RESEARCH ARTICLE

The Psychometric Properties of the English and Spanish Versions of the Bilingual Early Language Assessment

Lillian K. Durán, Mark S. Innocenti, and M. Brooke Robertshaw

Utah State University

The Bilingual Early Language Assessment (BELA) was developed by the Harvard Graduate School of Education for the Cambridge Public School District in 2002 to respond to a growing need for an assessment in multiple languages. The BELA is available in 11 different languages and can be downloaded at no cost. However there is no psychometric data available on the BELA to inform end users regarding the quality of the measure. This paper reports on a pilot study investigating the psychometric properties of the English and Spanish BELA including the internal consistency, test-retest reliability, and concurrent validity of the BELA with the Preschool Language Scale-4.

The population of Dual language Learners (DLLs) in Head Start has dramatically risen over the last 10 years. The complexity of providing culturally and linguistically responsive services to this growing population in Head Start is outlined in the report "Dual Language Learning: What does it take?" (U.S. Department of Health and Human Services, 2008). The report describes the findings of a national assessment of Head Start programs that included focus groups, conference calls, and multiple meetings with individuals within the Head Start community to determine program needs for serving DLLs. A major finding involves the need for better assessments:

> Local programs are required to conduct developmental screenings and ongoing assessments of enrolled children. Yet, there exist few valid and reliable assessment instruments for evaluating progress in language and literacy development, as well as development in other domains of learning for children who are learning two languages. Without accurate assessment information, staff are not able to properly support the child's development, identify progress, individualize the curriculum fully, or identify behavior or delays requiring further evaluation and possible intervention.
> (U.S. Department of Health and Human Services, 2008 p. 5).

Complicating this task further is the sheer number of languages spoken by children enrolled in Head Start nationally. Over 140 different languages are represented in the Head Start

Lillian K. Durán, Special Education, Utah State University; Mark S. Innocenti, Center for Persons with Disabilities, Utah State University; M. Brooke Robertshaw, Instructional Technology, Utah State University.

Correspondence concerning this article should be addressed to Lillian K. Durán, Department of Special Education and Rehabilitation, Utah State University, 2865 Old Main Hill, Logan, UT 84322-2865. E-mail: Lillian.duran@usu.edu

population with trends suggesting this diversity in languages may increase (U.S. Census Bureau, 2010; U. S. Department of Health and Human Services, 2008). Head Start programs are mandated to both screen all children in the program for developmental delays and to keep data on the progress of each child served. This requires Head Start staff have access to technically adequate screening and progress monitoring tools to use with *all* of the children enrolled. This also means that staff need to be able to determine the assessment language or languages to accurately capture children's abilities. In order to fulfill this mandate more effort needs to be invested in developing language proficiency measures and screening and progress monitoring tools that are available in languages other than English. Although there are psychometrically sound measures that can assess skills in English and Spanish (Barrueco, Lopez, Ong, & Lozano, 2012), gaps still exist in their psychometric characteristics and, more importantly, usability by Head Start and other programs. Measures for DLLs other than those in Spanish/English environments are very limited. There remains a critical need for measures that meet the Head Start mandate.

The focus of this paper is on one measure that is currently available in 11 different languages. The Bilingual Early Language Assessment (BELA: Tabors & Heise-Bagorria, 2004) was designed for use with 2.9 to 5 year olds for the Cambridge Public Schools. It provides a general measure of ability in the child's home language and English on language and conceptual skills. It is also designed as a progress monitoring tool that can be administered several times over the school year. The BELA can help practitioners understand what concepts children have in their home language and in English because it taps the exact same skills in each language. The test is not a screening tool or a formal language proficiency measure. Instead, it is meant to provide information to guide instruction. Little information is available regarding the development of the instrument, however the measure was presented at the National Dual Language Head Start Institute in 2008 (Heise-Bagorria, 2008*).

The test is available at no cost on-line, however there are materials that need to be purchased in order to administer the instrument (www.cpsd.us/BELA). The BELA is currently available in Arabic, Bangla, Chinese, English, Haitian-Creole, Portuguese, and Spanish. The Minnesota Department of Education has also recently translated the tool into Hmong, Oromo, Russian, and Somali and these versions are also available at no cost on-line (http://education.state.mn.us/MDE/Learning_Support/Early_Learning_Services/Early_Childhood_Programs/Help_Me_Grow_Prog_Serv/Administration/index.html). The authors of the BELA explicitly encourage translation of the BELA into other languages that can then be posted on their website.

