Research

■ ■ ■
■ ■ ■
■ ■ ■ Validity of the Spanish Preschool Language Scale-3 for Use With Bilingual Children

M. Adelaida Restrepo
Stacy W. Silverman
The University of Georgia, Athens

This study evaluated the validity of the Spanish edition of the Preschool Language Scale-3 (I. L. Zimmerman, V. G. Steiner, & R. E. Pond, 1993). As a preliminary step, the authors reviewed the test to determine whether it met psychometric criteria established by McCauley and Swisher (1984) and Hutchinson (1996). Examination of the test's psychometric characteristics revealed that the test met only 4 out of 10 criteria proposed by McCauley and Swisher and none of the additional criteria from Hutchinson. Problems were evident in the test's norming and in the lack of reliability and validity data. The authors then investigated whether, despite the test's psychometric shortcomings, it was useful for the assessment of Spanish/English-speaking children. Results revealed that the children studied performed approximately 1.5 *SD* below the mean. Moreover, the children's performance on the subtests did not reflect an even progression of item difficulty, indicating limited evidence of construct and content validity.

Key Words: validity, test, bilingual, language, children

T he present study assesses the validity of a current language measure, the Spanish edition of the Preschool Language Scale-3 (SPLS-3; Zimmerman, Steiner, & Pond, 1993), for use with preschool and early school-age children. Examination of the test's psychometric characteristics and results from a sample of Spanish-speaking children were used to evaluate its validity. The validity of the SPLS-3 for the identification of language disorders in Spanish-speaking children in the United States has not been established despite the growing consensus in speech-language pathology that available language assessments for Spanish-speaking and/or bilingual children are inadequate for determining the presence of a language disorder (e.g., Beaumont & Langdon, 1992; Erickson & Iglesias, 1986; Huang, Hopkins, & Nippold, 1997; Mattes & Omark, 1991; Taylor & Payne, 1983). Validity is of increasing importance; as the number of Hispanic children increases in schools, the need for reliable, valid measures for assessing this population becomes more urgent. Currently, 30% of the Hispanic population is younger than age 15, compared to 20% of the non-Hispanic white population (Hispanic Task Force, 1999). By the year 2015, the Hispanic population is expected to be the largest minority group in the United States (Pollard, 1999).

Assessing the Validity of Language Measures

A practical consequence of poor test validity is faulty identification of clients within the domain that the measure purports to test (e.g., for discussions, see Hutchinson, 1996; McCauley & Swisher; 1984). To assess the validity of a measure, Hutchinson described two types of evidence, logical and empirical. The assessment of logical evidence entails an analysis of the rationale and specifications of the test, whereas the assessment of empirical evidence entails determination of quantifiable relationships among scores on different parts of the test and in comparison with other tests (evidence that can be obtained by test developers or by independent researchers). Further, even if the logical and empirical evidence is strong, the test must also have good reliability and adequate norms to help the clinician determine its validity and its appropriateness for a specific population.

Several types of validity are described in the literature. *Content validity,* a type of logical evidence, is the determination that items are representative of a content domain and are relevant to the intent and model of the test. *Criterion-related validity,* a type of empirical validation, is evaluated by comparing a test's results with an established measure (criterion) that has been independently proven valid (e.g., for discussions, see Anastasi, 1988; McCauley & Swisher, 1984; Plante & Vance, 1994). *Construct*

validity, also a type of logical evidence, relates to the test's adherence to the specific theoretical bases of the domain being assessed (e.g., Anastasi, 1988; McCauley & Swisher, 1984). Construct validity can be measured by the extent to which a test's items and procedures are supported by the literature and the extent to which the test represents one or more specific theories.

Equally important to validity is whether tests are *reliable* in their measurement of performance (e.g., Anastasi, 1988; Hutchinson, 1996; Plante & Vance, 1984; Ventry & Schiavetti, 1986). Reliability alone does not validate a test; however, unreliability affects its validity. *Test-retest reliability*, for example, is a measure of the consistency of performance within individuals when a test is taken twice. *Split-half (or "odd-even") reliability*, which compares performance on half of the items to performance on the other half, is an assessment of consistency in item difficulty. *Inter-rater reliability* is a measure of the extent to which an examinee's performance varies with different examiners.

Besides issues of test validity and reliability, the method used to obtain normative samples is also a concern (Hutchinson, 1996). McFadden (1996) examined the samples selected in the development of normative data for 11 tests. She reported that only 6 of the 11 tests had normative samples that included participants who represented the full range of language abilities. The remaining 5 tests applied exclusionary criteria to homogenous samples. If, through exclusionary criteria, the children who would be expected to represent the low end of the distribution are not included, the subsequent normative data set will not accurately depict "average" performance (Hutchinson, 1996; McFadden, 1996). Rather, a child who performs at the 50th percentile on a test that is normed on a left-tail truncated sample actually has performed above the mean when compared with the full range of language abilities in the unrestricted population.

Spanish Language Assessment Measures

Examination of Spanish language tests for reliability and content, construct, and concurrent validity suggests caution in the use of these tests for measuring language performance. The first problem with available tests is the translation of tests from English to Spanish. Langdon (1992), for example, reported on 21 available Spanish language tests and found that several contained a majority of items that were translated directly from English. Such an approach automatically compromises content validity because immediate translations ignore linguistic differences across languages (Langdon, 1992; Mattes & Omark, 1991).

Another problem with the validity of Spanish language tests is that the extent to which these tests exhibit construct validity for the assessment of language skills is unknown. We can speculate, however, that if these tests have poor content validity, they will not adequately represent the constructs on which they were originally based. If, for example, one construct is that certain language behaviors increase in frequency of use or complexity with age, a test that is translated directly from English will have poor content validity and may not provide an age progression of

item difficulty, which is a developmental construct.

A third problem is that of presentation of reliability data. Langdon's (1992) review of Spanish language tests concluded that none of the 21 tests provided adequate reliability data (defined as test-retest, internal, and inter-rater reliability). Without this information, one can only speculate about the stability of children's test performance.

Adequate norming of samples in language tests for Spanish-speaking populations is also a significant problem (Erickson & Iglesias, 1986; Langdon, 1992). Langdon (1992) found that several of the tests she examined had limited norms as compared to those recommended in test development (100 individuals per age or language subgroup); others were translated and not normed, or provided such limited information about their norming sample that interpretation was difficult. In other cases, data regarding the normalcy of the sample, the children's language use, or their language levels were not discussed or presented in the manual (Langdon, 1992).

