Published in final edited form as:

Learn Disabil Res Pract. 2013 August ; 28(3): 113–128. doi:10.1111/ldrp.12014.

Accurately Predicting Future Reading Difficulty for Bilingual Latino Children at Risk for Language Impairment

Douglas B. Petersen and

Assistant Professor in the Division of Communication Disorders at the University of Wyoming

Ronald B. Gillam

Professor and the Ray L. and Eloise Hoopes Lillywhite Endowed Chair in Speech-Language Pathology in the Department of Communicative Disorders and Deaf Education at Utah State University

Abstract

Sixty-three bilingual Latino children who were at risk for language impairment were administered reading-related measures in English and Spanish (letter identification, phonological awareness, rapid automatized naming, and sentence repetition) and descriptive measures including English language proficiency, language ability, SES, and preschool attendance at kindergarten. At the end of first grade, English word-level reading and reading comprehension were measured. Results indicated that the Spanish predictor measures did not account for significant variance over and above the English predictor measures for any of the first grade outcome measures. Of the descriptive predictor measures, only English language proficiency and language ability were significantly predictive, accounting for unique variance in first grade reading comprehension. Sensitivity ranged from .67 to .86 and specificity ranged from .82 to .93 across the four first grade outcome measures.

Reading difficulty likely represents the lower tail of a normally distributed ability (Shaywitz, Escobar, Shaywitz, Fletcher, & Makuch, 1992). Thus, it is estimated that only a small percentage of children in any population will have reading problems (Shaywitz, 1998). The prevalence of reading problems, however, is a significant concern in the U.S., especially for children who are culturally and linguistically diverse. According to the National Assessment of Educational Progress (2011), nearly 50% of Latino children read below a basic level at fourth grade, and 81% cannot read proficiently. Research has indicated that children who are at-risk for reading problems in their early elementary school years often continue to have difficulty reading into adulthood, entailing adverse academic and vocational consequences (Lyon, 2004; Scarborough, 1998). The first step in reversing this alarming trend is to develop valid early reading assessments that are not culturally and linguistically biased, for it is the early, accurate identification and subsequent prevention of reading difficulty that holds the most promise for reducing the prevalence of reading problems (Durlak, 1997; Gersten & Dimino, 2006; Walker, 1996). Unfortunately, most assessment methods are relatively poor at identifying reading difficulties in culturally and linguistically diverse

children (Catts, Petscher, Schatschneider, Bridges, & Mendoza, 2009; Lindsey, Manis, & Bailey, 2003). The purpose of this study was to explore the extent to which multiple English and Spanish reading and language measures were predictive of later word-level reading and reading comprehension problems in Latino children at risk for language impairment. A principal focus of the study concerned predictive evidence of validity and classification accuracy. Although actual reading difficulty is found in a small percentage of the population, this is not the case for children who have language impairment, where reading problems are highly prevalent (Catts & Hogan, 2003). Thus, children at risk for language impairment were specifically selected to provide an optimal opportunity to identify reading difficulty.

Language Minority Research

In their recent review of the literature pertaining to literacy development and second-language learners August and Shanahan (2006) concluded that phonological awareness in English and in Spanish, along with working memory, rapid automatized naming and measures related to letter knowledge in English were predictive of word-level reading ability in this population of children. August and Shanahan also found that oral language proficiency, including vocabulary knowledge in English, listening comprehension, syntactic skills, and metalinguistic skills were closely related to reading comprehension in bilinguals. Other factors such as cognitive ability, word reading ability, socioeconomic status, home-language use, and literacy practices and educational instruction were factors that influenced reading comprehension.

August and Shanahan (2006) offered recommendations for future research applicable to the early identification of reading difficulty in English language learners. They suggested that information about first-language background be collected (e.g., the amount of Spanish and English language exposure in the home, community and school) for the purpose of investigating the role of second-language (English) proficiency as it relates to reading fluency and comprehension. They also recommended that oral language ability be explored more closely as a potential predictive variable for English language learners. Although it has been fairly well established that phonological processing skills are more predictive of reading ability over oral language with monolingual English speaking students, August and Shanahan noted that these findings have not been confirmed for English-language learners with varying English language ability. Additionally, August and Shanahan recommended that other contextual variables be considered more closely, such as SES, instructional methods used in school (e.g., bilingual, ESL support, mainstream) and quality of language instruction.

Research on Latino Children

We conducted a literature search to identify and evaluate recent longitudinal research on measures that predict reading ability (broadly defined) in Latino children. Dependent variables were disaggregated whenever possible and arranged to fit within intermediary reading component skills that are formative of reading (e.g., word identification, decoding, reading fluency, vocabulary, and comprehension). These intermediary reading component skills were divided into two broad constructs; word-level reading and reading

comprehension. These two constructs are particularly important to examine separately for culturally and linguistically diverse children because the relationship between word-level reading, including measures of fluency, and reading comprehension may not be as correlative as it is for monolingual children (Mancilla-Martinez & Lesaux, 2011; Mancilla-Martinez, Kieffer, Christodoulou, Biancarosa, & Snow, 2011; Mancilla-Martinez, & Lesaux, 2010). The relationship between the predictor and criterion measures used in the research are reported in coefficients of determination (R^2), reflective of the regression analyses used in those studies.

The results of these Latino-focused studies are represented in Table 1. Perhaps the most striking findings are related to the role that first language (Spanish) measures play in second language (English) reading. For bilingual Latino children, measures administered in English *and* in Spanish have accounted for unique variance in English formative intermediary reading component skills (Lindsey, Manis, & Bailey, 2003; Manis, Lindsey, & Bailey, 2004; Oh, Haager, & Windmeuller, (2007); Pérez & Rinaldi, 2006; Swanson, Sáez, Gerber, & Lefstead, 2004). These research findings indicate that Spanish-related pre-literacy measures and Spanish language ability are predictive of reading in English for bilingual Latino children, lending support to the linguistic interdependence hypothesis (Cummins, 1979; Cummins, Swain, Nakajima, Handscombe, Green, & Tran, 1984). It should be noted, however, that not all Spanish related tasks were found to positively correlate with intermediary reading component skills. Hammer, Lawrence, and Miccio (2007), and Swanson et al. (2004) found that Spanish receptive language was *negatively* correlated with English word-level reading and reading comprehension, while Pérez and Rinaldi (2006) found that expressive vocabulary in both English and Spanish was predictive of word-level reading.

With word-level reading as an outcome measure, the individual and linear combination of predictor measures examined in the research with bilingual Latino children typically yielded moderate to high R^2 values. Letter identification alone emerged as a robust predictor (Lindsay et al., 2003; Manis et al., 2004, Oh et al., 2007). Although letter identification was a strong predictor, a combination of measures consistently accounted for significant variance over and above letter identification. These combinations included Spanish letter I.D., phonological awareness, expressive vocabulary, sentence repetition, concepts about print, and rapid automatized naming (Lindsey et al., 2003), Spanish letter I.D., phonological awareness, rapid automatized naming, and Spanish expressive language (Manis et al., 2004), English expressive vocabulary, sentence repetition, and phonological awareness (Pérez & Rinaldi, 2006), and English vocabulary and English and Spanish memory tasks (Swanson et al., 2004). Oh et al., (2007) also found that phonological awareness alone accounted for moderately high variance in word-level reading. Only Lindsay et al., (2003) and Manis et al., (2004) included reading comprehension as an outcome measure. Their studies, which included letter identification alone and the combination of Spanish predictor measures previously listed, indicated that the variance in reading comprehension was accounted for to a moderate-large degree.

It has been difficult to find an acceptable balance of sensitivity and specificity using any combination of English and/or Spanish predictor measures. For example, Lindsey et al.

(2003) examined the predictive evidence of validity of the Spanish measures administered at kindergarten and reported sensitivity values between 63% to 77% and specificity values between 60% to 78% for predicting reading problems with measures of first grade English decoding, word identification, and English and Spanish reading comprehension. Clearly, there is a need to identify accurate measures of reading problems in bilingual Latino children.

Only one of the studies investigated the extent to which Spanish reading-related measures accounted for significant variance in reading outcomes over and above similar English reading-related measures. Pérez and Rinaldi (2006) found that when using a combination of similar English and Spanish predictor measures to identify later English word-level reading the Spanish measures did not account for significantly more variance over English measures. The extent to which Spanish measures account for significant, unique variance in reading ability needs to be investigated further.

For descriptive measures, none of the studies that we reviewed included English language ability or proficiency measures to predict reading comprehension. Although Hammer et al., (2007) and Swanson et al., (2004) found that measures of receptive language were not predictive of *word-level reading* for Latino children, a strong relationship between language and *reading comprehension* is expected (Kamhi, 2009; Catts, 2009). Children who have limited English language proficiency or who are at risk for language impairment are at significantly greater risk for reading problems than the general population (Bishop & Snowling, 2004; Catts & Hogan, 2003; Snowling, Bishop, & Stothard, 2003). Assessing the sensitivity and specificity of reading measures in bilingual Latino children who are at risk for language impairment could serve as an efficient first step in the accurate identification of reading comprehension difficulty. This disproportionate stratified sampling of a population may reduce assessment costs while allowing unbiased classification estimates (Elliott, McCaffrey, Perlman, Marshall, Hambarsoomians, 2009).