There are currently few measures related to instruction and not diagnostics available for use with preschoolers that specifically measure bilingual children's' abilities in Spanish and English. The tools that measure language proficiency related to instructional decision-making most widely used with young Spanish speakers include the Pre-IDEA Proficiency Test (Pre-IPT; Ballard, Tighe, & Dalton, 1991), the Pre-Language Assessment Scale (Pre-LAS; Duncan & DeAvila, 2000), and the Woodcock-Muñoz Language Survey-Revised (WMLS-R, Alvarado, Ruef, & Schrank, 2005). Although these measures are widely used they all have weaknesses that interfere with confidence in the validity of their use with young children. One weakness that is common to all of them is their reliance on measuring discrete aspects of language (i.e. phonology, syntax, morphology, & lexicon) without including any context-based or pragmatic communication which is especially important in early language development (Esquinca, Yaden, & Rueda, 2005). The Pre-LAS has also been shown to have weak validity in determining a

child's language abilities in Spanish and English and it has been found to over-identify children as "non-speakers" of both languages (MacSwan, Rolstad, & Glass, 2002). The WMLS-R is designed for use with 2-to 90-year-old individuals and the number of items at the three-to-five-year-old level is limited. This is true for many of the WMLS-R subtests that require the child to read by the third or fourth item on the test (e.g., the passage comprehension subtest). The few items available result in many children quickly reaching the assessment ceiling, frequently without establishing a basal, which limits findings and does little to facilitate instructional decision-making. Another significant issue is that these tools are available in only Spanish and English which leaves many languages for which Head Start practitioners have no measures available.

Head Start personnel also need to be provided with detailed information about appropriate assessment procedures for DLLs. It is currently considered best practice to measure bilingual children in each of their languages to more accurately describe their abilities across their languages. (NAEYC, 2005; Paradis, Genesee, & Crago, 2010; Peña & Halle, 2011). A child's overall abilities may be underestimated if only assessed in one of their languages (Bedore, Peña, García, & Cortez, 2005; Boyce, Gilliam, Innocenti, Cook, & Ortiz, in press; Peña & Kester, 2004). Children who speak more than one language will have skills distributed across all of their languages based on the context in which they use that language and the quality and quantity of input they have had in each (Hammer, Miccio, & Rodriguez, 2004). Young Spanish-English bilingual children in the U.S. have been found to have different vocabulary distributed across both of their languages (Pearson, Fernández, & Oller, 1993). If a child's vocabulary in one language is compared to monolingual norms they may appear delayed, but if the child's total vocabulary is added across both languages their total score might indicate no delay (Peña, Bedore, Rapazzo, 2003; Peña, Gillam, Bedore, & Bohman, 2011). In order to accomplish this however, valid and reliable assessment tools must be available in each of the languages spoken by the children to be assessed to accurately measure their overall language ability.

The purpose of this paper is to provide preliminary psychometric data on the Spanish and English versions of the BELA (Tabors & Heise-Bagorria, 2004). The vast majority (86%) of DLLs in Head Start speak Spanish. Providing a tool that can provide basic information to teachers about children's abilities in both languages to guide instruction and to be used across the year to monitor the children's progress in each language would be a valuable addition to the measures currently available. The BELA is starting to be used in Head Start programs (e.g. Minnesota) while little is known about its psychometric properties. This paper addresses three specific questions regarding the psychometric properties of the BELA.

1. What is the internal consistency of the Spanish and English BELA?
2. What is the test-retest reliability of the Spanish and English BELA?
3. What is the concurrent validity of the BELA with the Preschool Language Scale-4 (Zimmerman, Steiner, & Pond, 2002) in English and Spanish?

METHODS

Participants

Thirty-nine preschool children (ages 40 to 63 months; M = 50.79; Female: n = 17) in four classrooms, attending one Migrant Head Start school located in a rural community in the intermountain region of the U.S. participated in the study. Every child's parents completed a family language questionnaire either in Spanish or English (see Appendix A) and responses indicated that all used Spanish in their home with at least one primary care provider. All families were from Mexico. Based on data gathered on the family language questionnaire children were categorized into three broad language proficiency groups: English Dominant (ED, n=6); Spanish Dominant (SD, n=21); and Simultaneously Bilingual (SB, n=9).

Setting

All children were enrolled in a Migrant Head Start preschool program in the rural intermountain west. The preschool sessions were all full-day, full-time (5 days/week) and implemented a Montessori curriculum. Children were in four classrooms. Each classroom was staffed with one lead teacher and a teacher aide for 16-18 children. All four teachers were trained in the Montessori method, and all held associate degrees in child development. One classroom had a Spanish-speaking lead teacher and three had Spanish-speaking aides. There was wide variation across all four classrooms in the amount of Spanish and English used throughout the day and the activities in which Spanish and English were more likely to be used.

