# Validity of the first two subtests of the preschool language assessment scale as a language screener for Spanish-speaking preschool children

Stefano Rainelli [d],*, Rebecca J. Bulotsky-Shearer [a], Veronica A. Fernandez [a], Daryl B. Greenfield [a], Michael López [b]

[a] University of Miami, United States
[b] Abt Associates, United States
[d] Miami-Dade County Public Schools, United States

### ABSTRACT

Large-scale early childhood studies use the first two subscales of the Preschool Language Assessment Scale, "Simon Says" and "Art Show" (PreLAS2000; Duncan & De Avila, 1998) to guide decisions about the most appropriate language (or languages) researchers should use when directly assessing the academic skills of dual language learner (DLL) children. Large-scale studies use a cut-score derived from a total score on the two PreLAS subscales in English and/or Spanish in combination with parent or teacher reports of children's language abilities, to route children into the most appropriate language of assessment. However, limited research exists to support the use of these cut-scores as part of a language routing procedure with Spanish-speaking DLL preschool children from low-income backgrounds. The current study examined the validity of the two English PreLAS subscale scores for a local sample of children enrolled in Head Start ($N = 872$) and Hispanic children from the national FACES 2006 sample ($N = 935$). Rasch and DIF analyses supported the invariance of item difficulty values across the three- and four-year-old age groups in the overall sample. For a subsample of Spanish-speaking DLL children, receiver operator characteristic (ROC) analyses identified the most appropriate cut-scores on the PreLAS screener for both age groups. Findings provided evidence to support the validity of the use of the English PreLAS language screener score as part of a more comprehensive language routing procedure. Implications for policy, practice, and measurement development are discussed.

© 2016 Elsevier Inc. All rights reserved.

Decades of research findings document a pernicious, national achievement gap. At kindergarten entry, children from low-income households perform on average one standard deviation below their socioeconomically more advantaged peers in language, literacy, and mathematics (Benner & Crosnoe, 2011; Hair, Halle, Terry-Humen, Lavelle, & Calkins, 2006; Rouse, Brooks-Gunn, & McLanahan, 2005). Children from ethnic and linguistic minority families face additional challenges as they transition to kindergarten (Espinosa, 2007; Lareau, 2000). In particular, Hispanic children, the fastest growing ethnic minority group in the United States, are statistically more likely to come from non-English-speaking immigrant families living in poverty (DeNavas-Walt, Proctor, & Smith, 2009), to perform poorly in U.S. schools in reading and other academic areas upon kindergarten entry (Rathbun, West, & Germino-Hausken, 2004), and to be retained and drop out of school (Gersten & Woodward, 1994).

Early childhood educational programs serving low-income children from diverse backgrounds have the opportunity to provide high-quality educational services during a time in early development when learning is most malleable and critical to children's future academic success (Duncan, Ludwig, & Magnuson, 2007; Heckman & Masterov, 2007). Head Start, our nation's largest and most comprehensive early education program supporting the school readiness needs of children from low-income families, serves an increasing number of three- and four-year-old Hispanic children, who speak Spanish at home and are considered Dual Language Learners (DLLs). The term DLL is used to refer to young children who come from a home in which a language other than English is primarily spoken (Barrueco, Lopez, Ong, & Lozano, 2012; Espinosa & García, 2012). Currently, 30% of children enrolled in Head Start programs nationally are considered DLLs and the majority of these DLL children (85%) are from families where Spanish is the primary home language (Castro, 2011; U.S. Department

---

* Corresponding author.
*E-mail address:* stefano.rainelli@gmail.com (S. Rainelli).

of Health and Human Services, Administration on Children and Families, 2014).

As increasing numbers of Spanish-speaking three- and four-year-old DLL children enroll in early childhood education programs, a critical policy issue is how best to support and to assess their early learning (Center for Early Care and Education Research-Dual Language Learners [CECER-DLL], 2011; Chang et al., 2007). Federally- and state-funded early childhood programs, such as Head Start, are also facing increased accountability mandates to provide high-quality instruction to all three- and four-year-old children and to document children's academic progress regardless of language background (Espinosa, 2007; Meisels, 2007). However, there are gaps in the field's understanding of DLL children's language development and, subsequently, their academic skill development, as well as several methodological challenges in conducting research with this population (CECER-DLL, 2011).