In addition to language proficiency and language ability measures, none of the studies we reviewed included other measures that were identified by August and Shanahan (2006) as being important such as socioeconomic status (SES), home-language use and literacy practices, and prior educational instruction. These factors may be what is needed to increase the predictive evidence of validity of an early reading measure for Latino children.

It is unclear whether some of the independent variables identified from previous research relate to word-level reading. Many of the authors of the research reviewed defined reading through the simple view of reading (Gough & Tunmer, 1986), and did not specify whether decoding or comprehension were the target dependent variables. Nor did the authors in many cases attempt to determine whether a measure was predictive of both decoding and comprehension. Reading fluency was used as a criterion measure in only one study (Oh et al., 2007), and given that reading fluency is an integrated measure of word identification and decoding, it would seem appropriate to include that measure as a criterion variable in any study interested in predicting word-level reading.

The purpose of this study was to determine the extent to which descriptive measures, and reading-related measures administered to bilingual Latino children at risk for language impairment at kindergarten were predictive of carefully differentiated word-level reading and reading comprehension outcome measures at first grade. The primary research questions were:

1. What are the correlations between the descriptive and reading-related measures administered at kindergarten and the first grade reading criterion measures (nonsense word fluency, oral reading fluency, and word identification)?
2. How much variance in first grade reading criterion measures is accounted for by the kindergarten descriptive and reading-related predictor measures?
3. To what extent do Spanish predictor measures account for significant, unique variance in first grade reading criterion measures over and above the English predictor measures?
4. How well do the kindergarten measures classify children at risk for first grade word-level reading and reading comprehension difficulty?

METHOD

Participants

A disproportionate stratified sampling method was used to identify children at risk for language impairment (Sutcliffe, 1965; Koch, 1969). We preselected Latino, bilingual (Spanish-English) students who were at risk for language impairment before entering kindergarten. By only selecting English language learners who were at risk for language impairment, we increased the probability of including a higher proportion of children who would have future reading or reading-related learning disabilities. This was expected to yield a more accurate estimate of sensitivity and specificity because these were the children who would likely be referred for assessment and consideration for special education in school environments. Children with language impairment represent a large percentage of children with reading or reading-related learning disabilities (Catts & Hogan, 2003). Also, by reducing the number of participants, we were better able to efficiently administer a large battery of assessments and gather highly detailed information necessary to address our research questions.

The first phase of participant selection entailed the assessment of 249 Latino, bilingual children from three schools in a large urban school district in Utah. These children were participating in a large, multi-site study of assessment practices with bilingual children (Gillam, Peña, Bedore, Bohman and Perez, in press; Peña, Gillam, Bedore & Bohman, 2011). All 249 children were screened before entering kindergarten using the *Bilingual English Spanish Oral Screener* (BESOS, Peña, Bedore, Gutiérrez-Clellen, Iglesias, & Goldstein, in development). The screener consisted of four subtests: English Semantics, English Syntax, Spanish Semantics and Spanish Syntax. More information about the screener is provided by Bohman, Bedore, Peña, Gillam and Pérez (2010) and Peña, Gillam, Bedore and Bohman (2011). In the second phase of participant selection, we selected only those children at risk for language impairment. Children were considered to be at risk for

language impairment if they scored below the 30th percentile on one or more subtests in each language on the BESOS. Sixty-three Latino bilingual kindergarten children who were identified as at-risk for language impairment during pre-kindergarten screening participated in this longitudinal study. 15 children were lost to follow-up across the 2 years of the study leaving 48 participants who completed all the testing at each time point. Pertinent demographic characteristics of the sample and of the children lost at follow up are provided in Table 3.

The children attended general education kindergarten and first grade classrooms in which English was the primary language of instruction. The estimated median household income in the city was \$39,711, while the estimated median household income for the U.S. at the outset of the study was \$60,374. Over 75% of the children in the school district were classified as economically disadvantaged. Ethnicity in the school district was about 50% Hispanic and 50% European American. The three schools from which the participants were selected had a predominantly Hispanic population (over 80%), and had not met adequate yearly progress in reading for at least 2 years prior to the outset of the study. There were 29 girls and 34 boys with a mean age of 65.3 months (SD = 4.2 mos.). Socioeconomic information, and context-specific information such as English/Spanish use and exposure in the home, and language ability were collected for each subject (see Table 2). These variables were collected at kindergarten for the purpose of detailed participant description. Information obtained from the participants' parent/guardian and school indicated that they were from low socioeconomic status backgrounds as calculated using weighted Hollingshead scores (Hollingshead 1975), with mother and father years of education averaging 8.7 (SD 3.1) and 8.1 (SD 3.2) years respectively. The children's hearing was tested by school district personnel. None of the participants had a hearing impairment at the outset of the study. Preschool attendance was variable (M = 1.0 years; SD = 1.0 years). Each participant who had attended preschool had been enrolled in a Head Start program with instruction primarily in English. Based on yearly and daily expressive language information gathered from parent/guardian interviews, the participants were considered to be bilingual, Spanish dominant, English language learners (see Table 2).

Procedures

Phase 1 of this study took place in the late fall of the students' kindergarten school year. Sixty-three Latino kindergarten children received four reading-related assessments in English that targeted phonological awareness (PA), rapid automatized naming (RAN), sentence repetition (SR), and letter identification (LtID) as well as three reading-related assessment measures in Spanish (PA, RAN and SR). In addition, descriptive information such as socioeconomic status (SES), preschool attendance (PrAt), English/Spanish use and exposure in the home (Exp/Sxp), English and Spanish language proficiency (ELP/SLP), and language ability were collected for each subject. These measures were administered at kindergarten for the purpose of exploring their utility as predictor measures for English word-level reading and comprehension ability at first grade. The first author, a clinical supervisor (certified bilingual speech-language pathologists) and several graduate and undergraduate students enrolled in a speech-language pathology program administered the kindergarten assessments. The Spanish measures were administered by fluent bilingual

English/Spanish speakers, all of whom had lived for at least 2 years in a Spanish-speaking country and were fluent in oral and written Spanish. Those examiners who were English language dominant only administered English measures.

Phase 2 of this study took place almost 2 years after the battery of kindergarten measures were administered, near the end of the participants' first grade school year. This phase included the collection of school district administered word-level measures of nonsense word fluency, and reading fluency. The same team of examiners administered a standardized, norm-referenced assessment reflective of word-level reading (word identification) and reading comprehension. Three first grade word-level measures (nonsense word fluency, oral reading fluency, and word identification) were chosen to serve as formative criterion measures representative of the construct of word-level reading in English. The reading comprehension measure was used to serve as a criterion measure representative of the linguistic comprehension construct as measured in the context of written language.

Phase 1 Measurement: Kindergarten Descriptive Measures

Parent/Guardian Questionnaire: A parent/guardian questionnaire was administered in person or over the telephone in either English or Spanish according to the parents' request. The questionnaire collected information on (a) qualification for free school lunch, (b) parents' occupation, (c) parents' level of education, (d) history of language exposure at home and at school, (e) history of preschool attendance and (f) Spanish dialect spoken in the home (e.g., Puerto Rican, Mexican).

A portion of the information collected was used to calculate weighted Hollingshead scores, an index of socioeconomic status (SES; Hollingshead, 1975). An education score ranging from 1 to 7 (1 = less than 7th grade education) and an occupation score ranging from 1 through 9 (1 = farm labor/menial service jobs) were assigned for each of the participants' parents/guardians. The education and occupation scores were then weighted to obtain a single score for each parent/guardian, ranging from 8 to 66. Hollingshead scores below 30 were considered to be in the lower SES range.

Information about school lunch was collected to lend further insight into the participants' socioeconomic status. Eligibility for free school lunch was determined according to whether or not a student lived in a household that had a total income that fell below 130% of the federal poverty level, received TANF (Temporary Assistance for Needy Families), or received Food Stamps. Children meeting other special conditions including being homeless or who had parents that were migrant workers also qualified for free school lunch programs. Each of the participants qualified for free school lunch.

Home Language Profile/Familial Routine Survey: This survey was developed as part of the norming process for the Bilingual English-Spanish Assessment (BESA; Peña, et al., in development). It was designed for collecting information about what activities the child participated in each hour of the day during the week and during the weekend, with whom the child interacted, and in which language the child was spoken to and responded. The parent interview was conducted just prior to the child's enrolment in kindergarten. Parents were asked to recount a typical day during the week and during the weekend listing all of the

major activities their child participated in each hour (e.g., dinner time, watching T.V.) and specifying with whom their child interacted and in which language their child was exposed to *receptively*, and which language their child used *expressively*. In addition, the parents were asked to indicate yearly *expressive* language information to better understand the extent that the child used English, Spanish or both languages across their lifespan.