Procedures

Thirty-nine Spanish-speaking preschoolers were administered the BELA (Tabors & Heise-Bagorria, 2004) and Preschool Language Scale-4th Edition (PLS-4; Zimmerman, et al., 2002) in English and Spanish. The BELA was administered twice in each language to each child in November and then again in mid-December or early January due to the holiday break. All test administrators were provided two hours of training that included familiarity with the assessment measure, procedures for assessing young DLLs, and discussions of dialectical differences that might arise given the diversity in the Spanish speakers who were administering the assessments. If children provided responses in Spanish that were unfamiliar to the assessors they were to write down the response and then later confer with the child's Spanish-speaking teacher to discuss the dialectical appropriateness of the response. The lead author observed the first two administrations of the assessment by each assessor to ensure fidelity to the protocol and appropriate testing behaviors with young children. Assessors received explicit instructions to use only one language during the entire assessment. At no time was the same test given in English and Spanish on the same day to any child. Each assessor was assigned to either English or Spanish and did not switch to test the other language at any point so that children would clearly identify them as either English or Spanish speakers.

The English BELA was administered by four monolingual English-speaking undergraduates majoring in early childhood education or special education. The Spanish version was administered by one graduate research assistant and three undergraduates. The one graduate

research assistant was majoring in early childhood special education. She held a bachelor's degree in Spanish, had two years of experience working and living in Spanish-speaking countries including Spain, Mexico, and Nicaragua. The three undergraduates were majoring in business and international relations and were all native Spanish-speakers with two international students from the Dominican Republic and one U.S. born Colombian American.

The PLS-4 was administered once in English and Spanish in mid-December or January. The lead author and one graduate student in sociology administered all English PLS-4 assessments. The Spanish version was administered by four assessors including three native Spanish speakers; one held a degree in Elementary Education from Mexico and was previously employed by Migrant Head Start as a teacher. Two were undergraduate students; one was from the Dominican Republic and the other was a Mexican American with Spanish as her first language. One undergraduate student was majoring in business and the other in elementary education. The last assessor was the same graduate student with a degree in Spanish who also administered the BELA in Spanish.

Measures

The PLS-4 (Zimmerman et al., 2002) in English and Spanish includes the Auditory Comprehension (AC) and Expressive Communication (EC) subscales. The Spanish PLS-4 was normed on 1,188 Spanish-speaking children from all regions of the U.S. representing a broad range of Spanish dialects with both simultaneous and sequential bilinguals. Reliability coefficients range from .91 to .92 for 4- to 5-year olds.

The English PLS-4 standardization sample included 1,564 children with 18.1% being Hispanic. The PLS-4 English version also includes guidelines for scoring responses from children whose English is influenced by Spanish under a section in the examiner's manual entitled, "scoring guidelines for dialect speakers" (Zimmerman et al., 2002, p.19). Composite reliability coefficients range from .94 to .96 for 4- to 5-year-olds. Validity of the English PLS-4 with the Peabody Picture Vocabulary Test-3 (PPVT-3; Dunn & Dunn, 1997) was also reported as acceptable in a study including both English-speaking Hispanic and European American Head Start children (Qi & Marley, 2011). The PLS-4 was found to predict scores on the PPVT-3 equally well for both groups.

The BELA (Tabors & Heise-Bagorria, 2004) was developed by the Harvard Graduate School of Education for the Cambridge Public School District's early intervention providers as a tool for teachers to measure both basic language and pre-academic skills in the child's home language and English. The BELA was introduced at the National Head Start Dual Language Institute (2008), providing an endorsement for its use. The information obtained from this assessment is meant to guide both home-based and preschool intervention. There is no published information available on the development of the test or any publically available data available to describe its psychometric properties. In addition to the administration manual, test administration directions, scoring protocols, and a list of materials needed to administer the assessment, there is also a training CD available at no cost (http://www.cpsd.us/bela). The BELA includes 10 receptive and 9 expressive items. All items receive a dichotomous score of 1 or 0 except for item 2.1 in which you add the total number of correct responses with 4 being the maximum score. The items on the receptive subtest include tasks requiring the child to

identify common objects, colors, size, quantity, shapes, and actions. On the expressive subtest children are asked to:

a) Provide basic personal information such as their name, age, favorite color and play activity.
b) Repeat sounds, phrases and sentences.
c) Label colors, body parts, common objects, and actions in response to stimulus items.
d) Rote count to ten.

Overall the items on the BELA represent a sampling of tasks found on many popular early language assessments commonly used in the field.

The Spanish version is a direct translation of the English version and most items are exactly the same on the Spanish and English versions with the exception of the items requiring the child to repeat sounds, phrases and sentences. No published information could be found regarding the procedures followed in developing the Spanish version.