The ideal approach for early childhood researchers and practitioners is to use a dual-language assessment approach or a conceptual scoring approach, to assess children's academic skill development in both English and Spanish (Barrueco et al., 2012; CECER-DLL, 2011). However, few commercially available, psychometrically sound measurement tools currently are available to assess and to track preschool children's academic progress in both English and Spanish. In addition, due to limited resources, it is not always possible in large-scale studies to assess children using a dual-language approach or conceptually scored measure. Therefore, researchers recommend the use of a comprehensive multi-step language routing procedure to help correctly route children into the most appropriate language(s) of assessment. This approach combines parent- and teacher-report of children's primary language, as well as the use of a minimum proficiency cut-score on a brief screener of children's language proficiency skills (Barrueco et al., 2012). Children receive a brief language screener and their resultant score routes them into a series of English-language (and/or Spanish-language) direct academic assessments (Barrueco et al., 2012). For example, large-scale early childhood studies such as the Early Childhood Longitudinal Study, Kindergarten Cohort (ELCS-K), the Early Childhood Longitudinal Study, Birth Cohort (ELCS-B), and the Head Start FACES 2006 and 2009 cohorts, have relied on this screening approach (CECER-DLL; Aikens, Atkins-Burnett, & Bandel, 2012). However, a limitation in this approach is that few validated language-screening tools are available to incorporate into a routing procedure for three-year-old children (Espinosa, 2005). As three-year-old DLL children increasingly enroll in early childhood education programs nationally, validation of screening tools for use in large-scale studies for this age group is needed.

## 1. Use of the PreLAS as a language screener

The Preschool Language Assessment Scale (PreLAS2000; Duncan & De Avila, 1998) is a commercially available, norm-referenced language proficiency battery for use with children aged four to six years. It was developed to measure expressive and receptive language skills, including syntax, vocabulary, and command of grammatical phrases. The scale is available in English and Spanish. The full measure, used to derive a proficiency classification, is comprised of five subtests. The full battery can be administered in English or Spanish (or both) and has been standardized for administration to children between the ages of four to six years. The purpose of the scores from the full assessment are to help determine whether there are delays in certain areas of language development. This full assessment provides normative-referenced scores for expressive and receptive skills and can facilitate a child's

access to early intervention services should there be an identified delay (Duncan & De Avila, 1998).

As noted above, large-scale early childhood studies have adapted the PreLAS2000 measure to be used for screening purposes when resources for completing a full battery in both Spanish and English are limited, or when equivalent Spanish and English assessments are not available (Barrueco et al., 2012; CECER-DLL, 2011). These studies have used an abbreviated version of the PreLAS2000. Beginning with the ECLS-K, 1998-99, the first three subscales, "Simon Says" (which measures receptive language skills), the "Art Show" (that measures expressive language skills), and "Let's Tell Stories" (that measures natural speech) were used to route four-year-olds into an English language assessment battery (National Center for Education Statistics [NCES], 2002; Rock & Pollack, 2002). The researchers selected the PreLAS2000 subtests because of their brevity, ease of administration and scoring, sound predictive validity and because of its similarities in format and administration procedures to the other measures used in the study (NCES, 2002). In consultation with the developer of the original PreLAS2000 (Dr. De Avila), the first three subscales of the measure were combined to develop the Oral Language Development Scale (OLDS; NCES, 2002) in English.

Subsequent large-scale studies of preschool children have used just the first two ("Simon Says" and "Art Show") subtests as an English language screener to route both three- and four-year-old children into the most appropriate language of assessment (Head Start FACES, 2006, 2009). For example, the FACES 2003 and FACES 2006 studies used the first two subtests of the PreLAS2000 (which from this point forward will be referred to as "PreLAS") in combination with the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007) to route children into either the English- or Spanish-language assessment battery. If a child, from a Spanish-speaking background, answered five consecutive items incorrectly on each subscale of the English PreLAS screener and made five to eight errors on the PPVT, the assessor discontinued the screener and routed the child out of the English assessments. Researchers generalized this "stoppage rule" from the PreLAS2000 manual, which states that "if a [child] misses five consecutive items in two consecutive parts," the test should be stopped (Duncan & De Avila, p.14, 1998). However, this published rule applies for administration of the PreLAS2000 assessment battery as a whole, and not for the two subtests as a screener. In addition, this stoppage rule was not validated for use with three-year-old children. The stoppage rule applies for administration of the full battery only for children ages four years and older, as represented in the normative sample. Few studies have subsequently examined whether this stoppage rule (5 consecutive items incorrect) is empirically valid for both three- and four-year-old DLL children.

Further, the use of the first two subtests of the PreLAS2000 as a language screener differs from the original purpose of the full measure, which is as a clinical assessment of a child's English- or Spanish-language skills. Accordingly, the Standards for Educational and Psychological Testing indicate that validity evidence is needed to support the use of a measure outside its intended purpose (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014). Only one study to date has examined the psychometric properties of the cut-score derived from the two subtests of the PreLAS2000 for use as part of a routing procedure. In the First 5 Los Angeles Universal Preschool Childhood Outcomes Study (UPCOS; Vogel et al., 2008), an empirically derived cut-score was used from the total score on the first two subscales of the PreLAS2000. Receiver operator characteristic curve analysis (ROC) identified a cut-score of 15 out of 20 items from the "Simon Says" and "Art Show" subtests in that sample (Fawcett, 2006; Murphy et al., 1988). In combination with infor-