices, PD for early educators, and family engagement. The study is examining the range of practices, by age, setting type, and diverse language groups, and how various practices are supportive of child and family outcomes. The study includes 16 counties selected to be broadly representative of the state's DLL population: Butte, Calaveras, Contra Costa, Fresno, Los Angeles, Monterey, Orange, Riverside, Sacramento, San Diego, San Francisco, Santa Barbara, Santa Clara, Sonoma, Stanislaus, and Yolo. The study is being conducted by AIR and its partners at Juárez & Associates; CRI; School Readiness Consulting; Allen, Shea & Associates; and Stanfield Systems, Inc.; with guidance from the DLL Input Group, which comprises stakeholders, advocates, and state and national experts on DLLs.

For more information about the study and to read other study briefs and reports:

<https://californiadllstudy.org/>

www.cafc.ca.gov/

Copyright © 2021 American Institutes for Research®. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, website display, or other electronic or mechanical methods, without the prior written permission of the American Institutes for Research. For permission requests, please use the Contact Us form on www.air.org.



2800 Campus Drive | Suite 200
San Mateo, CA 94403
650.376.6300 | www.air.org



2389 Gateway Oaks Drive | Suite 260
Sacramento, CA 95833
916.263.1050 | www.first5california.com