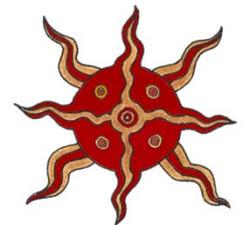# Assessing Dual Language Learners: Critically Examining Our Measures

Sandra Barrueco, Ph.D.
The Catholic University of America

Michael L. López, Ph.D.
National Center for Latino Child & Family Research

Present Times

- Increasingly diverse population

- Growth of early intervention programs

- Focus on inclusion and accountability

The Complete Assessment of DLLs

- <u>Approaches</u>
  - Portfolios and Work Sampling
  - Formative and Dynamic Assessments
  - Family Interviews and Conferencing
  - Norm- and Criterion-Referenced Standardized Assessments
  - Medical and Specialty Assessments

- <u>Goals (e.g., Shepard et al, 1998; Epstein et al, 2004):</u>
  - Instructional Planning
  - Clinical
  - Research
  - Policy

Standardized Assessments with DLLs

- Growth of multilingual measures

- Frequent errors in field
    - Conducting one's own translation
    - Family member administer assessment
    - Selecting other language measure out of convenience, familiarity or because they appear adequate in English.

- Potential Consequences for inappropriate selection and use
    - Misdiagnosis
    - Program Defunding
    - Invalid research findings on DLLS
    - Inappropriate policy decisions

Selecting Appropriate Measures that are culturally, linguistically, and psychometrically appropriate with DLLs

- A basic translation of a measure from English to Spanish is neither adequate nor sufficient

- Spanish measure may be published and readily available to the field but it is no guarantee of
  - solid psychometric properties, or
  - that the necessary array of cultural and linguistic measurement development approaches were undertaken.

Selecting Appropriate Measures that are culturally, linguistically, and psychometrically appropriate with DLLs

- <u>Selecting the Intensity</u>
  - Screener vs full evaluation
  - Sensitivity and Specificity

- <u>Selecting the Domain(s)</u>
  - What is the measure intended to measure?

- <u>Logistical</u>
  - Who can administer it and for what purpose?
  - How much training and time is needed?

Selecting Appropriate Measures that are culturally, linguistically, and psychometrically appropriate with DLLs

- <u>Selecting the Language(s)</u>
  - English Language Fluency and Identification for educational services
  - Identifying Dominant vs. Non-dominant Language and relative levels of Fluency
  - *Assessment approaches in one language*
    - English
    - Spanish
    - Dominant Language
  - *Dual Language Approaches*
    - Language Screening/ Routing
    - Asymmetrical Dual Language Assessment
    - Symmetrical Dual Language Assessment

Selecting Appropriate Measures that are culturally, linguistically, and psychometrically appropriate with DLLs

Development and Standardization

- How were DLL/Spanish-speaking children included when developing the measure? How many of these children were included?

- What is the demographic composition of the normative samples across ethnicity, socioeconomic status, language, and region, among others?

- How well does the norming sample match with the population under study?

Selecting Appropriate Measures that are culturally, linguistically, and psychometrically appropriate with DLLs

Psychometric Features

- How strong are its psychometric properties (across the different types of reliability and validity)?

- How do these compare or differ across the English and Spanish versions?

Psychometric Features

- *Reliability* refers to how precise or trustworthy a test score is in capturing the skills, attitudes or abilities it is purported to measure.
  - Test-retest reliability
  - Alternate forms reliability
  - Internal consistency
  - Inter-rater or inter-scorer agreement

Psychometric Features

- *Validity* refers to the degree to which all accumulated evidence supports the interpretation and use of test scores for a particular purpose
  - Face validity
  - Content validity
  - Construct validity (Internal and External)
  - Criterion validity (including Concurrent and Predictive)

Cultural and Linguistic Considerations

- How strong are its cultural and linguistic measurement properties for use with Spanish-speakers or DLLs?

Cultural and Linguistic Considerations

Content equivalence

Are constructs & operational definitions pertinent for cultural group?