For descriptive purposes the participants were classified into five language proficiency groups following procedures similar to those used by Peña, et al., (2011), Hammer, Miccio, and Rodríguez (2004) and Kohnert, Bates, and Hernández (1999). Based on yearly language use data from the language questionnaire, the five language groups were: functional monolingual English (i.e., 80% of expressive language history was English), functional monolingual Spanish (i.e., 80% expressive language history was Spanish), bilingual English dominant (60–80% expressive language history was English), bilingual Spanish dominant (60–80% expressive language history was Spanish and balanced bilingual (40–60% expressive language history was English and Spanish). This analysis indicated that 46% ($n = 29$) of the participants were functionally monolingual Spanish, 16% ($n = 10$) were bilingual Spanish dominant, 3% ($n = 2$) were bilingual English dominant, and 38% ($n = 24$) were balanced bilinguals. None of the children were functionally monolingual English (see Table 2). These language data were collected 1 week prior to the participants' enrollment in kindergarten, however, the predictor measures were administered 3 months following the language survey, and the participants' language of instruction was primarily English during those months of kindergarten instruction. Thus, English language proficiency was likely underestimated to some degree.

Language Proficiency and Language Ability: In the context of bilingualism, and consistent with Cummins (1984), we use language proficiency to mean the extent that someone uses and understands a specific language (e.g., English as a second language). We use language ability to mean the extent that someone is able to learn any language, including his or her first language. The English and Spanish semantics and morphosyntax subtests of the Bilingual English-Spanish Assessment (BESA; Peña, Gutierrez-Ciellen, Iglesias, Goldstein & Bedore, n.d.) were used to help determine the language proficiency and language ability of the 63 children. The BESA semantics and morphosyntax subtests in Spanish and English contain a greater number of items than the BESOS screener subtests, and are designed to more accurately represent the construct of language and to accurately identify language impairment. The format of the subtests in each language is similar, but the test items are representative of known markers of language impairment and critical aspects of speech and language development in the target language. Across the subtests in Spanish, coefficient alpha values range from .784 to .840. Coefficient alpha values for the English subtests range from .812 to .918, indicating good internal consistency within subtests for both languages. The BESA has preliminary normative information derived from a sample of over 1500 bilingual Spanish/English Latino children.

Scores on the BESA subtests in English provided information on English language proficiency. Information from the parent questionnaire, which yielded language dominance classifications, and information from the BESA English and Spanish subtests were cross-referenced to reflect language ability. Those children who were classified as functional

monolingual Spanish, and bilingual Spanish dominant, and who scored below the 30th percentile on the Spanish Semantics *and* Spanish Morphosyntax subtests of the BESA were classified as having low language ability. Likewise, children who were classified as functional monolingual English and bilingual English dominant and who scored below the 30th percentile on the English semantics *and* English morphosyntax subtests were classified as having limited language ability. Children who were balanced bilingual had to score below the 30th percentile on all of the English and Spanish subtests to be considered to have limited language ability. Eight of the 63 participants (12.7%) were classified as having limited language ability. None of the children who were classified as bilingual scored below the 30th percentile on all of English and Spanish BESA semantics and morphosyntax subtests. The mean language ability z-score derived from the combined English or Spanish (depending on language dominance) BESA subtests was .97 with a standard deviation of 1.46 and a range of -2.63 to 3.51, indicating a wide distribution of language ability among the participants (see Table 2). Based on these results, we classified the participants into two different language ability groups: those having typical language ability (TL) and those having language difficulty (LgD). These classifications are considered to be a gross estimate of language ability in the domains of semantics and morphosyntax in the dominant language.

Phase 1 Measurement: Kindergarten Reading-Related Measures—The following reading-related measures were included in the kindergarten screening battery based off of the comprehensive review of literature that implicated such measures as being predictive of reading ability for Latino children. Equivalent English and Spanish reading-related measures were administered. Whenever possible, best language scores (BLS) were also calculated. The BLS represented the highest raw score obtained on either the English or Spanish measures. Instead of providing information from two separate language systems, the BLS measure provided a conceptual score that reflected a child's system as a whole (Kester & Peña, 2002; Pearson, Fernandez, & Oller, 1992, 1993). Those measures that were in Spanish and English were administered in random order, with English or Spanish measures administered first about 50% of the time. If Spanish measures were to be administered first, the examiner spoke with the participant in Spanish prior to the testing session and throughout the Spanish assessment process. Likewise, if English measures were to be administered first, the examiner spoke English with the participant prior to, and throughout the administration of those tasks.

Letter identification (LtID): A letter identification task was designed and administered in English, however Spanish responses were accepted. The letter identification task included a total of 35 upper and lower case letters printed in various typefaces displayed in random order. For this task, the children were asked to name each letter following from left to right on the form. The task was not timed, and errors were recorded. Examiners awarded 1 point for each correctly named letter. Reliability of the letter identification measure was analyzed using Cronbach's alpha, which was .95. Convergent-correlational evidence of validity with the kindergarten DIBELS Letter Naming Fluency measure was .65, $p < .0001$.

Phonological awareness (PA) in English and Spanish: English and Spanish deletion subtests were administered to each participant. The assessments were modifications of tasks

developed by Catts, Fey, Zhang, and Tomblin (2001) and Rosner and Simon (1971). Children were told to say a word, and to then say the word again with a part omitted. English and Spanish instructions were identical (e.g., “Say newspaper. Now say newspaper, but don’t say news;” “Di porfavor. Ahora di porfavor, pero no digas por”). The Spanish version was not a direct translation of the English measure; the Spanish phonological awareness words were chosen based on two criteria: (a) the words had to be relatively familiar to kindergarten-age Spanish speaking children, and (b) some of the words had to be segmentable syllabically. And, as with the English PA task, the Spanish task required the children to omit increasingly more minute parts of the words, moving from the syllabic to phonemic level (e.g., *sacapuntas*, *secar*, *feo*). The Spanish words ranged from two to four syllables and the English words ranged from one to three syllables. Cronbach’s alpha reliability coefficients were as follows: .89 for English phonological awareness and .86 for Spanish phonological awareness. Convergent-correlational evidence of validity for the English phonological awareness measure with the kindergarten DIBELS Initial Sound Fluency measure was .52, $p < .0001$, and with the kindergarten DIBELS Phoneme Segmentation Fluency measure .44, $p < .0001$.

Phonological awareness best language score: Phonological awareness best language scores (PA-BLS) represented the highest raw score obtained on either the English or Spanish phonological awareness subtest. This score was possible to obtain because of the equivalent number of test items across the two different phonological awareness subtests. The PA-BLS was purposed to reflect a child’s true phonological awareness ability with language removed as a confounding factor. The percentage of PA English scores and Spanish scores that contributed to the PA-BLS measure are shown in Table 4.

Sentence repetition (SR) in English and Spanish: Sentence repetition measures were designed in English and Spanish. Examiners asked the participants to imitate sentences after them, and the examiner was not allowed to repeat the sentence. Sentences were designed to increase in length and in syntactic complexity. The Spanish SR assessment was not a direct translation of the English assessment, however, just as in English, sentences were composed with the intent of moving from simple to complex syntactic and semantic content, and at times, were similar in content to the English sentences (e.g., *El perro esta corriendo. No me dijiste que ibas a Mexico. Ayer el raton cayó en las garras del gato hambriento*). MLU was 8.3 for the English sentences and 7.8 for the Spanish sentences. Cronbach’s alpha reliability coefficients were .79 for the English sentence repetition task and .67 for the Spanish sentence repetition. Convergent-correlational evidence of validity for the English sentence repetition measure with the English BESA composite language score was .62, $p < .0001$. Convergent-correlational evidence of validity for the Spanish sentence repetition measure with the BESA Spanish composite language score was .42, $p = .001$.

Sentence repetition best language score (SR-BLS): The SR-BLS represented the highest raw score obtained on either the English or Spanish sentence repetition subtest. This score was possible to obtain because of the equivalent number of test items across the two different subtests. The SR-BLS was designed to represent a child’s ability to repeat sentences with language removed as a confounding factor. The percentage of SR English

scores and Spanish scores that contributed to the SR-BLS measure are shown in Table 4. Convergent-correlational evidence of validity for the sentence repetition BLS measure with language ability was .42, $p = .001$.

Rapid automatized naming (RAN) in English and Spanish: These RAN tasks were based on the Rapid Automatized Naming of Animals assessment used in the Catts et al. (2001) study. Objects that would be familiar to children in both English and Spanish and that had the same number of syllables in English and the same number of syllables in Spanish were used. Simple drawings of a cat, a house, and a car (*gato, casa, carro*) were used in conjunction with the colors blue, black and red (*azul, negro, rojo*). The three colors were randomly assigned to the simple drawings, and the drawings were randomly placed in rows. There were six items in four rows yielding a total of 24 items in each language. Instructions were identical in English and in Spanish; “Say the color and name of the pictures as fast as you can;” “Dime el nombre y el color de los dibujos tan rapido como pueda.” The children then named as many items as possible in one minute.

Rapid automatized naming best language score (RAN-BLS): The rapid automatized naming best language scores (RAN-BLS) represented the highest raw score obtained on either the English or Spanish RAN subtest. This score was possible to obtain because of the equivalent number of test items across the two different subtests. The RAN-BLS was purposed to remove language as a confounding factor. In almost every case, the better rapid automatized naming score was derived from the English subtest. The percentage of RAN English scores and Spanish scores that contributed to the RAN-BLS measure are shown in Table 4.