Data Analysis

Data analysis proceeded in four stages. First we conducted an analysis of descriptive statistics including means, standard deviations, and score ranges on the BELA and PLS-4 subscale and total scores in English and Spanish. The descriptive statistics are provided in Table 1. Given the dichotomous scoring of each item on the BELA, we calculated the internal consistency of the BELA using the Kuder-Richardson 20 (KR-20) formula (Pedhazur & Schmelkin, 1991). We used the Spearman's Rho formula to calculate the test-retest reliability of the BELA given the non-normal distribution of the data (Sheskin, 2004). Lastly, to more accurately investigate the concurrent validity of the BELA with the PLS-4 we first converted scores on both instruments to z-scores to standardize a scale between the two instruments. We then calculated the correlation between scores on the two instruments' expressive and receptive language sections as well as the total scores in English and Spanish using the Spearman's Rho formula (Sheskin, 2004). We also conducted a factor analysis to investigate how subscales (i.e. receptive and expressive scales) between the two measures in both languages grouped together.

TABLE 1
Descriptive Statistics for scores* on Time 1 and Time 2 of the BELA and PLS-4

Assessments		Mean	*SD*	Range
Time 1:	BELA English			
	Receptive	6.97	2.90	0-10
	Expressive	5.46	3.69	0-11
	Total	12.49	6.13	0-21
Time 1:	BELA Spanish			
	Receptive	8.46	1.45	10-May
	Expressive	7.03	2.97	12-Jan
	Total	15.51	3.97	22-Aug
Time 2:	BELA English			
	Receptive	7.95	2.05	10-Mar
	Expressive	7.00	3.20	12-Jan
	Total	14.95	4.88	22-Jun
Time 2:	BELA Spanish			
	Receptive	8.49	1.39	10-May
	Expressive	7.69	2.72	12-Jan
	Total	16.18	3.74	22-Sep
PLS-4:	English			
	Auditory comprehension	41.21	9.50	19-57
	Expressive	38.00	12.32	20-63
PLS-4:	Spanish			
	Auditory	47.15	4.91	36-57
	Expressive	46.23	8.14	25-59

* Raw scores are presented.

RESULTS

Preliminary Analysis

In general there were broad ranges in performance on both the BELA and the PLS-4 in both English and Spanish. The standard deviations as well as the range in scores indicate a significant amount of variability in the Spanish and English skills in this group of Spanish-speaking preschoolers. This is particularly evident in the expressive communication scores in both English and Spanish on the PLS-4. This finding is consistent with other studies of young Spanish speakers as performance in each language is dependent on the quantity, quality, and timing of exposure in each language and the specific skills that are being measured in each language (Kohnert, Bates, & Hernandez, 1999; Kohnert & Bates, 2002; Kovelman, Baker, & Petitto, 2008; Páez, Tabors, & López, 2007; Tabors, Páez, & López, 2003). Overall however the group had stronger language skills in Spanish.

Internal Consistency of the BELA

The KR-20 values were calculated for the time 1 administration of both the English and Spanish BELA. The English version of the BELA showed acceptable internal consistency reliability on the total scale (α=.90), the receptive subtest (α=.86), and the expressive subtest (α=.81). The Spanish BELA had low levels of reliability with KR-20 values of α=.53 on the receptive subtest, α=.49 on the expressive subtest and α=.59 on the total scale. Reliability scores between α=.7 and α=.79 are considered acceptable for exploratory work, while scores above α=.8 are considered acceptable for non-exploratory work (Nunnally & Bernstein, 1994). See Table 2 for results.

TABLE 2
Internal Consistency

Subscale	KR-20 (α)
BELA English	
Receptive	0.86
Expressive	0.81
Total	0.90
BELA Spanish	
Receptive	0.53
Expressive	0.49
Total	0.59

Test-Retest Reliability of the BELA

The test-retest reliability of the BELA was calculated using Spearman's Rho (ρ) between the first and second administration of the test; approximately one month. The English receptive subtest received a test-retest reliability coefficient of ρ=.75. The expressive subtest was ρ=.80 and the overall total scale achieved a value of ρ=.85. The Spanish receptive subtest received a value of ρ=.69, the expressive subtest ρ=.84 and the total scale was ρ=.85. See Table 3 for all results.

TABLE 3
Test-Rest Reliability

Subscale	ρ
BELA English	
Receptive	0.75
Expressive	0.80
Total	0.85
BELA Spanish	
Receptive	0.69
Expressive	0.84
Total	0.85

Concurrent Validity

Concurrent validity is a form of validity that compares a test against a benchmark test to examine whether the two tests are measuring the same construct (Nunnally & Bernstein, 1994). Concurrent validity analyses were performed on the PLS-4 and the second administration of the BELA as they were administered during the same time frame. Each child's performance on the PLS-4 to the BELA was compared on both the expressive and receptive subtests through correlation analyses and an exploratory factor analysis using z-scores (DeVellis, 2003). Validity between the PLS-4 and BELA was initially calculated using Spearman's Rho (ρ) and these results are presented in Table 4. Spearman Rho values for related subscales (e.g., English BELA expressive with English PLS-4 Expressive Communication) ranged from $\rho=.69$ to $\rho=.84$. These relations demonstrate acceptable validity correlations.