- Literature review
- Expert panel
- Observations
- Interviews
- Focus Groups

Cultural and Linguistic Considerations

Semantic Equivalence

- How was the measure translated and subsequently adjusted?
- Upon translation, did the items conform to *semantic* or *content* equivalents?
  - Translation/Back-Translation methods
  - Substitution: culturally concordant items or construct equivalents
  - Field Tests & Statistical analysis of item difficulty, order, etc.
  - Feedback from cultural informants: interviews, focus groups, panel comprised of community and experts

Cultural and Linguistic Considerations

- <u>Structural Consistency across English and Spanish versions:</u>
  - How similar are the items, length and format of the measure across the languages?

- <u>Standardization:</u>
  - Was the measure standardized with Spanish-speaking children or DLL children?
  - How many?
  - Was the measure statistically equated? How?

Cultural and Linguistic Considerations

- <u>Technical Equivalence in Reliability:</u>
  - How similar is the reliability of the Spanish and English versions?

- <u>Technical Equivalence in Validity:</u>
  - How similar is the validity of the Spanish and English versions?

Assessing Young Latino Children Within and Across Two Languages: Approaches and Measures
(Barrueco, López, Ong, & Lozano, in press)

- Critical examination of the psychometric, linguistic, and cultural properties of the measures currently available for use with preschool-aged language minority children.

- Intent: Provide independent evaluations of measures that clinicians, teachers, or researchers are considering utilizing.
  - The ultimate selection by these individuals is guided by the purpose for the assessment, along with the characteristics they need in a measure.

- Focused on direct language and literacy measures for use with Spanish-speaking preschoolers

Selection Process and Inclusion Criteria

- **Step 1:**
  - Initial identification of > 1000 measures relating to early childhood, language, or literacy
  - Multiple sources: MMY, publishing companies, articles, reports, and internet searches.
- **Step 2:**
  - Identified measures with:
    - target age range including 3-5 year olds,
    - ample coverage of language and/or literacy domains,
    - publication of English and Spanish forms, and
    - direct child assessment.

Selection Process and Inclusion Criteria

Step 3: Final Selected Measures

- The last step consisted of identifying the final measures for review. The selected measures possessed the following characteristics:

  a) availability in English and Spanish for similar age groups,

  b) included a direct child assessment component,

  c) useful with the general population, rather than solely for children with disorders,

  d) had been standardized or extensively used in last decade, *and*

  e) could be administered by teachers, examiners, and/or researchers.

19 Assessments Reviewed

- Battelle Developmental Inventory (BDI-2)
- Bilingual Vocabulary Assessment Measure (BVAM)
- Boehm–3 Preschool
- Brigance Preschool Screen-II
- CELF-Preschool-2
- Compton: Preschoolers Screening Evaluation
- Developmental Indicators for the Assessment of Learning (DIAL-4)
- Early Literacy Skills Assessment (ELSA)
- Early Screening Inventory-Revised
- Expressive One-Word Picture Vocabulary Test (EOWPVT)

19 Assessments Reviewed

- FirstStep: Screening Test for Evaluating Preschoolers
- Get Ready to Read! Screening Tool
- Merrill-Palmer-Revised Scales of Development (Merrill-Palmer-R)
- Peabody Picture Vocabulary Test-4 /Test de Vocabulario Imágenes Peabody (PPVT/TVIP)
- Pre-LAS 2000
- Preschool Language Scale-5 (PLS-5)
- Receptive One-Word Picture Vocabulary Test (ROWPVT)
- Woodcock-Muñoz Language Survey-Revised
- Young children's Achievement Test (YCAT)

Discussion

- Consider assessment goals and question at hand

- One needs to examine and consider reliability, validity, and linguistic and cultural equivalence across both English and Spanish versions

- More measures being developed adequately for use with Spanish ELL preschoolers, but gaps still exist

Full Reference and Gratitude

Barrueco, S., López, M.L., Ong, C. & Lozano, P. (in press). *Assessing Young Latino Children Within and Across Two Languages: Approaches and Measures.*

Baltimore, MD: Brookes Publishing.

Thanks to CUA, First 5 Los Angeles, and Pew Charitable Trusts' Early Childhood Accountability Project.

Contact: barrueco@cua.edu