Measurement: Phase 2 First Grade Measures—Measures representing the formative, intermediary components of word-level reading and reading comprehension were used as first grade criterion measures. These subtests were administered when the participants were in the first grade, approximately 1.5 years following the collection and administration of the kindergarten reading-related and descriptive predictor measures. Forty-eight participants were available for word-level reading assessment and 38 participants were available for reading-comprehension assessment.

Decoding: nonsense word fluency (NWF): The Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Kaminski & Good, 1996) Nonsense Word Fluency (NWF) subtest was the measure selected to represent first grade decoding ability. The DIBELS NWF is a standardized assessment designed to record the number of nonsense words a child can produce in 1 minute. The test is comprised of randomly ordered VC and CVC nonsense words (e.g., rav, sig, ov) and is scored according to the number of correct sounds produced. The test-retest reliability was reported to be .83 (Good et al., 2002). The convergent-correlational evidence of validity of the DIBELS NWF when compared to a curriculum-based measure of reading fluency administered 4 months later was .82 (Good et al., 2002).

Oral Reading fluency (ORF): The Dynamic Indicators of Basic Early Literacy Skills Oral Reading Fluency (ORF) subtest was used to represent reading fluency. The ORF subtest required that children read a passage aloud for 1 minute. Words substituted, omitted, and any

hesitations that last longer than 3 seconds were scored as errors. The total number of words read correctly in 1 minute comprised the oral reading fluency score. The test-retest reliability ranged from .92 to .97 and the alternate form reliability of different passages taken from the same level ranged from .89 to .94 (Tindal, Marston, & Deno, 1983). Convergent-correlational evidence of validity measured across 8 different studies ranged from .52 to .91 (Good & Jefferson, 1998).

Word Identification (Word ID): The Woodcock Reading Mastery Test-Revised (WRMT-R; Woodcock, 1987) Word Identification subtest (Word ID) was used to represent word identification. The Word ID subtest involved the recognition of sight words that were presented in isolation. The split half reliability for first grade Word Identification was .98. The internal consistency reliability (Cronbach's alpha; Cronbach, 1951) was reported to be .96. Convergent-correlational evidence of validity with the Woodcock-Johnson Reading Test was .69.

Reading Comprehension (RC): The WRMT-R (Woodcock, 1987) passage comprehension subtest was administered. Standard scores and percentiles were calculated. The passage comprehension subtest uses a cloze procedure that requires children to fill in a series of blanks according to the meaning of the surrounding sentences or phrases. The internal consistency reliability ranged from .68 to .98 and the split-half reliability ranged from .86 to .99. The WRMT-R norms are based on a sample population of approximately 3,700, reflective of the 1994 U.S. census.

Fidelity of Administration and Interrater Reliability—The first author directly observed 25% of the kindergarten and first grade measures that were administered by the research team. Fidelity of administration for each measure was at or above 93%. Interrater reliability was calculated for 20% of the kindergarten and first grade measures administered by the research team. Reliability was calculated point-for-point whenever possible. Interrater reliability was at or above 93% for each measure.

RESULTS

Descriptive Statistics and Correlations

Table 5 summarizes descriptive statistics for the kindergarten predictor measures and first grade criterion measures. We calculated bivariate Pearson product-moment correlation coefficients (r) to answer our first question. Results indicated that not all kindergarten predictor measures were significantly correlated with the first grade reading measures. Only those predictor measures that were significantly correlated were used in future regression analyses. Intercorrelations among the kindergarten measures and first grade measures are shown in Table 6.

Multiple Regression Analysis

Multiple regression analysis can help identify the influence of different traits or abilities from a large set of variables. This extraction of influential variables has bearing on the predictive evidence of validity of those measures on the constructs of interest (in this case

word-level reading and reading comprehension), to the extent that specification error has been curtailed. Specification errors in multiple regression, which include the failure to verify several assumptions, can lead to errors of inference, which in turn lead to the misapplication of test results, resulting in poor validity. Regression analysis requires that several assumptions be met before the sample-derived results can be projected inferentially to a population. Violation of the assumptions for regression analysis can lead to bias. Assumptions required for regression analysis were examined and any necessary transformations were conducted.

Based on the results of the correlation analysis, our regression analyses included the following kindergarten reading-related measures; letter identification (LID), English and Spanish phonological awareness (PA), English and Spanish rapid automatized naming (RAN), and English and Spanish sentence repetition (SR), and the following kindergarten descriptive measures; English language proficiency (ELP) and language ability (LA). These predictor measures were analyzed with the following first grade word-level reading and reading comprehension criterion measures; DIBELS nonsense word fluency (NWF), DIBELS oral reading fluency (ORF), the Woodcock Reading Mastery Tests-Revised word identification subtest (Word ID), and the Woodcock Reading Mastery Test-Revised reading comprehension subtest (RC). Results are organized according to respective research questions and are shown in Table 7.

Four multiple regression models were designed to answer our second research question. Models 1 and 2 investigated the extent that English and Spanish reading-related measures accounted for variance in the first grade reading criterion measures. The kindergarten English reading-related measures found to be significantly correlated with the first grade word-level reading and reading comprehension measures were entered simultaneously into the regression model. Results were significant for all four outcome measures; NWF, ORF, Word ID, and RC, accounting for approximately 20%, 17%, 41%, and 48% of the variance respectively. Likewise, when kindergarten Spanish reading-related measures were simultaneously entered into the regression model, results were significant across all four outcome measures, accounting for approximately 25%, 23%, 50%, and 54% of the variance respectively. Model 3 was conducted to determine how much the best English or Spanish descriptive measure scores (BLS) accounted for variance in the first grade criterion measures. Results were significant across all four outcome measures, accounting for approximately 19%, 15%, 42%, and 47% of the variance respectively. Model 4 was designed to identify the extent that the descriptive measures ELP and LA predicted the first grade reading criterion measures. Results were non-significant for NWF and ORF, and significant for Word ID and RC, accounting for approximately 1%, 3%, 16%, and 35% of the variance respectively.

Unique Variance—Model 5 was conducted to answer our third question, which was to determine whether there was a significant change in R^2 following the inclusion of the reading-related Spanish measures after the reading-related English measures had been entered first in the regression analysis. The results of this analysis indicated that the reading-related Spanish measures did not account for significant, unique variance over and above the reading-related English measures for any of the first grade reading criterion measures.

Model 6 was designed to answer our fourth question, which was to determine the extent that the 1descriptive kindergarten predictor measures accounted for significant, unique variance in first grade reading criterion measures over and above the reading-related kindergarten predictor measures. Results indicated that the descriptive measures did not significantly account for unique variance over and above the reading-related measures for any of the first grade word-level reading measures. However, the descriptive measures did account for unique variance in the Reading Comprehension measure over and above the reading-related measures.

Classification Analysis

One of the main purposes of the kindergarten descriptive and reading-related measures was to predict which Latino children would have future reading difficulties. Poor word-level reading status in first grade was determined by scoring at or below the 20th school district percentile on the NWF or ORF measure or at or below the 20th percentile on the WRMT-R word identification measure, using the test's published norms. Poor reading comprehension status in first grade was determined by scoring at or below the 20th percentile on the WRMT-R reading comprehension measure, also using the test's published norms.

Sensitivity, the ability of a kindergarten measure to identify those first grade students who scored at or below the 20th percentile on the reading measure, was calculated by dividing the number of true positives by the sum of the true positives and false negatives. Specificity, the ability of the kindergarten measure to identify students who were reading above the 20th percentile on a first grade reading measure, was calculated by dividing the number of true negatives by the sum of the true negatives and false positives.

Based on the results of the regression analysis, which indicated that the descriptive measures accounted for unique variance over and above the reading-related measures only for reading comprehension and not for any of the word-level reading criterion measures, we conducted discriminant function analyses with the LrID and BLS reading-related measures for the first grade reading criterion measures, and the letter ID, BLS reading-related measures plus the descriptive measures for the first grade reading comprehension measure. The discriminant function analysis for NWF was significant ($\Lambda = .54$, $\chi^2(4, N=48) = 27.35$, $p < .0001$), yielding 86% sensitivity and 90% specificity. The discriminant function analysis for ORF was not significant ($\Lambda = .86$, $\chi^2(4, N=48) = 6.69$, $p = .15$), yielding 67% sensitivity and 82% specificity. The discriminant function analysis for Word ID was significant ($\Lambda = .42$, $\chi^2(4, N=38) = 29.34$, $p < .0001$), yielding 75% sensitivity and 91% specificity. The discriminant function analysis for reading comprehension was significant ($\Lambda = .58$, $\chi^2(6, N=38) = 17.34$, $p < .01$), yielding 67% sensitivity and 93% specificity. The sensitivity and specificity results of the classification analyses are presented in Table 8.