TABLE 4
BELA and PLS-4 Correlations

BELA		PLS-4	
		Expressive Communication (ρ)	Auditory Comprehension (ρ)
English			
	Expressive	0.69	
	Receptive		0.77
Spanish			
	Expressive	0.71	
	Receptive		0.84

An exploratory factor analysis was also completed due to a) the exploratory nature of this psychometric investigation, b) the low reliability on the Spanish BELA, and c) the small sample size. Our hypothesis was that a four factor solution would explain the most variance. The four factors we hypothesized that would load together were: English BELA receptive and English PLS-4 Auditory Comprehension, Spanish BELA receptive and Spanish PLS-4 Auditory Comprehension, English BELA Expressive and English PLS-4 Expressive Communication, and Spanish BELA Expressive and Spanish PLS-4 Expressive Communication. An orthogonal, varimax, rotation was applied to the solution in order to maximize the amount of variance on each factor and variable (Nunnally & Bernstein, 1994). Three rounds of analysis were completed and factor loadings above .4 were considered acceptable (DeVellis, 2003). The first round was purely exploratory and no set number of factors were set; a two factor solution was produced. Based on our hypotheses, the second analysis was set to 4 factors. This did not produce acceptable loadings for all subscales. A third solution was completed to confirm the two factor solution produced in the first analysis. As can be seen in Table 5, the English subscales on the BELA and PLS-4 loaded together, and the Spanish subscales on the BELA and PLS-4 loaded together. Factors did not result that loaded the receptive and expressive subscales on the two measures in each language respectively as we had initially hypothesized.

TABLE 5
Factor loadings for examination of concurrent validity between the BELA and PLS-4

Scale	Factor 1	Factor 2
BELA English Receptive	0.86	
BELA English Expressive	0.86	
BELA Spanish Receptive		0.51
BELA Spanish Expressive		0.79
PLS-4 English Expressive Communication	0.79	
PLS-4 English Auditory Comprehension	0.80	
PLS-4 Spanish Expressive Communication		0.80
PLS-4 Spanish Auditory Comprehension		0.75

DISCUSSION

The primary objective of this study was to provide pilot data on the psychometric properties of the BELA (Tabors & Heise-Bagorria, 2004). We specifically investigated the internal consistency, test-retest reliability, and the concurrent validity of the BELA with the PLS-4. However, given the small sample size included in this study these data can only provide preliminary evidence regarding the validity and reliability of the Spanish and English versions of the BELA with Spanish-speaking preschoolers in Head Start.

The English BELA demonstrates good internal consistency with KR-20 (α) values ranging from .81 to .90. The Spanish BELA has moderately low internal consistency with the receptive section receiving a value of α=.53 and expressive section α=.49. These findings support the construction of the English BELA, but raises questions about the Spanish BELA. More information is needed on the Spanish version given its lack of adequate internal consistency. It is not currently considered best practice in bilingual measurement to simply translate an instrument into another language without considerations of key linguistic and cultural features that might influence performance on the assessment (Anderson, 2002). In order to improve the internal consistency of the Spanish BELA it may be necessary to redesign the instrument and to develop a version that considers features unique to Spanish language development and to the cultural contexts of young Spanish speakers in the U.S. (Peña, 2007). This may also have implications for the translation of the BELA into languages other than Spanish as well.

Test-retest reliability was acceptable across both subtests and in both languages ranging from ρ=.84 to ρ=.69. Although it would be preferable to have reliability coefficients above .80, Nunally and Bernstein (1994) suggested the reliability levels similar to those achieved in this study are acceptable for preliminary research. However, more research with a larger sample is needed and the BELA should be used cautiously until it meets approved standards in the field (Standards for Educational and Psychological Testing, 1999).

The BELA holds promise as a progress monitoring tool. This is encouraging given the dearth of publically available progress monitoring tools available for use with young Spanish-English bilinguals. The tools that are publically available in Spanish and English such as the Get Ready to Read! (GRTR; Lonigan, 2003; Lonigan & Wilson, 2008), and the *Circle Phonological Awareness Language and Literacy Screener (CPALLS;* Landry, Assel, Gunewig, Swank, 2004) focus exclusively on discrete early literacy skills such as one word picture naming and

phonological awareness. The BELA includes items that tap pre-academic content knowledge about colors, shapes, size, and numbers. Additionally, the BELA includes items that probe language skills beyond picture naming such as repeating sounds and phrases and answering open-ended questions such as, "What do you like to do?" Measuring children's performance in these broader areas could prove to provide more information about the child's overall abilities, knowledge, and level of development in each language which in turn could facilitate better instructional decision-making in areas other than early literacy.