Evidence of validity in alternatives to Spanish normed-referenced assessments. At present, there are few validated measures of Spanish-speaking children's language abilities, which in turn affects our ability to obtain criterion-related validity. Recently, Anderson (1996) evaluated the concurrent validity of the Spanish Structured Photographic Expressive Language Test-Preschool (SSPELT-P; Werner & Kresheck, 1989b). She compared it with a non-standardized task of her own design and found that 4-year-old Spanish-speaking children, without exception, performed better on the set of grammatical targets in her task than on the SSPELT-P, even though the grammatical targets were the same across measures. She attributed this difference to the fact that her task allowed the children more opportunities to produce target structures than the SSPELT-P did, and that her tasks were more interactive, engaging the children more than the items on the SSPELT-P.

More recently, Restrepo (1998) used discriminant function analyses to determine the extent to which performance on several measures predicted Spanish-speaking children's status in terms of specific language impairment versus normal language. The measures were the Spanish Structured Photographic Expressive Language Test-II (SSPELT-II; Werner & Kresheck, 1989a), parental concerns, family history of speech and language problems, number of grammatical errors per terminable unit (T-unit), mean length of T-unit, novel vocabulary learning, novel grammatical rule generalization, and the Developmental Assessment of Spanish Grammar (Toronto, 1976). The SSPELT-II, used alone or in combination with other measures, did not discriminate well between children with and without language impairment. It discriminated with a sensitivity of 65.22% (percentage of children with language disorders identified as such) and a specificity of 91.30% (percentage of children with normal language identified as such), indicating an average accuracy of 78.26%, which translates to a misidentification rate of 21.74%.

Restrepo's (1998) results indicated that no individual measure alone provided good discrimination (80% or greater for total discrimination; Plante & Vance, 1994). However, when used in combination, two measures

(number of grammatical errors per T-unit and parental concerns) and four measures (number of grammatical errors per T-unit, parental concerns, family history of speech/language problems, and mean length of T-unit) of the eight measures discriminated with greater than 90% accuracy. Based on these and other results (Gutierrez-Clellen, Restrepo, Bedore, Peña, & Anderson, 2000), the use of spontaneous language sample analyses as the primary means of assessment for this population, combined with parental report, are the statistically supported standards for the identification of Spanish-speaking children with language disorders. Also, they are well-accepted methods for assessing English-speaking children of different ages and backgrounds (e.g., Dale, 1991; Dunn, Flax, Sliwinski, & Aram, 1996; Gavin, Klee, & Membrino, 1993; Scott & Windsor, 2000). However, these procedures are time consuming, and the use of norm-referenced measures is required for qualification of children in some school districts. Thus, it seems that more efficient, standardized measures would be preferred by clinicians.

Further research is therefore needed on the validity and efficiency of norm-referenced measures for identifying language disorders in Spanish-speaking children. With this in mind, we selected for evaluation one norm-referenced test, the SPLS-3, which is used to assess the language abilities of Spanish-speaking children. This test is a translation of the English version, Preschool Language Scale–3 (Zimmerman, Steiner, & Pond, 1993), which is designed for the assessment of preschool and early school-age children. The SPLS-3 uses the same booklet and norms as the English version, although the manual and test forms are different. We first examined how the SPLS-3 meets the psychometric criteria for test adequacy provided by McCauley and Swisher (1984) and, more recently, Hutchinson (1996). We then examined the validity of the SPLS-3 using a local sample of bilingual Spanish/English-speaking children.

Method

Procedures

To evaluate the extent to which the SPLS-3 meets the psychometric criteria for test adequacy, using the criteria described below, both investigators in the current study independently reviewed the SPLS-3 manual. Following independent review, the investigators conferred and arrived at a consensus about each criterion addressed in the test manual. The investigators determined whether each criterion was met, not met, or marginally met (addressed, but with some problems). The criteria used to analyze the psychometric value of the test included the 10 criteria from McCauley and Swisher's (1984) work and 5 supplemental criteria from Hutchinson (1996) that did not duplicate McCauley and Swisher's criteria. Table 1 provides the list of criteria used and subsequent results.

Based on the results of this review, we posed further questions regarding the validity of the SPLS-3:

* To what extent are the SPLS-3 norms representative of a local sample of bilingual Spanish/English-speaking children?

TABLE 1. Performance of the Spanish Preschool Language Scale on psychometric criteria.

Criteria	Results
McCauley & Swisher	
1. Description of normative sample	Marginal
2. Sample size	No
3. Item analysis	Yes
4. Means and standard deviations	Marginal
5. Concurrent validity	No
6. Predictive validity	No
7. Test-retest reliability	No
8. Inter-examiner reliability	No
9. Description of test procedures	Yes
10. Description of tester qualifications	No
Hutchinson	
11. Purpose of the test explicitly stated?	No
12. Construct or model explicitly defined?	No
13. Supportable rationale for test content?	No
14. Sample behavior at the extremes?	No
15. Norms represent performance at the extremes?	No

* To what extent does the SPLS-3 have construct and content validity, based on item analysis data from a local sample of bilingual Spanish/English-speaking children?

* To what extent does the SPLS-3 have criterion validity, and specifically, concurrent validity, when comparing an independent sample of children's performance on the SPLS-3 with their performance on two descriptive Spanish language measures, a spontaneous language analysis, and a parent interview?

Specifically, it was hypothesized that Auditory Comprehension subtest scores of the SPLS-3 would correlate significantly with a criterion-referenced receptive (CRR) measure, and that the Expressive Comprehension subtest scores of the SPLS-3 would correlate significantly with mean length of terminable unit (MLTU) and the criterion-referenced expressive (CRE) measure. Furthermore, it was hypothesized that the Spanish Preschool Language Assessment Scale (PreLAS; Duncan & DeAvila, 1986), an assessment of Spanish language proficiency, and the total SPLS-3 scores would correlate significantly with each other.

Participants

Thirty-seven children participated in the study (18 boys and 19 girls), although, due to participant attrition, not all of the children participated in all of the measures described, as will be indicated in the results. The ages of the children ranged from 4;4 to 6;6 (years;months), with a mean age of 5;6 ($SD = .61$). All of the children were bilingual speakers of Spanish and English who attended English-only instruction classrooms in Georgia and who were enrolled at the time of testing in pre-kindergarten and kindergarten programs. The children in the current study did

not receive English as a second language (ESL) support because such programs are only mandated beginning in the first grade in Georgia. In addition, the children did not receive any formal instruction in their first language.

Participant selection criteria. To participate in the study, the children were required to meet the following criteria:

- The children had to be bilingual. They had to score the same in Spanish and English or better in Spanish than in English in the PreLAS language proficiency test. Using this criterion ensured that the children's most proficient language was Spanish or that they had similar proficiency in both languages.

- The children had to achieve a score of at least 2 or better on the Spanish subtest of the PreLAS. A score of 2 indicates that the child has some fluency in the language but makes errors in some aspects of the language. This score was judged to be an appropriate cutoff score because it can include children who are fluent in Spanish but have language disorders (which would affect the score). This measure is used in schools throughout the United States for bilingual program placement decisions. It helped to ensure that each child spoke enough Spanish to complete the SPLS-3.