Results Summary

Although the kindergarten reading-related English, reading-related Spanish, and reading-related BLS measures (Letter ID, PA, RAN, and SR) accounted for significant variance in all four of the first grade criterion measures, the reading-related Spanish measures did not account for significant variance over and above the reading-related English measures.

Kindergarten descriptive measures (ELP and LA) significantly accounted only for the first grade word ID and reading comprehension criterion measures, and those predictor measures only accounted for significant, unique variance over the reading-related measures for the first grade reading comprehension criterion measure. Across all four first grade reading measures, sensitivity was lower than specificity values indicating that these measures were better at correctly identifying children who did not have future reading difficulty than correctly identifying children who did have future reading difficulty.

DISCUSSION

Bivariate Correlations

Correlation coefficients were permitted considerable influence in this study. We used the initial correlation coefficient matrix to aid in sorting the multiple kindergarten measures into two groups; those that were significantly correlated with the first grade word-level reading measures and the reading comprehension measure, and those that were not significantly correlated. These multiple, pairwise correlations produced both expected and unexpected results.

Previous research has indicated that letter identification (LtID), phonological awareness (PA), rapid automatized naming (RAN), and sentence repetition (RA) significantly correlate with reading outcome measures. We also, as expected, found these variables to be moderately to strongly correlated with reading outcome measures for our sample of bilingual Latino children who were at risk for language impairment.

We did not include Spanish language proficiency measures in our regression analyses because of the nonsignificant correlations with our first grade reading criterion measures. Lindsay et al. (2003) and Manis et al., (2004) included measures of Spanish language along with several reading-related Spanish measures to predict English word identification and decoding, with their overall models yielding significant results. Yet, with so many other variables included, the extent to which their Spanish language measures uniquely contributed to their predictive models is not clear. Hammer et al., (2007) specifically isolated Spanish language proficiency as a predictor variable and found that Spanish language was negatively related to English word-level reading ($R^2 = -.24$). Their findings align with our non-significant correlation results.

Based on the results of their meta-analysis, August and Shanahan (2006), suggested that descriptive variables such as socioeconomic status and preschool history may be predictive of reading ability for culturally and linguistically diverse children. We found that neither of those measures were significantly correlated with first grade word-level reading ability or reading comprehension. This finding was unexpected.

The SES measure was calculated using a four-factor Hollingshead scale, which has evidence of reliability and validity (for a review see Cirino, Chin, Sevcik, Wolf, Lovett, & Morris, 2002). The correlation between the Hollingshead SES score and the word-level reading criterion measures were low, ranging from $r = .12$ to $.18$. The correlation between the Hollingshead SES score and the reading comprehension measure was higher ($r = .30$), yet

also non-significant. The mean weighted Hollingshead score was 14.3 with a standard deviation of 7.2, which was well within the criteria for low SES status even when scores were above 2 standard deviations from the mean. Thus, a lack of variance in the SES measure could account for the low correlation statistic. It is also possible that the relatively small number of participants impacted the significance of the correlations.

More variance was noted among the participants who had attended preschool. The number of years of preschool attendance ranged from 0 to 3, with 27 participants having never attended Head Start, 13 participants having attended 1 year of Head Start, 17 participants having attended 2 years of Head Start, and 4 participants having attended 3 years of Head Start. These preliminary results indicated that attendance in a Head Start preschool did not account for significant variance in first grade word-level reading ability or reading comprehension ($r = -.03$ to $.05$). There is mixed evidence on the short- and long-term benefits of Head Start (Barnett & Hustedt, 2005), and the results of this study align with other studies that have found limited effects, at least in the domain of reading outcomes.

Multiple Regression Analyses

First Grade Word-Level Reading—The combination of reading-related English, Spanish, and BLS predictor measures accounted for significant moderately-large variance in each of the first grade word-level reading criterion measures, with R^2 results similar to those reported in previous research (e.g., Lindsay et al., 2003; Oh et al., 2007; Pérez & Rinaldi, 2006). These findings provide additional evidence that LfID, PA, RAN, and SR in both English and Spanish are significantly predictive of English word-level reading for bilingual Latino children who were at risk for language impairment.

Even though our regression models with these combined predictors were significant, many of the predictor variables were not individually significant within the model. Unfortunately, because of the non-significant values for many of the individual variables, it is difficult to interpret each model more specifically. That is, it is difficult to determine which, if any of the predictor measures could be eliminated from each model while still maintaining the same or very similar significant R^2 results. In post-hoc forced step-wise regression analyses using word-level reading as the criterion measure, we found that no single predictor measure accounted for unique variance over and above any of the other predictor measures. A relatively small sample size can possibly explain this phenomenon, and further research should be conducted with a larger sample of bilingual Latino children who were at risk for language impairment to better understand the individual contribution of our predictor variables to later word-level reading.

Multicollinearity between the predictor variables is also often singled out as the reason for regression model significance combined with individual variable non-significance. When two or more predictor variables are highly correlated, they may also account for similar variance in an outcome measure, and although combined, these measures significantly account for variance, their overlapping nature makes it impossible to single out one variable as significant over the other. On the surface it appears that this was the case with our regression models. However, Variance Inflation Factor (VIF) analyses, which are used to identify multicollinearity, indicated that none of the predictor measures were highly collinear,

with VIF values never higher than 1.93. VIF values of 4 or 5 typically raise concern. We suspect that LfID, PA and RAN, while not particularly collinear, account for the same variance in word-level reading, and that little is gained from a predictive standpoint in using all three measures combined. This is not to say that each measure may not uniquely inform instruction, which is a topic beyond the scope of this particular study.

The reading-related Spanish measures did not significantly account for unique variance of first grade English word-level reading performance over and above the reading-related English measures. This finding has considerable implications for the assessment of risk for English word-level reading difficulty for bilingual kindergarten children. It appears that the inclusion of Spanish measures may not add meaningful predictive evidence of validity over and above English measures for English word-level reading at first grade. These findings align with findings from Pérez and Rinaldi (2006) who found that when using a combination of similar English and Spanish predictor measures to identify later English word-level reading the Spanish measures did not account for significantly more variance over English measures.

Although our English language proficiency and language ability measures did not account for significant variance in first grade nonsense word fluency or oral reading fluency, our results indicated that those measures did significantly account for a moderate amount of variance ($R^2 = .16$) in first grade word identification. Hammer et al. (2007) likewise found that English receptive language (a measure of English language proficiency) accounted for significant variance in later word identification. Our additional analyses however, indicated that the English language measures did not account for unique variance in first grade word-level reading over and above the reading-related measures, something that Hammer et al. (2007) did not explore.

This is the first study to explore the predictive relationship between language and oral reading fluency measures for bilingual Latino children who were at risk for language impairment, and although oral reading fluency is often used as a proxy measure for reading comprehension (e.g., DIBELS), our findings suggest that, at least for this sample of children, the relationship between kindergarten language measures and first grade oral reading fluency is tenuous. Our findings align with research from Mancilla-Martinez & Lesaux, (2010; 2011) and Mancilla-Martinez, et al., (2011), who reported weak relationships between language and word-level reading for ELL children. Our findings also align with those reported by Nakamoto, Lindsey, & Manis (2007) and Saez (2007), who found that that Spanish speaking children typically learn to decode a second language at an expected rate of acquisition, yet fall behind significantly in second language comprehension when compared to monolingual peers.

Our results suggest that Spanish and English language proficiency and language ability may not be predictive of English first grade word-level reading, including oral reading fluency, for bilingual Latino children who were at risk for language impairment. These findings are important to highlight because of the limited number of fluent Spanish-speaking examiners in the US who could reliably administer such measures in Spanish. If measures of English language proficiency are capable of accounting for the same or more variance in English

word-level reading as Spanish proficiency measures, then there is little point in using Spanish proficiency as a predictor of English word-level reading.

First Grade Reading Comprehension—To our knowledge, no research to date has explored the extent that English reading-related measures predict English reading comprehension in bilingual Latino children who were at risk for language impairment. Our results indicated that the English reading-related measures were a robust predictor of first grade English reading comprehension, accounting for 48% of the variance. Spanish reading-related measures accounted for an even greater amount of variance in first grade English reading comprehension ($R^2 = .54$). This finding was surprising given the weaker results reported by Lindsey et al., (2003) and Manis et al. (2004), who used similar Spanish reading-related measures to predict English reading comprehension, with R^2 results ranging from .15 to .21. In contrast to word-level reading outcomes, we found that English and Spanish reading-related measures were significantly predictive of reading comprehension outcomes.

These findings highlight the probability that word-level reading and reading comprehension pertain to two distinct constructs that are connected in a unilaterally dependent manner. That is, decoding can take place without comprehension, yet reading comprehension cannot take place without accurate decoding. The dependent relationship between decoding and comprehension only exists to the extent that it is necessary to accurately decode any form of language (e.g., speech, gestural, visual) before it can be converted to language.

Classification

The purpose of this longitudinal study was primarily predictive, and while seemingly counterintuitive, studies of this nature are concerned more with the degree that an assessment is predictive as opposed to *why* it is predictive. Thus, the preponderant verification of construct validity was evidenced inferentially in the measures' classification accuracy.