The BELA showed promising concurrent validity with the PLS-4. The correlational and exploratory factor analyses showed that the English language BELA and English PLS-4 and the Spanish language BELA and Spanish PLS-4 measure similar constructs. The correlational analyses demonstrated expected patterns. However, the factor analysis did not show discrimination among the subscales beyond grouping English and Spanish language domains. This suggests promise for future validation studies. A larger sample is needed to confirm these findings, and to explore whether the subscales, by language, will show groupings as hypothesized.

The PLS-4 was specifically selected as the criterion measure in this study given the rigor of its validity and reliability, the extensive use of the instrument in the field with young Spanish-English bilinguals, and the careful attention the authors paid to cross-linguistic issues in the development of both the Spanish and English versions. Importantly, the authors of the English PLS-4 included Spanish-English bilinguals in the standardization sample and they provide scoring guidance for Spanish influenced English (Zimmerman et al., 2002). Recently Qi and Marley (2011) also investigated the validity and reliability of the English PLS-4 with English-speaking Latino preschoolers in Head Start. It is important to note that all children in the study spoke only English and therefore findings relate to potential cultural bias in the items on the English version and not linguistic bias. Internal consistency analyses of both the auditory comprehension and expressive communication subscales compared values achieved when sampling European-American (n=63) versus Latino (n=387) preschoolers and found values for both higher than .90. Validity was established through finding acceptable, moderate correlations when comparing PLS-4 scores with performance on the PPVT-III. They also found no statistically significant difference in the correlations between the PLS-4 and the PPVT-III between the European-American and Latino subgroups, indicating that the two tests generally ranked children's performance the same across the two groups. However, the cut-off scores for delay differ between the two instruments and the PPVT-III identified more Latinos as language delayed than the PLS-4 (Qi & Marley, 2011). In 2009 Qi and Marley in a similar study using a differential item functioning analysis of the PLS-4 with English-speaking Hispanic and European low-income children found few items functioning differentially. Between these two studies they concluded there is no strong evidence of cultural bias in the English PLS-4. However, more work could be done to investigate linguistic bias when children are bilingual Spanish-English speakers.

The Spanish version of the PLS-4 is not simply a translation of the English version. The Spanish PLS-4 standardization sample included 1,188 Spanish-speaking children from all regions of the U.S. representing a variety of Spanish dialects. There are lists of alternate vocabulary on the record form that are acceptable responses given regional differences in the Spanish spoken in the U.S. The standardization sample also included both simultaneous and sequential bilinguals making it applicable to a broad range of young Spanish speakers in the U.S. Overall the authors created a test in Spanish that incorporates linguistic features that are unique

to Spanish and to the language development of young Spanish speakers in the U.S. (Zimmerman et al., 2002). The Spanish version also has strong psychometric properties with reliability coefficients range from.91 to .92 for 4- to 5-year olds. However, the Spanish PLS-4 has not been subject to the same level of research as that completed on the English PLS-4.

Although there currently are no publically available measures of early language development in English and Spanish that have addressed all current recommendations regarding bilingual measurement such as addressing the functional, linguistic, metric, and cultural equivalence between the English and Spanish versions (Peña, 2007), the PLS-4 is one of a few language measures available in English and Spanish for use with three- to five –year-old children that has at least demonstrated adequate reliability and validity with a Spanish-speaking bilingual/bicultural population. Therefore, the strong correlations between the BELA and the PLS-4 found both between the Spanish and English versions bode well for the utility of the BELA in the field and for the confidence with which it can be used.

The BELA was designed to be used for instructional planning, progress monitoring, and as a broad measure of language proficiency. Given that the instrument is currently available at no cost on-line, it is important that practitioners at least have access to pilot data so that they can be aware of the limitations of the instrument. The findings reported here suggest that some caution should be taken when using the BELA with young Spanish speakers. Questions remain about the reliability of the BELA, especially the Spanish version. The BELA shows some promise as a progress monitoring and instructional planning tool, but more research is needed. Replication of this study with a larger sample size is warranted to assess the validity and reliability of the BELA in a stronger analytic design in the hope of providing definitive recommendations regarding the appropriate uses of the BELA for practitioners.

REFERENCES

Alvarado, R., Ruef, M.L., & Schrank, F.A. (2005). *Woodcock-Muñoz Language Survey-Revised*. Itasca, IL: Riverside Publishing.

Anderson, R. (2002). Practical assessment strategies with Hispanic students. In A. E. Brice (Ed.), *The Hispanic child* (pp. 143–184). Boston: Allyn & Bacon.

Ballard, W. S., Tighe, P. L., & Dalton, E. F. (1991). Pre-*IDEA language proficiency test.* Brea, CA: Ballard & Tighe.

Barrueco, S., Lopez, M., Ong, C., & Lozano, P. (2012). Assessing Spanish-English bilingual preschoolers. Baltimore: Brookes Publishing.