- The children were not previously identified as having a language disorder; however, in order to maintain a range of language skills, children who were at risk for language disorders per teacher report were not eliminated from the study (Hutchinson, 1996; McFadden, 1996).

- The teachers must have identified the children as primarily Spanish speaking. Our recruitment procedure provided additional assurance that the children spoke relatively fluent Spanish. Teachers were asked to refer Spanish-speaking children who came from Spanish-speaking homes.

- The children each had to pass a hearing screening at 20 dB in the following frequencies: 500, 1000, 2000, and 4000 Hz.

Procedures

Participants were seen by research assistants (RAs) in a quiet room in their schools. For the Spanish testing, the RAs were fluent in Spanish and English and spoke only Spanish to the children during the assessment. For the English testing, the RAs spoke only English during the English assessment of the child, although they had some knowledge of Spanish. In order to maintain a consistent language context for the children, no RA tested the same child in English and Spanish. Testing was conducted on at least 2 days. For most children, the Spanish testing required 2 days and the English testing required 1 day. To protect against fatigue, the children were not seen for more than 1 hour at a time. One of the investigators or a graduate assistant supervised 20% of the sessions to ensure that the RAs were following the appropriate procedures. Furthermore, reliability data were obtained on 10.8% (4 of 37) of the children for the SPLS-3 and the parent interview.

Measures

Parent interview, Spanish language sample analyses, and criterion-referenced measures were selected instead of standardized measures because there is no validated standardized measure and there are inherent drawbacks to available tests for the bilingual Spanish-English population. Further, parent interview and language sample analyses were selected because both have been validated for the identification of primarily Spanish-speaking children with specific language impairment (Restrepo, 1998).

Participants were evaluated using the following measures given in random order to avoid an order effect: SPLS-3, CRR and CRE morphology measures, and MLTU (Hunt, 1965; see adaptation by Gutierrez-Clellen & Hoffstetter, 1994). MLTU was obtained from the spontaneous language samples in Spanish, which were audiotaped for transcription and analyses. The spontaneous language samples were obtained in three different formats for each child: spontaneous conversation with preselected topics, free play with preselected toys, and story retelling from one of two frog books by Mayer (1967, 1969).

Language samples. The language samples were transcribed by a native Spanish speaker and segmented into T-units. The transcriber checked each sample twice and then gave the sample to a second transcriber, who listened to the tape and made corrections. If there were disagreements in the transcription, they were corrected by consensus. A T-unit is defined as a main clause and any subordinate clauses (Hunt, 1965). Gutierrez-Clellen and Hoffstetter's (1994) adaptation to Spanish was used, in which all conjoined subjectless clauses are counted as separate T-units rather than one T-unit. Elliptical phrases and T-units with unintelligible words were not counted. Mean length in words was used because the number of words can be counted accurately even when the child code-switches into English. Calculating the number of morphemes across languages as the child code-switches within the sample is problematic because of the presence or absence of different morphological markers in each language (Gutierrez-Clellen et al., 2000).

Parent interview. Parents participated in an interview in their own homes that addressed their children's current language skills (adapted from Restrepo, 1998). Parents were interviewed by fluent Spanish speakers, and all of the interviews were conducted in Spanish. Parents were asked to determine whether their child had speech or language problems in comparison to other siblings or children the same age. This section contained 18 yes/no questions, and the maximum score for parental concerns was 17 (possible yes responses), as two questions asked the same information in opposite form (17, 18). See the Appendix for a translation of the questions used in the questionnaire.

SPLS-3. The SPLS-3 was administered according to the directions in the manual so that base and ceiling scores were obtained for each child.

CRR. The CRR measure, developed by the first author, consisted of 20 items in which the child pointed to one picture out of a set of four pictures after the RA read a sentence. Each item contained a sentence that was best described by one of four pictures. The items probed the

child's understanding of several grammatical morphemes: plurals, articles, past tense, gender, person and number agreement, negatives, and embedded sentences. This measure contained no base or ceiling and was scored online.

CRE. The CRE measure addressed the same grammatical morphemes as the CRR, but the child had to describe a picture, given a question or sentence probe, for a total of 45 items. For example, to elicit the plural form, the child was asked in Spanish, "Here there is a duck; here there are _____ (picture of two ducks)." The areas evaluated were determined to be acquired by preschool age in typically developing Spanish-speaking children. This measure was scored online because it required only short answers that could be judged quickly for correctness.

PreLAS. A language proficiency measure, PreLAS (Duncan & DeAvila, 1986), was administered to all participants in order to determine whether children had enough Spanish proficiency to take the SPLS-3 and were more proficient in Spanish or equally proficient in Spanish and English to justify a Spanish assessment. The measure has several subtests: story retelling, sentence imitation, sentence completion, following directions, and pointing to pictures. The measure assigns children to language levels ranging from 1 to 5, with 1 indicating limited language proficiency and 5 indicating full language proficiency.

Reliability

For the SPLS-3 and parent interviews measures, 4 families and 4 children (10.8%) were scored online independently by an RA and one of the investigators. Results indicated that the point-to-point reliability for the SPLS-3 was 94.7%, and for the parent interview was 98.1%. For the language samples, all tapes were checked twice by one transcriber, and then once by a second transcriber. All disagreements were resolved by consensus.

Results

Psychometric Criteria

The results of the independent analyses are summarized in Table 1 and explained in more depth below.

1. Do the authors provide a description of the normative sample? The investigators determined by consensus that the test's description of its normative sample was marginally adequate. The first problem identified was that the test authors had an experimental Hispanic group of 181 children, which they do not call a normative sample. Instead, the authors refer test administrators to the PLS-3 (English version) to obtain standard scores. This latter normative group, on which the PLS-3 is based, is described as having Hispanic representation with no further description of that subgroup of participants, and their language background is not provided. Although there is a limited description of the Hispanic experimental group, there is no description of this group's language proficiency and normalcy.

2. Is the normative sample adequate, including 100 children per age group? The investigators independently agreed that this criterion was not met. For the Hispanic

experimental group results, there were as few as 20 children per age group, for a total of 181 children. As discussed in Criterion 1, the authors should at least provide language proficiency norms, and the norms should be tested to determine whether there are dialectal differences. If so, separate norms should be provided. If the English norms of the PLS-3 are used, as recommended by the developers, they will not be appropriate for Hispanic children, given that they included only 143 Hispanic children (in the whole normative sample) for the entire age range. The sample was therefore judged to be inadequate.