The addition of descriptive measures to traditional reading-related measures did not add to our predictive model for word-level reading in the regression analyses. Thus only the reading-related measures were used in the discriminant function analyses for the word-level criterion measures. We verified this decision by using a step-wise discriminant function analysis with the language measures included with the reading-related measures. With each word-level outcome measure, the language measures were removed from the model because they did not significantly add to the classification accuracy. Using reading-related measures, classification accuracy for NWF was excellent, yet for oral reading fluency and word identification, the classification accuracy was poor to moderate, with acceptable specificity but unacceptably low sensitivity. In accordance with the regression findings, we did include the language measures along with the reading-related measures in the discriminant function analyses to predict first grade reading comprehension. The results were similar to those found for ORF and Word ID, where specificity was excellent, but sensitivity was unacceptably low. Research that focuses on refined measures or alternative measures needs

to be conducted to improve the early and accurate classification of bilingual Latino children who have reading difficulties.

Limitations and Future Research

Interpretation and the inferential application of the results of this study should be tempered by several limitations. Our decision to include a relatively small number of participants from a specific subgroup of bilingual Latino children at risk for language impairment allowed us to obtain highly detailed information about each child's SES, language ability, and other contextual factors hypothesized to account for reading ability. Furthermore, by only including children who were pre-identified as at risk for language impairment, we increased our chances of having children with reading difficulty, which is important when conducting classification analyses. While the accumulation of this detailed information and the strategic participant sampling is a strength of the study, the smaller sample size with reduced participant variation and a relatively high number of predictor variables is also a limitation. Our sample of students potentially narrowed the range of many of the predictor measures used. SES was only represented at the lower end, which made it impossible to determine the extent that higher SES accounted for variance in the outcome reading measures. A larger sample of participants, or a more diverse sample of participants should be recruited for future research in this area. For example, the participants in this study were receiving reading instruction in English. Researchers may find different results if participants are learning to read in Spanish. Further, although the regression analysis assumptions of normality and linearity for the criterion reading measures were met, a larger sample would allow for more patent interpretation and generalization to the larger population and might allow for clearer interpretation of which specific predictor measures uniquely contribute to reading outcome measures.

Acknowledgments

This research was funded, in part, by grant R01DC007439 from the National Institute on Deafness and Other Communication Disorders (NIDCD). We wish to thank all of the interviewers and testers for their assistance with collecting the data and the school districts for allowing us access to the participants.

References

- August, DL., Shanahan, T. Developing literacy in second-language learners: Report of the National Literacy Panel on Language Minority Children and Youth. Mahwah, NJ: Erlbaum; 2006.
- Barnett WS, Hustedt JT. Head Start's Lasting Benefits. *Infants & Young Children*. 2005; 18(1):16–24.
- Bishop DVM, Snowling M. Developmental dyslexia and specific language impairment: Same or different? *Psychological Bulletin*. 2004; 130:858–886. [PubMed: 15535741]
- Bohman TM, Bedore LM, Peña ED, Mendez-Perez A, Gillam RB. What you hear and what you say: Language performance in early sequential Spanish English bilinguals. *International Journal of Bilingual Education and Bilingualism*. 2010; 13:325–344. [PubMed: 21731899]
- Catts HW. The narrow view of reading promotes a broad view of comprehension. *Language, Speech, and Hearing Services in Schools*. 2009; 40:178–183.
- Catts HW, Fey ME, Zhang X, Tomblin JB. Estimating the risk of future reading difficulties in kindergarten children: A research-based model and its clinical implication. *Language, Speech, and Hearing Services in Schools*. 2001; 32:38–50.
- Catts HW, Hogan TP. Language basis of reading disabilities and implications for early identification and remediation. *Reading Psychology*. 2003; 24:223–246.

- Catts HW, Petscher Y, Schatschneider C, Bridges M, Mendoza K. Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*. 2009; 42(2):163–176. [PubMed: 19098274]
- Cirino P, Chin C, Sevcik R, Wolf M, Lovett M, Morris R. Measuring socioeconomic status: Reliability and preliminary validity for different approaches. *Assessment*. 2002; 9:145–155. [PubMed: 12066829]
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951; 16:297–334.
- Cummins, J. Bilingualism and special education: Issues in assessment and pedagogy. Clevedon, UK: Multilingual Matters; 1984.
- Cummins J. Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*. 1979; 49(2):222–251.
- Cummins, J., Swain, M., Nakajima, K., Handscombe, J., Green, D., Tran, C. Research and application. Clevedon, UK: Multilingual Matters; 1984. Linguistic interdependence among Japanese and Vietnamese immigrant students.
- Durlak, JA. Successful prevention programs for children and adolescents. New York: Plenum; 1997.
- Elliott MN, McCaffrey D, Perlman J, Marshall GN, Hambarsoomians K. Use of expert ratings as sampling strata for a more cost-effective probability sample of a rare population. *Public Opinion Quarterly*. 2009; 73(1):56–73. [PubMed: 20936050]
- Gersten R, Dimino J. RTI (Response to Intervention): Rethinking special education for students with reading difficulties (yet again). *Reading Research Quarterly*. 2006; 41:99–108.
- Gough PB, Tunmer WE. Decoding, reading, and reading disability. *RASE: Remedial and Special Education*. 1986; 7:6–10.
- Good, RH., Jefferson, G. Contemporary perspectives on curriculum-based measurement validity. In: Shinn, MR., editor. *Advanced applications of curriculum-based measurement*. New York: Guilford; 1998. p. 61-88.
- Good, RH., Kaminski, RA., editors. *Dynamic indicators of basic early literacy skills*. 6. Eugene, OR: Institute for the Development of Educational Achievement; 2002. Retrieved March 2, 2010, from dibels.uoregon.edu
- Hammer CS, Lawrence FR, Miccio AW. Bilingual children's language abilities and early reading outcomes in Head Start and kindergarten. *Language, Speech, and Hearing Services in Schools*. 2007; 38:237–248.
- Hammer, C., Miccio, A., Rodriguez, B. Bilingual Language Acquisition and the Child Socialization Process. In: Goldstein, B., editor. *Bilingual Language Development and Disorders in Spanish-English Speakers*. Baltimore: Brookes; 2004.
- Hollingshead, AB. Unpublished manuscript. New Haven, CT: Yale University; 1975. Four-factor index of social status.
- Kaminski RA, Good RH. Toward a technology for assessing basic early literacy skills. *School Psychology Review*. 1996; 25:215–227.
- Kamhi AG. The case for the narrow view of reading. *Language, Speech, and Hearing Services in Schools*. 2009; 40:174–177.
- Kester ES, Peña ED. Language ability assessment of Spanish-English bilinguals: future directions. *Practical Assessment, Research & Evaluation*. 2002; 8(4)
- Koch G. The effect of non-sampling errors on measures of association in 2x2 contingency tables. *Journal of the American Statistical Association*. 1969; 64:852–863.
- Kohnert KJ, Bates E, Hernandez AE. Balancing bilinguals. *Journal of Speech, Language, and Hearing Research*. 1999; 42:1400–1413.
- Lindsey KA, Manis FR, Bailey CE. Prediction of first-grade reading in Spanish-speaking English-language learners. *Journal of Educational Psychology*. 2003; 95:482–494.
- Lyon GR. Reading disabilities: Why do some children have difficulty learning to read? What can be done about it? *Perspectives*. 2004; 29(2) Retrieved September 15, 2009 from <http://www.interdys.org>.
- Mancilla-Martinez J, Lesaux NK. The gap between Spanish-Speakers' word reading and word knowledge: A longitudinal study. *Child Development*. 2011; 82:1544–1560. [PubMed: 21848955]