Bedore, L. M., Peña, E. D., García, M., Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference? *Language, Speech, and Hearing Services in the Schools*, 36, 188-200.

Boyce, L.K., Gillam, S.L., Innocenti, M.S., Cook, G.A., & Ortiz, E. (in press). An examination of language input and vocabulary development of young Latino dual language learners living in poverty. *First Language.*

DeVellis, R. (2003). *Scale development: Theory and application*. Thousand Oaks, CA: Sage.

Duncan, S. E*., & De Avila, E. A. (2000). PreLAS 2000. Monterey,CA: CTB/McGraw-Hill.*

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test—Third Edition*. Circle   Pines,   MN: American Guidance Service.

Esquinca, A., Yaden, D., & Rueda, R. (2005). In J. Cohen, K.T. McAlister, K. Rolstad, J. Macswan (Eds.) Current language proficiency tests and their implications for preschool English language learners. *Proceedings of the 4th International Symposium on Bilingualism* (pp. 674-680). Somerville, MA: Cascadilla press.

Hammer, C., Miccio, A.W., & Rodriguez, B.L. (2004). Bilingual language acquisition and the child socialization process. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp.). Baltimore: Brookes.

Heise-Bagorria, C. (2008, October). The Bilingual Early Language Assessment. In Sharon

Yandian (Chair), *National Head Start Dual Language Institute*. Symposium conducted at the meeting of the National Head Start Association, Washington D.C.

Kohnert, K. & Bates, E. (2002). Balancing bilinguals II: Lexical comprehension and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research*, 45, 347-359.

Kohnert, K. Bates, E. & Hernandez, A. (1999). Balancing bilinguals: lexical-semantic production and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research*, 42, 1400-1413

Kovelman, I., Baker, S.A., Petitto, L. (2008). Age of first language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, 11(2), 203-223.

Landry, S., Assel, M., Gunewig, S., Swank, P. (2004). *Circle phonological awareness language and literacy screener*. Houston, TX: Ridgeways.

Lonigan, C. J. (2003). *Technical report on the development of the NCLD Spanish-language Get Ready to Read! screening tool*. Retrieved from http://www.getreadytoread.org.

Lonigan, C. J., & Wilson, S. (2008). *Report on the Revised Get Ready to Read! Screening Tool: Psychometrics and Normative Information*. New York: National Center on Learning Disabilities.

MacSwan, J., Rolstad, K., & Glass, G. V. (2002). Do some school-age children have no language? Some problems of construct validity in the Pre-Las Español. *Bilingual Research Journal*, 26 (2), 395-420.

National Association for the Education of Young Children (2005). *Screening and assessment of young English-language learners: Supplement to the NAEYC and NAECS/SDE joint position statement on early childhood curriculum, assessment, and program evaluation*. National Association for the Education of Young Children. Available at: http://www.naeyc.org/files/naeyc/file/positions/ELL_Supplement_Shorter_Version.pdf

Nunnally, J. & Bernstein, I. (1994). *Psychometric Theory*. Columbus, OH: McGraw-Hill.

Páez, M.M., Tabors, P.O., & López, L.M. (2007). Dual language and literacy development of Spanish-speaking preschool children. *Journal of Applied Developmental Psychology*, 28, 85-102.

Paradis, J., Genesee, F. & Crago, M. B. (2010). *Dual language development and disorders: A handbook on bilingualism and second language learning.* Baltimore, MD: Brookes.

Pearson, B., Fernández, S.C., & Ollers, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, 43, 93-120.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child Development*, 78(4), 1255-1264.

Peña, E. D., Bedore, L., & Rapazzo, C. (2003). Comparison of Spanish, English, and bilingual children's performance across semantic tasks. *Language, Speech, and Hearing Services in Schools*, 34, 5-16

Peña, E. D., Gillam, R. B., Bedore, L.M., Bohman, T. M. (2011). Risk for performance on a language screening measure for bilingual preschoolers and kindergarteners. *American Journal of Speech-Language Pathology*, 2, 302-314.

Peña, E. D., & Halle, T. G. (2011). Assessing Preschool Dual Language Learners: Traveling a Multiforked Road, *Child Development Perspectives*, 5 (1), 28-32.

Peña, E. & Kester, E. S. (2004). Semantic development in Spanish-English bilinguals: Theory, assessment, and intervention. In B. A. Goldstein (Ed.), *Bilingual language development and disorders in Spanish-English speakers* (pp. 187-212). Baltimore: Brookes.

Qi, C.H., & Marley, S.C. (2009). Differential item functioning analysis of the Preschool Language Scale-4 between English-speaking Hispanic and European American children from low-income families. *Topics in Early Childhood Special Education, 29*, 171-180.