3. Do the authors provide an item analysis of the test items? The investigators independently agreed that the test provided an adequate item analysis. A table in Appendix A of the SPLS-3 manual addressed this criterion adequately by listing, for each item and for each version of the test (English and Spanish), the age at which the item was marked correct for 50%, 75%, and 90% of the children. The table indicated that, in general, the earlier an item appeared on the test, the higher the accuracy of the item. Most test items were acquired approximately 6 months later in the Spanish group than the age predicted for English, indicating a problem with the construct of the test, which is based on an English language development model.

4. Do the authors provide means and standard deviations? The investigators independently agreed that the presentation of means and standard deviations was marginally adequate. The manual did report means and standard deviations for each age interval, each subtest (Auditory Comprehension and Expressive Communication), and each version of the test (Spanish and English). However, the manual directs the clinician to use the English norms that, by themselves, lead to inappropriate use of the test. The authors did not report means and standard deviations for the separate dialect or language proficiency groups represented, or for the Hispanic group, which conceivably affected the norms.

5. Is there evidence of concurrent validity? The investigators independently agreed that this information was not provided for the Spanish version. Therefore, there is no evidence provided to judge concurrent validity.

6. Is there evidence of predictive validity? The investigators independently agreed that this information was not provided for the Spanish version. Therefore, there is no evidence provided to judge predictive validity.

7. Did the authors obtain test-retest reliability? The investigators independently agreed that, although the test provided coefficients of internal reliability, there was no statement of test-retest reliability for the Spanish version.

8. Did the authors obtain inter-examiner reliability? The investigators independently agreed that the test authors did not provide inter-examiner reliability for the Spanish version.

9. Do the authors provide a description of test procedures? The investigators independently agreed that the manual provided a description of the test administration procedures.

10. Do the authors provide a description of tester qualifications? The investigators independently agreed that this criterion was not met, especially in the context of

evaluating bilingual children. The only statement concerning the qualifications of the examiner read, "you must have knowledge of test administration and score interpretation" (original PLS-3 *Examiner's Manual*, p. 18). In the view of the investigators, appropriate evaluation of a Spanish-speaking child requires fluency in oral and written Spanish.

11. Is the purpose of the test explicitly stated? The investigators independently agreed that this criterion was not met. Although the purpose of the English version of the PLS was to assess receptive and expressive language skills in infants and young children, the supplemental manual for the Spanish version does not address the purpose. Rather, the examiner is left to infer that the purpose is to assess Spanish language development.

12. Is the construct or model explicitly defined and does it relate to the stated purpose? The investigators independently agreed that this criterion was not met. Again, the two versions of the test must be addressed. The English version refers to Lahey's (1988) language construct of form, content, and use, and English language development. The manual for the Spanish version does not refer to such constructs as a theoretical base. Moreover, the same theoretical base should not be inferred; clinicians cannot assume that the theoretical constructs developed for English apply to Spanish as well, given that the development of aspects of spatial relations, morphology, vocabulary, and language use differ across languages. For example, we should not assume that the vocabulary, semantic, and grammatic indicators of development that are referenced in the SPLS-3 are universally applicable to all Latin American cultures. Test items such as stamps, wagons, prepositions, time concepts, colors, shapes, and money are not necessarily culturally or developmentally appropriate. Therefore, the defined construct (English language development) does not appear to match the (unstated) purpose of the SPLS-3 (Spanish language development).

13. Is there a clear, supportable rationale for the selection of test content? This question relates to Criterion 12. The investigators independently agreed that the answer is negative because the items were translated from the English version of the test, with no Spanish developmental data provided to support the selection of items. Moreover, in the absence of developmental data, authors would be expected to at least provide a rationale for the inclusion of each item. Such information is not provided in the SPLS-3 manual.

14. Does the test adequately sample behavior at the extremes? The investigators came to a consensus that there is not enough information to answer this question, given the problems described with item translation and lack of use of Spanish developmental norms. This question is addressed further through our own item analysis in Phase Two (see the Discussion section).

15. Do the norms represent performance at the extremes? The researchers independently agreed that the SPLS-3 manual did not attempt to address this issue. The test manual described the English and Spanish samples, but it did not address the issue of sampling children at the low end of the continuum, although it was designed for these children.

Emperical Questions

In addition to the psychometric criteria, we answered the following empirical questions to examine the validity of the SPLS-3:

1. To what extent are the SPLS-3 norms representative of a local sample of bilingual Spanish/English-speaking children? The local, or study, sample was compared to that of the test, using standard scores, the mean, and the standard deviation (mean standard score = 100, SD = 15). Each child's test results were converted to a standard score according to the SPLS-3 manual. Scores were compared with the English norms as directed in the manual. The study's sample mean (n = 37) for total SPLS-3 language score was 77.16 (SD = 14.07), which placed the mean for this group at 1.52 SD from the test mean. The mean for the Auditory Comprehension subtest score was 78.38 (SD = 9.5), which placed the study sample group at –1.44 SD from the test mean. The mean for the Expressive Communication subtest score was 80.38 (SD = 19.40), which placed the study sample group mean at –1.31 SD from the test mean. Furthermore, 81% (30/37) of the study sample children scored below 1 SD from the test's means.

As noted earlier, the SPLS-3 manual also reports means and standard deviations for an experimental Hispanic group (n = 181), but the examiner is referred to the English test norms and is urged to interpret the results with caution. To compare the means of the test's experimental Hispanic group with those of the study's Hispanic group, Table 2 provides the means and standard deviations for the study sample, the test's experimental Hispanic sample, and the test's English sample. The mean scores for children in the study sample are below those of the test's experimental Hispanic group scores. Furthermore, 51% (19/37) of the study sample children scored –1 SD below the SPLS-3 experimental Hispanic group's mean, indicating that the children in the study scored at the low end of the test's distribution, even when compared with the experimental Hispanic group.

2. To what extent does the SPLS-3 have construct and content validity? An item analysis from the study sample was performed. The numbers of participants who answered an item correctly and incorrectly were calculated to examine the overall progression of item difficulty on each subtest. This type of analysis is crucial given that the SPLS-3 is a developmental scale that assumes an age progression and uses base and ceiling scores.

Results indicated that the ratio of children who answered the item correctly to children who were administered the item did not decrease systematically across the items of either subtest. Figures 1 and 2 illustrate this uneven item progression across both subtests. For example, in the Auditory Comprehension portion, Items 31, 32, 33, and 36 were answered correctly by fewer than 50% of the children who were administered those items, whereas Items 34, 37, and 38 were answered correctly by more than 80%. Figure 1 shows the item analysis of the Auditory Comprehension portion.