- Mancilla-Martinez J, Kieffer MJ, Christodoulou JA, Biancarosa G, Snow C. Investigating English reading comprehension growth in adolescent language minority learners: Some insights from the Simple View. *Reading and Writing: An Interdisciplinary Journal*. 2011; 24:339–354.
- Mancilla-Martinez J, Lesaux NK. Predictors of reading comprehension for struggling readers: The case of Spanish-speaking language minority learners. *Journal of Educational Psychology*. 2010; 102:701–711. [PubMed: 20856691]
- Manis FR, Lindsey KA, Bailey CE. Development of reading in grades K2 in Spanish-speaking English-language learners. *Learning Disabilities Research and Practice*. 2004; 19:214–224.
- NAEP Assessment Tool. NAEP Questions. 2011. Retrieved February 22, 2012, from http://nationsreportcard.gov/reading_2012/
- Nakamoto J, Manis FR, Lindsey KA. A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading*. 2008; 12(4):351–371.
- Oh DM, Haager D, Windmueller M. A longitudinal study of predicting reading success for English-language learners from Kindergarten to grade one. *Multiple Voices for Ethnically Diverse Exceptional Children*. 2007; 10(1&2):107–124.
- Páez M, Rinaldi C. Predicting English word reading skills for Spanish-speaking students in first grade. *Topics in Language Disorders*. 2006; 26(4):338–350. [PubMed: 20396623]
- Pearson BZ, Fernandez MC, Oller DK. Measuring bilingual children's receptive vocabularies. *Child Development*. 1992; 63:1012–1221. [PubMed: 1505238]
- Pearson BZ, Fernandez MC, Oller DK. Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*. 1993; 43:93–120.
- Peña, ED., Bedore, LM., Gutiérrez-Clellen, VF., Iglesias, A., Goldstein, B. Bilingual English Spanish Oral Screener. (in development)
- Pena ED, Gillam RB, Bedore LM, Bohman TM. Risk for Poor Performance on a Language Screening Measure for Bilingual Preschoolers and Kindergarteners. *American Journal of Speech Language Pathology*. 2011; 20:302–314. 1058-0360_2011_1010-0020. [PubMed: 21821821]
- Peña, ED., Gutierrez-Clellan, V., Iglesias, A., Goldstein, B., Bedore, L. Bilingual English-Spanish Assessment: Screener. n.d. Manuscript in preparation
- Rosner J, Simon D. The Auditory Analysis Test: An initial report. *Journal of Learning Disabilities*. 1971; 4:384–392.
- Saez, Marina. Doctoral Dissertation. University of Pittsburgh; 2007. does L2 word decoding imply L2 meaning activation? Relationships among decoding, meaning identification, and L2 oral language proficiency in reading Spanish as a second language.
- Scarborough HS. Predicting future achievement of second graders with reading disabilities: Contributions of phonemic awareness, verbal memory, rapid naming and IQ. *Annals of Dyslexia*. 1998; 48:115–136.
- Shaywitz SE. Current concepts: Dyslexia. *The New England Journal of Medicine*. 1998; 338:307–312. [PubMed: 9445412]
- Shaywitz SE, Escobar MD, Shaywitz BA, Fletcher JM, Makuch R. Evidence that dyslexia may represent the lower tail of the normal distribution of reading ability. *New England Journal of Medicine*. 1992; 326:145–150. [PubMed: 1727544]
- Snowling MJ, Bishop DVM, Stothard SE. Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry and Allied Disciplines*. 2003; 41:587–600.
- Sutcliffe JP. A probability model for errors of classification: General considerations. *Psychometrika*. 1965; 30:73–96. [PubMed: 14293193]
- Swanson HL, Saez L, Gerber M, Lefstead J. Literacy and cognitive functioning in bilingual and nonbilingual children at or not at risk for reading disabilities. *Journal of Educational Psychology*. 2004; 96(1):3–18.
- Tindal, G., Marston, D., Deno, SL. Res Rep No 109. Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities; 1983. The reliability of direct and repeated measurement.
- Walker, BJ. Diagnostic teaching of reading: Techniques for instruction and assessment. 3. Upper Saddle River, NJ: Prentice Hall; 1996.

Woodcock, RW. Woodcock reading mastery test-revised. Circle Pines, MN: American Guidance Service; 1987.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1Variance (R^2) Estimates From a Review of Latino Predictive Reading Research

Study Predictor Variables	English Outcome Measures			
	Reading Comprehension	Word-Level Reading		
		Word I.D.	Decoding	ORF
Lindsay et al., (2003)				
Spanish Letter I.D., Phonological Awareness, Vocabulary, Sentence Repetition, RAN, & Print Concepts	.21	.35	.20	
Spanish Letter I.D.	.15	.25	.13	
Manis et al., (2004)				
Spanish Letter I.D., Phonological Awareness, RAN, & Expressive Language	.20	.26		
Spanish Letter I.D.	.16	.19		
Oh et al., (2007)				
Letter I.D.			.34	.26
Phonological Awareness			.08	.22
English Decoding			.27	.38
Páez & Rinaldi (2006)				
English Vocabulary, Sentence Repetition, & Phonological Awareness		.23		
English Vocabulary		.21		
Swanson et al., (2004)				
Spanish Vocabulary, Reading, Memory; English Memory		.33		
Hammer et al., (2007)				
English Receptive Language		.15	.21	
Spanish Receptive Language		-.24	-.23	

Note. Word I.D. = Word Identification; ORF = Oral Reading Fluency; Letter I.D. = Letter Identification; RAN = Rapid Automatized Naming.

Table 2

Participant Characteristics

Characteristic	<i>n</i>	Mean	SD
Male	29 (46%)		
Female	34 (54%)		
Hispanic/Latino	63 (100%)		
Age (months)	63	65.3	4.2
Mother's education	60	8.7	3.1
SES (Weighted Hollingshead Score)	63	14.3	7.2
Years in preschool	61	1.0	1.0
Years expressive Spanish	62	3.3	2.8
Years expressive English	62	0.2	1.1
Years expressive English/Spanish	62	2.4	2.7
Spanish Monolingual	29 (46%)		
English Monolingual	0 (0%)		
Bilingual Spanish Dominant	24 (38%)		
Bilingual English Dominant	10 (16%)		
Balanced Bilingual	24 (38%)		
Language Difficulty in Dominant Language	8 (13%)	.97*	1.47*

Note. Data collected one week prior to kindergarten;

* z-scores.

Table 3

Participant Characteristics of the Attrition Group and the Longitudinal Group, and First Grade Reading Performance for the Longitudinal Group.

Measure	<i>n</i>	Mean	<i>SD</i>	Median
Longitudinal Group				
Male	48	22 (46%)		
Female	48	26 (54%)		
Age (months)	48	65.1	3.8	64.5
SES (Weighted Hollingshead Score)	48	14.6	7.4	14.8
Years in preschool	48	1.0	1.0	1.0
Attrition Group				
Male	15	7 (47%)		
Female	15	8 (53%)		
Age (months)	15	66.2	5.3	66.0
SES (Weighted Hollingshead Score)	15	13.7	6.7	12.5
Years in preschool	15	0.8	1.0	0.0

Note. The Attrition Group is comprised of the students who left the school district prior to the end of first grade testing. The Longitudinal Group is comprised of those students who were assessed at kindergarten and then again at the end of first grade. Group differences were nonsignificant.

Table 4

Percentage of English Scores and Spanish Scores That Comprised the Best Language Score

BLS Measure	English	Spanish	Equivalent
PA-BLS	30%	49%	21%
RAN-BLS	97%	1.5%	1.5%
SR-BLS	40%	48%	12%

Note. PA-BLS = Phonological Awareness Best Language Score, RAN-BLS = Rapid Automatized Naming Best Language Score, SR-BLS = Sentence Repetition Best Language Score.

Table 5

Descriptive Statistics of Kindergarten Measures and First Grade Criterion Measures

Measure	<i>n</i>	Mean	<i>SD</i>	Median
Kindergarten				
Letter identification	63	28.5	7.9	32.0
Phonological awareness English	63	2.4	2.7	1.0
Phonological awareness Spanish	63	2.8	2.6	2.0
Phonological awareness BLS	63	3.3	2.7	3.0
Rapid Automatized Naming English	63	194.8	50.3	213
Rapid Automatized Naming Spanish	63	51.2	66.7	0.0
Rapid Automatized Naming BLS	63	193.4	50.1	212.0
Sentence repetition English	63	2.4	2.0	2.0
Sentence repetition Spanish	63	2.6	2.2	2.0
Sentence repetition BLS	63	3.6	2.1	3.0
SES: Hollingshead Score	63	14.3	7.2	14.5
Language Ability	63	.97	1.47	1.07
Number of years of preschool	61	1.0	1.0	1.0
First Grade				
Nonsense word fluency	48	53.9	17.9	51.0
Oral reading fluency	48	19.7	12.9	16.0
Word identification	38	39.2	12.1	40.5
Reading Comprehension	38	15.31	7.7	15.0

Note. BLS = Best Language Score

Table 6
Correlations and Intercorrelations Among Kindergarten Measures and First Grade Criterion Measures

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Letter Identification	-																		
2. Phonological Awareness English	.20	-																	
3. Phonological Awareness Spanish	.27*	.77*	-																
4. Phonological Awareness BLS	.28*	.92*	.93*	-															
5. Rapid Automatized Naming English	.54**	.35*	.39*	.42*	-														
6. Rapid Automatized Naming Spanish	.18	.33*	.29*	.33*	.32**	-													
7. Rapid Automatized Naming BLS	.52**	.25*	.31*	.34*	.97**	.35*	-												
8. Sentence Repetition English	.16	.53*	.36*	.50*	.27*	-.00	.18	-											
9. Sentence Repetition Spanish	.26*	.28*	.38*	.36*	.16	.45*	.17	.01	-										
10. Sentence Repetition BLS	.39**	.57*	.47*	.57*	.35**	.32*	.29*	.65**	.66**	-									
11. SES: Hollingshead Score	.16	.08	.04	.05	.10	-.13	.05	.26*	-.22	.12	-								
12. English Language Proficiency	.19	.31*	.19	.27*	.31*	-.02	.26*	.62**	-.20	.32**	.32*	-							
13. Spanish Language Proficiency	.21	.14	.14	.17	-.01	.28*	-.00	-.18	.42**	.12	-.06	-.13	-						
14. Language Ability	.28*	.24	.14	.20	.18	.07	.18	.40**	.07	.42**	.33	.61**	.44**	-					
15. Number of Years of Preschool	-.09	-.15	.02	-.08	-.08	-.03	-.05	-.04	-.07	-.10	.01	-.08	.02	-.01	-				
16. Nonsense Word Fluency (1 st Grade)	.40**	.31*	.33*	.32*	.42**	.40*	.41*	-.10	.18	.36*	.12	.17	.06	.21	-.03	-			
17. Oral Reading Fluency (1 st Grade)	.37**	.36*	.33*	.34*	.31*	.40*	.31*	-.01	.36*	.38**	.15	.10	.20	.12	.00	.81**	-		
18. Word Identification (1 st Grade)	.63**	.37*	.46*	.45*	.53**	.43*	.54*	.19	.26	.40*	.15	.31	.29	.44**	.05	.79**	.73**	-	
19. Reading Comprehension (1 st Grade)	.63**	.44*	.54*	.51*	.58**	.41*	.58*	.34*	.23	.49**	.30	.56**	.23	.53**	.11	.69**	.70**	.87**	-

* $p < .05$,

** $p < .01$;

BLS = best language score; SES = socioeconomic status.