Qi, C.H., & Marley, S.C. (2011). Validity study of the Preschool Language Scale-4 with English-speaking Hispanic and European American children in Head Start programs. *Topics in Early Childhood Special Education, 31*, 89-98.

Sheskin, D.J. (2004). *Handbook of parametric and non-parametric statistics* (3rd ed.). Boca Raton, FL: Chapman Hall.

Standards for Educational and Psychological Testing (1999). Joint Committee of the American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, AERA: Washington: DC.

Tabors. P. O., & Heise-Baigorria, C. (2004). *Bilingual Early Language Assessment*. Retrieved from
www.cpsd.us/BELA/

Tabors, P. O., Páez, M. M., & López, L. M. (2003). Dual language abilities of bilingual four-year olds: Initial findings from the early childhood study of language and literacy development of Spanish-speaking children. *NABE Journal of Research and Practice, 1,* 70-91.

U.S. Census Bureau. (2010).*Overview of race and Hispanic origin: 2010*. Retrieved from
http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. (2008). *Dual Language Learning: What does it take?* Retrieved from
http://eclkc.ohs.acf.hhs.gov/hslc/ecdh/eecd/Individualization/Learning%20in%20Two%20Languages/DLA NA_final_2009%5B1%5D.pdf

Zimmerman, I.L., Steiner, V., & Pond, R. E. (2002) *Preschool Language Scale -4*. Birmingham, AL: Psych Corp.

Appendix A

Questionnaire for the Family Home Language and Education

Date____/____/____

Child's Name_____
Child's Date of Birth____/____/____ Child's age:_____
Name of the person completing this form:_____
Your relationship to the child:_____

1. (a) At what age did your child begin to attend Head Start? _____
(b) In what type of educational programs has you child participated? (for example preschool, childcare) Were they bilingual programs or English-only settings?

2. (a) Does your child speak Spanish? YES   NO
(b) If so, at what age did your child begin to speak Spanish? _____
(c) Does he/she speak Spanish often? YES   NO
(d) With whom does your child speak Spanish?

3. (a) Does your child speak English? YES   NO
(b) If so, at what age did your child begin to speak English? _____
(c) Does he/she speak English often? YES   NO
(d) With whom does your child speak English?

4. (a) In what languages do you speak at home?_____
(b) Please list the people that live in your home, and the languages that each person speaks. (For example: grandmother-Spanish, older brother-English and Spanish, etc.)

5. Does your child have any medical or cognitive conditions that could limit his/her participation in academic activities? YES   NO

6. In comparison with other children of the same age, do you feel that your child has any problems in speaking? YES   NO

7. Do you have any concerns regarding your child's development, behavior, language or learning level? YES   NO

How many years of schooling has the child's mother completed?
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21+
In what country?_____

How many years of schooling has the child's father completed?
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21+
In what country?_____

Cuestionario Para La Familia Idioma Y Educación del Hogar

Fecha___/___/___

Nombre del niño/a_____

Fecha de nacimiento del niño/a ___/___/____ Edad del niño/a:_____

Nombre del informante_____

Relación con el niño/a: _____

1. (a) ¿A qué edad empezó su hijo/a a asistir a Head Start? _____

(b) ¿Cuales programas educativo ha participado su hijo/a? (Por ejemplo clases prescolar o cuidado de niños) Eran estos programas bilingües o solo inglés?

2. (a) ¿Su hijo/a habla español? SI   NO

(b) ¿Si habla español, a qué edad lo empezó a hablar? _____

(c) ¿Lo habla con mucha frecuencia? SI   NO

(d) ¿Lo habla bien, en comparación a otros niños de su edad? SI   NO

(e) ¿Con quién(es) habla el español?

3. (a) ¿Su hijo/a habla inglés? SI   NO

(b) ¿Si habla inglés, a qué edad lo empezó a hablar? _____

(c) ¿Lo habla con mucha frecuencia? SI   NO

(d) ¿Lo habla bien, en comparación a otros niños de su edad? SI   NO

(e) ¿Con quién(es) habla el inglés?

4. (a) ¿En qué idioma(s) hablan en casa?_____

(b) Favor de nombrar las personas que viven en casa, y los idiomas que hablan (por ejemplo: abuela- español, hermano mayor- inglés y español etc.):

5. ¿Su hijo/a tiene alguna debilidad médico o cognitivo que podría limitar su participación en actividades escolares? SI   NO

6. ¿En comparación a otros niños de la misma edad, piensa usted que su hijo/a tiene problemas en usar el lenguaje para expresarse o hacerse entender? SI   NO

7. ¿Usted tiene alguna preocupación actual acerca del desarrollo, conducta, lenguaje o aprendizaje de su hijo/a? SI   NO

¿Cuántos años escolares recibió la mamá del niño?
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21+
¿En qué país?_____

¿Cuántos años escolares recibió el papá del niño?
1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 18 19 20 21+
¿En qué país?_____