On the Expressive Communication portion, the expected decrease of correct responses from the subtest's start to its finish was even less evident than on the

TABLE 2. Children's performance compared to the test's sample on the Preschool Language Scale–Spanish version.

| | Spanish Preschool Language Scale Scores | | | | | |
| | Auditory Comprehension | | Expressive Language | | Total Score | |
Sample	M	SD	M	SD	M	SD
Test–English (n = 451)	100.00	15.00	100.00	15.00	100.00	15.00
Test–Hispanic (n = 181)	89.90	14.50	90.00	15.40	88.90	14.80
Study (n = 37)	78.38	9.50	80.38	19.40	77.16	14.07

Auditory Comprehension subtest. For example, although Expressive Items 23, 24, and 29 were answered correctly by fewer than 50% of the children who were administered those items, later items, such as 30, 31, 36, 40, and 48, were answered correctly by more than 60%. In fact, the last item, Number 48, was answered correctly by 80% of the children to whom it was administered. Item 48 probes irregular plurals in the English version of the scale, whereas in the Spanish version, the item probes regular plural forms, explaining the ease with which Spanish-speaking children would respond to this question. Figure 2 shows the item analysis of the Expressive Communication subtest.

Items were further analyzed for their appropriateness and relevance in the assessment of Spanish language development. Several items contained vocabulary that was judged to be developmentally inappropriate. For example, Item 36 of the Auditory Comprehension portion was determined to include high-level vocabulary with which a Hispanic child may not have experience: parachutes, wheelbarrows, or stamps. Further, although the use of prepositions is an important concept in English, for Spanish speakers, it is a less specific concept and therefore is not marked as precisely as in English (Slobin & Bocaz, 1988). As would be expected, the children in the study had greater difficulty with Item 33 in Auditory Comprehension, which evaluates the comprehension of spatial concepts.

Further, Item 23, which less than 10% of the children passed, requires children to talk about an animal in sentences containing 4–5 words, a task that may have been unfamiliar to the children.

Similarly, in the Expressive Communication portion of the test, children are asked to perform tasks that are probably unfamiliar to them, such as providing definitions and completing analogies. Low performance on Items 31 (use of auxiliaries) and 32 (answering obvious "when" questions) is likely to reflect difficulty with the tasks, although they require early-acquired language skills in Spanish. Moreover, aspects of the language that are important for the identification of language disorders in this population are not addressed, such as gender agreement in the noun phrase, person agreement in the verb phrase, the use of clitics, and the use of a variety of tenses and moods (e.g., Restrepo & Gutierrez-Clellen, 2001).

3. To what extent does the SPLS-3 have criterion validity, and specifically concurrent validity, when comparing an independent sample of children's performance on the SPLS-3 with their performance on two descriptive Spanish language measures, a spontaneous language analysis, and a parent interview? Four correlation analyses with the MLTU, the CRR and CRE measures, and the Spanish PreLAS were obtained. An alpha level of .05 was selected and subjected to the Bonferroni

FIGURE 1. Item difficulty analysis of the Auditory Comprehension subtest.
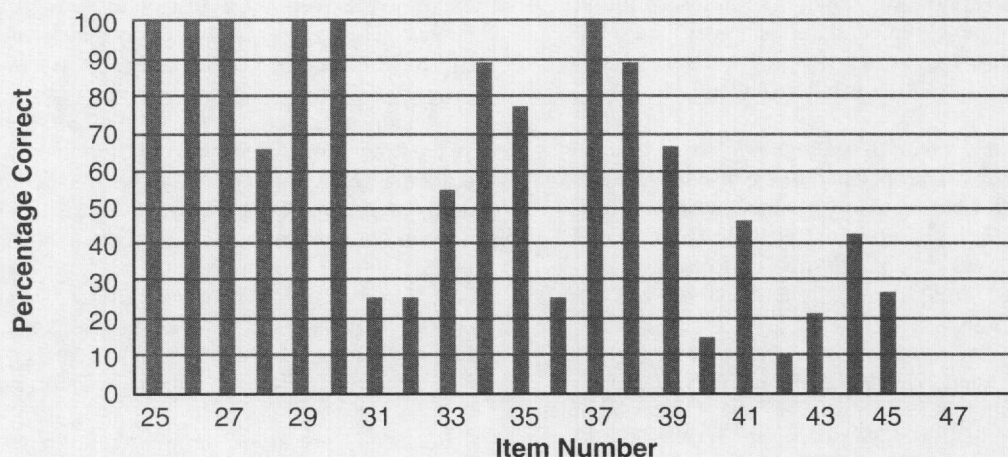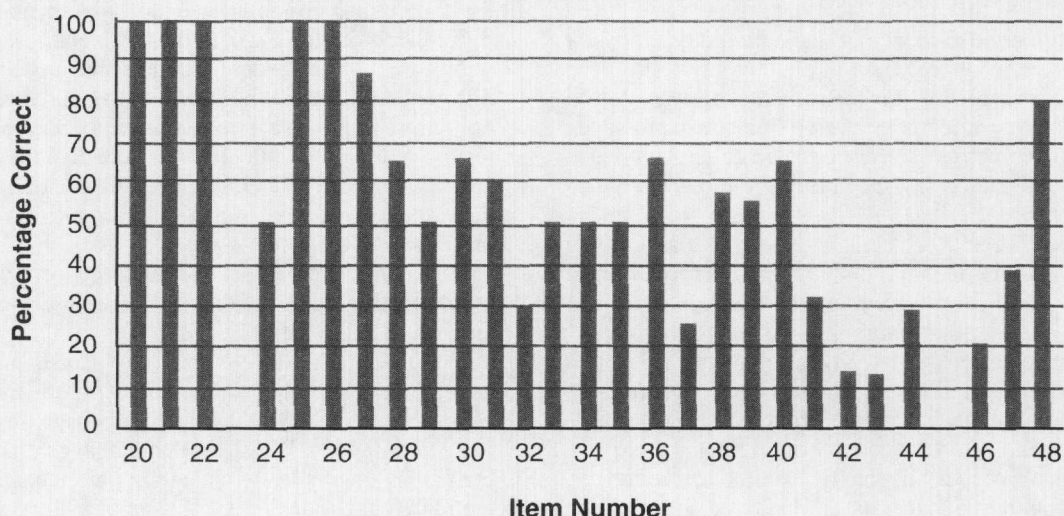
FIGURE 2. Item difficulty analysis of the Expressive Communication subtest.

adjustment of .0125 for the correlations. Results indicated that the Auditory Comprehension portion of the SPLS-3 correlated significantly with the CRR measure ($r = .44, p = .006$). The Expressive Comprehensive portion of the SPLS-3 did not significantly correlate with the two expressive measures, MLTU ($r = .33 p = .08$) and CRE ($r = .05, p = .75$). The total score of the SPLS-3 did not correlate significantly with the Spanish PreLAS ($r = .40, p = .015$). Table 3 lists the means and standard deviations for all of the measures.

No correlations were run for the parental interview measure because results on this measure were bimodal, which would affect the correlations. Therefore, children were classified according to "concern" or "no concern" based on the parent interview. Those children whose parents indicated concern about a language disorder (i.e., with a score of 7 or greater on the parent interview measure) were examined further to determine whether they scored below –1 SD on the SPLS-3 study's sample mean. (A score of –1 SD of the current sample was selected because of the possibility of including children with language problems.) Results indicated that, of 37 children, 6 children qualified as at risk on the parent interview measure (score >7). Of those children, the means for the total SPLS-3 score were as follows: 66, 57, 74, 73, 73, and 83. Of these 6 children, 1 child scored below –1 SD (9.17), or below 64.25 from the study's sample mean (73.42), and the other 5 scored within 1 SD. According to the English norms, all of the children with scores greater than 7 on the parent interview would score below –1 SD, but so would most of the children in the study (30 of 37 or 81% of the children).

Discussion

Psychometric Criteria

To address whether the SPLS-3 has evidence of logical validity through psychometric criteria, we examined the manual and used 15 criteria from McCauley and Swisher (1984) and Hutchinson (1996). The SPLS-3 met 4 of 15 psychometric criteria (Criteria 1, 3, 4, and 9). Of the 4, 2 were judged by the researchers to be marginally adequate (Criteria 1 and 4). Further, the researchers disagreed on 2 of the 15, therefore arriving at a decision through consensus (Criteria 1 and 14). The criteria included a description of the normative sample, provision of an item analysis, provision of means and standard deviations, and description of test procedures.

TABLE 3. Means and standard deviations for age and performance on measures.

Measure	n	M	SD
Age (years)	37	5.56	0.61
English PreLAS (1–5 levels)	37	2.00	1.22
Spanish PreLAS (1–5 levels)	37	3.57	0.93
Criterion-referenced receptive measure (out of 45 possible correct responses)	37	12.35	2.77
Mean length of terminable unit (words/T-unit)	29	5.00	1.26
Criterion-referenced expressive measure (out of 20 possible responses)	37	17.22	5.55
Parent report of language problems (number of yes responses in 17 questions)	22	4.32	4.42

Note. PreLAS = Preschool Language Assessment Scale; T-unit = terminable unit.

Results suggested that the lack of psychometric criteria was most likely related to the limited information provided in the manual, and thus we assume that the SPLS-3 presents with limited evidence of logical validity. The lack of evidence then leads to difficulty making informed decisions when reviewing this measure and to possible misuse or inappropriate interpretation of the test. However, the major psychometric issues of concern can be divided into two general categories: the norming sample and the construct of the test.

The norming sample. The characteristics of the norming sample are a major concern in the use and interpretation of the SPLS-3. Participants were recruited from six states and Puerto Rico, representing roughly five dialect groups and many different countries with many different cultures. Although cultural variation may not be an issue for some aspects of language, vocabulary and concept development may differ depending on the country of origin and the educational background. To remedy this problem, test developers would need to test for differences among dialect groups to determine if separate norms are warranted, and, if not, to include this information in the manual.

Further, the children's language history and background are of great importance in determining the children's strongest language and how they perform, depending on whether English or Spanish is the most proficient language. According to the description of the experimental group, most, but not all, of the sample consisted of bilingual children. However, it is unclear what the test developers mean by this because the manual does not state whether, or how, Spanish language proficiency was established before administration of the SPLS-3. This problem could be addressed by either controlling for language proficiency or dividing children into groups to be tested separately for differences. At the very least, the authors might provide norms for strictly monolingual children, mostly Spanish proficient children, mostly English proficient children, and English-Spanish proficient children.

Determining language proficiency is an important aspect of bilingual language assessments because, as is noted in the SPLS-3 manual, bilingual children in general present a different profile of development than do monolingual children. Thus, the test results of the experimental group may be based on a linguistically heterogeneous sample and may not apply to children with varying degrees of English proficiency.

The size and description of the norming sample represent an additional limitation because the Hispanic sample is smaller ($n = 143$ for the whole age range) than the recommended 100 children per language and age group (McCauley & Swisher, 1984). The implication of a smaller sample size is that the larger populations may not be well represented in the norms, potentially leading to over- or under-identification of disorders.

The second issue related to the norming sample concerns how to interpret results based on the reported sample's means and standard deviations. When these data are provided, clinicians still must address several questions. For example, should reliance be placed on the English-normed sample that, as stated above, does not have adequate Hispanic representation, dialect representation, and language proficiency information? If the clinician does use the English norms, what is considered a meaningful difference from the mean to qualify a child as presenting with a language disorder? Should clinicians count the difference from the experimental Hispanic means or the normative English sample? Because there was no apparent testing of children with disorders and no evidence of concurrent validity testing, clinicians do not have the data necessary to arrive at informed decisions, thus limiting the usefulness of the test.

Test construct. The second major concern with the psychometric quality of the test is with the construct of the test. As stated earlier, direct application of the English model of language development to Spanish language development is problematic. Moreover, this limitation can be applied to the identification of language disorders in these populations. We know that language disorders vary, both cross-linguistically and among bilingual versus monolingual children (e.g., Restrepo, 1998; Restrepo & Gutierrez-Clellen, 2001; Restrepo & Kruth, 2000). Therefore, universal areas of difficulty cannot be assumed, and the test may not be appropriate for the identification of Spanish-speaking children with language disorders.

Construct problems are further evidenced in the item analyses. Given that this test uses a base and ceiling format, item difficulty for its translations is particularly relevant. Again, item difficulty is not necessarily equivalent from language to language; therefore, a lack of logical item difficulty progression can affect how children perform, and when or whether they obtain adequate base and ceiling scores.

The results of the SPLS-3 item analyses further suggest that there is a gap in children's development between the English and the Spanish versions of approximately 6 months. In some cases, however, there is a gap of as much as a year, with some skills not acquired at all by the Spanish-speaking children. This appearance of a lag in development is problematic in test interpretation and logical validity; it appears that the development of language skills specific to Spanish-speaking children was not considered in developing SPLS-3 items, resulting in problems of content and construct validity.

In conclusion, the construct of the test was judged to be inadequate because it was based on English rather than Spanish language developmental data. This impression was confirmed by an item difficulty analysis. Additionally, the SPLS-3 presented a limited norming sample with no dialectal or language proficiency information. Finally, the manual did not address many of the issues that are considered to be essential components of a standardized measure. We concluded that many important psychometric issues appear to have been neglected in the test's construction, leading to its questionable usefulness as a measure of Spanish language development.

Empirical Validity

Comparison to the norms. We examined the validity concerns with our own empirical data, using an independent

sample of Spanish-speaking children and independent measures, some of which have been independently validated for this population (Gutierrez-Clellen et al., 2000; Restrepo, 1998). To address the first question, related to how the current study sample is represented in the norms, performance on this independent sample was compared with the norms. Results revealed that this group's performance on the SPLS-3 was well below the test's norms, indicating that the test norms are not representative of our study's sample. Consequently, the use of the SPLS-3 can result in the over-identification of language impairment in Spanish-speaking children who have normal language skills, at least for children in the southeast United States who attend English-only programs. Further, even if norms were adapted, scores would not be interpretable given the possible overlap between children in the lower end of the normal distribution and those with language disorders.

The low performance of the current sample, compared to the SPLS-3 Hispanic experimental means and standard deviations reported in the manual, is a concern. The total score for the children in the current sample was at least 15 points below the total score for the test's experimental Hispanic sample. There are two possibilities that may explain the low performance. First, the children in the study attend English-only schools. Studies have shown that such an environment does not promote development of a child's first language, either socially or academically, if that language is other than English. This in itself is not necessarily a problem, unless the children do not have support for first language use in their homes, through rich social- and literacy-based language experiences provided in relevant contexts. Given the situation of some of the children in our study, Spanish-language-rich experiences may not be possible. Parents often worked two jobs or evening jobs when the children were at home, and the children were cared for by other siblings in the family. These findings support the need to provide an adequate description of the language history, proficiency, and background when norming the test so that meaningful comparisons can be made.

A second possible explanation for low performance of our sample group may be the inclusion of children with language disorders. Although no child had been identified with language disorders, and no child scored low on both parent interview and MLTU, the current sample may have included a few children who did have a disorder. Nevertheless, it is clear that separate norms should be obtained for children with varying language backgrounds and levels of proficiency, given that there is so much variability in their language experience and use. Finally, we underscore again the need for test developers to focus on Spanish and bilingual language development as the theoretical framework for the construction of measures.

Item analysis. The second question addressed the extent to which the SPLS-3 has construct and content validity, based on item analysis data from a local sample. The item analyses of the children's performance in the study indicated several problems with the content of specific items. Test performance may be affected by factors that are not related to language disorders, such as lack of familiar-

ity with the test format, difficulty comprehending poorly translated items (i.e., Item 31 in Auditory Comprehension), difficulty with the vocabulary used as the level of difficulty increases (i.e., Item 32 in Expressive Communication) or decreases (i.e., Item 48 in Expressive Communication) with translation, and scoring that penalizes children's culturally appropriate responses (i.e., Items 28 and 32 in Auditory Comprehension). Items that were translated directly from English to Spanish without identification of potential differences in the difficulty level of linguistic concepts across the two languages might be expected to be problematic. Progression of item difficulty relates to both content and construct validity concerns, issues that are not trivial for the clinician engaged in test selection (for discussions, see Hutchinson, 1996; McFadden, 1996).

Criterion validity. Criterion-related validity analyses indicated that the SPLS-3 total score did not significantly correlate with the Spanish PreLAS and the expressive language measures, although the SPLS-3 Auditory Comprehension subtest significantly correlated with the CRR measure. However, this correlation is low, indicating that the SPLS-3 does not measure aspects of Spanish language development in language comprehension. Similarly, results of the expressive language measures suggested that the SPLS-3 does not measure grammatical development in Spanish. Further descriptive analysis of the SPLS-3, given parent concerns, indicated that the SPLS-3 is not sensitive to language development and disorders in Spanish-speaking children given the variability in performance of the children judged at risk in the parent interview.

Clinical Implications

Although many school systems require the use of standardized test scores to support qualification for services and determination of eligibility, the use of standardized scores is not required for children for whom there are no such validated measures. In fact, the use of such measures may violate children's rights to an appropriate and unbiased assessment. In the case of the SPLS-3, the test could actually result in over-identification of children as having language impairments that, in turn, could cost the school system in terms of the provision of unnecessary services. In our view, then, clinicians who are pressured by administrators to use standardized measures can defend their clinical decisions as complying with the Individuals With Disabilities Education Act Amendments of 1997. This law states that a child's examination must consist of procedures that have been validated for the specific purpose for which they are used (in this case, identification of a language disorder). In addition, these procedures must be selected and administered so as not to be discriminatory on a racial or cultural basis.

As early as 1983, Taylor and Payne (1983) commented on the lack of effort to make measures of handicapping conditions linguistically and culturally valid. Now, some 17 years later, we echo this concern. The number of published tests for Spanish-speaking children has increased; however, these tests have no norms, nor do they meet accepted psychometric criteria (Langdon, 1992), and

research with these tests is limited (e.g., Anderson, 1996; Restrepo, 1998). Until standardized measures are developed to meet reliability and validity criteria, Spanish-English bilingual children should be evaluated with alternative measures, that is, using a thorough parent interview and language sample analyses (mean length of utterance and number of grammatical errors per utterance, as well as structured elicitation). Also, dynamic assessment procedures can be used to identify language disorders in preschool- and kindergarten-age bilingual Hispanic children in the United States (Anderson, 1996; Gutierrez et al., 2000; Peña, 1996; Restrepo, 1998). Although further research with these measures and procedures is warranted, they have been validated for this purpose (Anderson, 1996; Gutierrez et al., 2000; Peña, 1996; Restrepo, 1998). Although these procedures require a greater time investment than standardized measures, they can provide information about the strengths and weaknesses of the child's language that may actually save the clinician intervention planning time. Furthermore, the physical, emotional, and social cost of over-identifying language disorders in children through the use of inappropriate measures certainly justifies a more time-consuming, but also more thorough and accurate, assessment approach. Researchers who work with culturally diverse populations should continue to describe language development and the characteristics of language disorders, and continue to investigate the use of measures that are less culturally and linguistically biased toward these populations. Clinicians and test developers must then use this information in their professional roles.

Conclusions

The results of the current study indicate that the SPLS-3 not only has poor evidence of logical validity, but also has limited evidence of empirical validity, poor description of the standardization and norms, and poor evidence of reliability—areas that have been determined to be critical for the development of an adequate test. Having used 15 psychometric criteria to evaluate the SPLS-3, the authors believe that this measure is no better or worse in terms of its psychometric adequacy than the majority of tests reviewed by McCauley and Swisher in 1984. Given the current knowledge base, however, we should expect better results 17 years later. At least we currently have informal measures that are validated for the identification of language disorders, and there is increased interest among national funding agencies in supporting this type of research.

Finally, the results of the current study should be evaluated with caution as well. The number of participants, for example, was small, and participants came from backgrounds that were different from those in the original test. Furthermore, the lack of valid measures for the identification of language disorders in Hispanic children limits the validation of measures for this purpose and makes the process of participant selection more complex. Continued research is needed to validate measures that are accurate and valid for the identification of language disorders in Hispanic children.

References

Anastasi, A. (1988). *Psychological testing* (3rd ed.). New York: Macmillan.

Anderson, R. (1996). Assessing the grammar of Spanish-speaking children: A comparison of two procedures. *Language, Speech, and Hearing Services in Schools, 27,* 333–345.

Beaumont, C., & Langdon, H. W. (1992). Speech-language services for Hispanics with communication disorders: A framework. In H. W. Langdon (Ed.), *Hispanic children and adults with communication disorders: Assessment and intervention* (pp. 1–19). Gaithersburg, MD: Aspen.

Dale, P. S. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Research, 34,* 565–571.

Duncan, S. E., & DeAvila, E. (1986). *Preschool Language Assessment Scale.* Monterey, CA: McGraw-Hill.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39,* 643–654.

Erickson, J., & Iglesias, A. (1986). Assessment of communication disorders in non-English proficient children. In O. L. Taylor (Ed.), *Nature of communication disorders in culturally and linguistically diverse populations* (pp. 181–217). San Diego, CA: College-Hill Press.

Gavin, W. J., Klee, T., & Membrino, I. (1993). Differentiating specific language impairment from normal language development using grammatical analysis. *Clinical Linguistics & Phonetics, 7,* 191–206.

Gutierrez-Clellen, V. F., & Hoffstetter, R. (1994). Syntactic complexity in Spanish narrative: A developmental study. *Journal of Speech and Hearing Research, 37,* 645–654.

Gutierrez-Clellen, V. F., Restrepo, M. A., Bedore, L., Peña, L., & Anderson, R. (2000). Language sample analysis: Methodological considerations. *Language, Speech, and Hearing Services in Schools, 31,* 88–98.

Hispanic Task Force. (1999). *Report of the Hispanic Task Force to the Board of Regents University System of Georgia.* Athens, GA: Author.

Huang R., Hopkins, J., & Nippold, M. A. (1997). Satisfaction with standardized language testing: A survey of speech-language pathologists. *Language, Speech, and Hearing Services in Schools, 28,* 12–29.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). Champaign, IL: National Council of Teachers of Education.

Hutchinson, T. A. (1996). What to look for in the technical manual: Twenty questions for users. *Language, Speech, and Hearing Services in Schools, 27,* 109–121.

Individuals With Disabilities Education Act, 20 U.S.C. § 1400 *et seq.* (1997).

Lahey, M. (1988). *Language disorders and language development.* New York: MacMillan.

Langdon, H. W. (1992). Speech and language assessment of LEP/bilingual Hispanic students. In H. W. Langdon (Ed.), *Hispanic children and adults with communication disorders: Assessment and intervention* (pp. 201–271). Gaithersburg, MD: Aspen.

Mattes, L. J., & Omark, D. R. (1991). *Speech and language assessment for the bilingual handicapped.* San Diego, CA: College-Hill Press.

Mayer, M. (1967). *A boy, a dog, and a frog.* New York: Dial Books for Young Readers.

Mayer, M. (1969). *Frog where are you.* New York: Dial Books for Young Readers.

McCauley, R. J., & Swisher, L. (1984). Psychometric review of language and articulation tests for preschool children. *Journal of Speech and Hearing Disorders, 49,* 34–42.

McFadden, T. U. (1996). Creating language impairments in typically achieving children: The pitfalls of "normal" normative sampling. *Language, Speech, and Hearing Services in Schools, 27,* 3–9.

Peña, E. D. (1996). Dynamic assessment: The model and its language applications. In K. N. Cole, P. S. Dale, & D. J. Thal (Eds.), *Assessment of communication and language* (pp. 281–307). Baltimore, MD: Brookes.

Plante, E., & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25,* 15–24.

Pollard, K. (1999). U.S. diversity in more than black and white. *U.S. Population Data Sheet. Web site for the Population Reference Report.* www.prb.org/Content/NavigationMenu/PRB/aboutPRB/Population_Bulletin/Americas_Racial_and_Ethnic_Minorities.htm [Acessed August 2000].

Restrepo, M. A. (1998). Identifiers of Spanish-speaking children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1398–1411.

Restrepo, M. A., & Gutierrez-Clellen, V. F. (2001). Article production in bilingual children with specific language impairment. *Journal of Child Language, 28,* 433–452.

Restrepo, M. A., & Kruth, K. (2000). Grammatical characteristics of a bilingual student with specific language impairment. *Journal of Children's Communication Development, 21,* 66–76.

Scott, C., & Windsor, J. (2000). General performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *Journal of Speech and Hearing Research, 43,* 324–339.

Slobin, D. I., & Bocaz, A. (1988). Learning to talk about movement through time and space: The development of narrative abilities in Spanish and English. *Lenguas Modernas, 15,* 5–24.

Taylor, O. L., & Payne, K. T. (1983). Culturally valid testing. *Topics of Language Disorders, 3,* 8–20.

Toronto, A. S. (1976). Developmental assessment of Spanish grammar. *Journal of Speech and Hearing Disorders, 41,* 150–171.

Ventry, I. M., & Schiavetti, N. (1986). *Evaluating research in speech pathology and audiology.* New York: Macmillan.

Werner, E. O., & Kresheck, J. D. (1989a). *Spanish Structured Photographic Expressive Language Test–II.* DeKalb, IL: Janelle.

Werner, E. O., & Kresheck, J. D. (1989b). *Spanish Structured Photographic Expressive Language Test-Preschool.* Sandwich, IL: Janelle.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1993). *Preschool Language Scale-3: Spanish Edition.* San Antonio, TX: Psychological.

Contact author: M. Adelaida Restrepo, PhD, Department of Communication Sciences and Disorders, 557 Aderhold Hall, The University of Georgia, Athens, GA 30602. E-mail: arestrep@coe.uga.edu

Appendix

Translated Parent Interview Sample Questions

1. In comparison to other children of the same age, do you think your child has problems understanding or expressing himself/herself?

2. In comparison to other children of the same age, do you think your child has speech problems?

3. For your child's age, do you think your child has difficulty saying correct phrases?

4. For his/her age, does your child say very short phrases?

5. Do you think your child has grammar problems?

6. When your child talks about something that happened, does your child have difficulty explaining what happened or using different tenses? For example,...

7. When your child talks about people, does your child have difficulty using the correct pronouns? For example, ...

8. Does your child have difficulty differentiating if he or she is talking about a man or a woman?

9. In comparison to other children the same age, does your child use many general and nondescriptive words such as this, thing, that?

10. Does your child have difficulty finding exact words to express himself/herself?

11. Does your child have difficulty explaining or describing things?

12. Is your child frustrated because he or she cannot express him or herself well?

13. Do you have to repeat things to your child more often than to other children?

14. Do you think your child has difficulty learning new words?

15. In comparison to children the same age, does your child have a very low vocabulary?

16. Do you think your child has a learning problem?

17. In comparison to children of the same age, does your child have difficulty pronouncing words correctly?

18. Is your child's pronunciation easy to understand?