Table 7

Multiple Regression Analysis Models

Model & kindergarten variable	<i>B</i>	Std error	Beta	<i>p</i>
<i>Model 1: English Only Reading-related Measures</i>				
First Grade Nonsense Word Fluency				
Letter Identification	0.48	0.34	0.23	.16
Phonological Awareness English	1.33	0.96	0.19	.17
Rapid Automatized Naming English	0.07	0.05	.23	.17
$R = .50, R^2 = .25, R^2_{adj} = .20, F(3, 44) = 4.84, p < .01$				
First Grade Oral Reading Fluency				
Letter Identification	0.41	0.25	0.27	.10
Phonological Awareness English	1.38	0.70	0.28	.06
Rapid Automatized Naming English	0.02	0.04	0.07	.69
$R = .47, R^2 = .22, R^2_{adj} = .17, F(3, 44) = 4.16, p = .01$				
First Grade Word Identification (n = 38)				
Letter Identification	0.64	0.22	0.48*	.01
Phonological Awareness English	0.94	0.67	0.19	.17
Rapid Automatized Naming English	0.04	0.04	0.17	.33
$R = .68, R^2 = .46, R^2_{adj} = .41, F(3, 34) = 9.65, < .001$				
First Grade Reading Comprehension (n = 38)				
Letter Identification	0.34	0.13	0.41*	.01
Phonological Awareness English	0.61	0.42	0.19	.16
Rapid Automatized Naming English	0.03	0.02	0.21	.18
Sentence Repetition English	0.77	0.55	0.18	.17
$R = .73, R^2 = .53, R^2_{adj} = .48, F(4, 33) = 9.47, < .001$				
<i>Model 2: Spanish Only Reading-related Measures</i>				
First Grade Nonsense Word Fluency				
Letter Identification	0.60	0.29	0.28*	.04
Phonological Awareness Spanish	1.30	0.88	0.20	.15
Rapid Automatized Naming Spanish	0.09	0.04	0.30*	.03
$R = .54, R^2 = .29, R^2_{adj} = .25, F(3, 44) = 6.10, p < .01$				
First Grade Oral Reading Fluency				
Letter Identification	0.38	.021	0.25	.08
Phonological Awareness Spanish	1.00	0.64	0.21	.13
Rapid Automatized Naming Spanish	0.07	0.03	0.31*	.03
$R = .53, R^2 = .28, R^2_{adj} = .23, F(3, 44) = 5.79, p < .01$				
First Grade Word Identification (n = 38)				
Letter Identification	0.67	0.17	0.50**	.0001
Phonological Awareness Spanish	1.08	0.55	0.25	.06
Rapid Automatized Naming Spanish	0.05	0.02	0.26*	.04
$R = .74, R^2 = .54, R^2_{adj} = .50, F(3, 34) = 13.34, < .0001$				
First Grade Reading Comprehension (n = 38)				

Model & kindergarten variable	<i>B</i>	Std error	Beta	<i>p</i>
Letter Identification	0.43	0.10	0.51**	.0001
Phonological Awareness Spanish	1.00	0.34	0.36*	.02
Rapid Automatized Naming Spanish	0.03	0.02	0.28	.01
Sentence Repetition Spanish	-0.50	0.42	-0.15	.24
$R = .77, R^2 = .59, R^2_{adj} = .54, F(4, 33) = 11.80, < .0001$				
Model 3: Reading-related Measures Best Language Scores				
First Grade Nonsense Word Fluency				
Letter Identification	0.47	0.34	0.21	.18
Phonological Awareness BLS	1.19	0.94	0.18	.21
Rapid Automatized Naming BLS	0.08	0.05	0.23	.15
$R = .49, R^2 = .24, R^2_{adj} = .19, F(3, 44) = 4.59, p < .01$				
First Grade Oral Reading Fluency				
Letter Identification	0.37	.025	0.24	.15
Phonological Awareness BLS	1.12	0.70	0.23	.11
Rapid Automatized Naming BLS	0.03	0.04	0.11	.52
$R = .45, R^2 = .20, R^2_{adj} = .15, F(3, 44) = 3.69, p < .05$				
First Grade Word Identification (n = 38)				
Letter Identification	0.62	0.22	0.46*	.01
Phonological Awareness BLS	0.98	0.65	0.21	.14
Rapid Automatized Naming BLS	0.03	0.04	0.15	.38
$R = .68, R^2 = .47, R^2_{adj} = .42, F(3, 34) = 9.97, < .0001$				
First Grade Reading Comprehension (n = 38)				
Letter Identification	0.30	0.14	0.36*	.04
Phonological Awareness BLS	0.66	0.42	0.23	.12
Rapid Automatized Naming BLS	0.03	0.02	0.20	.23
Sentence Repetition BLS	0.49	0.50	0.14	.33
$R = .73, R^2 = .53, R^2_{adj} = .47, F(4, 33) = 9.15, p < .0001$				
Model 4: Descriptive Measures				
First Grade Nonsense Word Fluency				
English Language Proficiency	0.74	1.86	0.07	.70
Language Ability	2.20	2.32	0.17	.35
$R = .22, R^2 = .05, R^2_{adj} = .01, F(2, 45) = 1.12, p = .34$				
First Grade Oral Reading Fluency				
English Language Proficiency	0.35	1.37	0.05	.80
Language Ability	0.84	1.70	0.90	.63
$R = .12, R^2 = .02, R^2_{adj} = -.03, F(2, 45) = .34, p = .72$				
First Grade Word Identification (n = 38)				
English Language Proficiency	0.77	1.37	0.10	.58
Language Ability	3.65	1.69	0.39*	.04
$R = .45, R^2 = .20, R^2_{adj} = -.16, F(2, 35) = 4.46, p < .05$				
First Grade Reading Comprehension (n = 38)				

Model & kindergarten variable	<i>B</i>	Std error	Beta	<i>p</i>
English Language Proficiency	1.89	0.76	0.39*	.02
Language Ability	1.88	0.94	0.32*	.05

$R = .62$, $R^2 = .39$, $R^2_{adj} = -.35$, $F(2, 35) = 11.06$, $p < .0001$

Model 5: Spanish Reading-related Measures Over and Above English Reading-related Measures

First Grade Nonsense Word Fluency

$R = .50$, $R^2 = .25$, $R^2_{adj} = .20$, $F(3, 44) = 4.84$, $p < .01$

R^2 change = .07, $F(5, 42) = 3.98$, $p = .12$

First Grade Oral Reading Fluency

$R = .47$, $R^2 = .22$, $R^2_{adj} = .17$, $F(3, 44) = 4.16$, $p = .01$

R^2 change = .09, $F(5, 42) = 3.79$, $p = .08$

First Grade Word Identification (n = 38)

$R = .68$, $R^2 = .46$, $R^2_{adj} = .41$, $F(3, 34) = 9.65$, $p < .0001$

R^2 change = .08, $F(5, 32) = 7.61$, $p = .07$

First Grade Reading Comprehension (n = 38)

$R = .73$, $R^2 = .53$, $R^2_{adj} = .48$, $F(4, 33) = 9.47$, $p < .0001$

R^2 change = .09, $F(7, 30) = 7.09$, $p = .09$

Model 6: Descriptive Measures Over and Above BLS Reading-related Measures

First Grade Nonsense Word Fluency

$R = .49$, $R^2 = .24$, $R^2_{adj} = .19$, $F(3, 44) = 4.59$, $p < .01$

R^2 change = .01, $F(5, 42) = 2.73$, $p = .83$

First Grade Oral Reading Fluency

$R = .45$, $R^2 = .20$, $R^2_{adj} = .15$, $F(3, 44) = 3.69$, $p < .05$

R^2 change = .01, $F(5, 42) = 2.24$, $p = .78$

First Grade Word Identification (n = 38)

$R = .69$, $R^2 = .47$, $R^2_{adj} = .41$, $F(4, 33) = 7.32$, $p < .0001$

R^2 change = .02, $F(6, 31) = 5.01$, $p = .52$

First Grade Reading Comprehension (n = 38)

$R = .73$, $R^2 = .53$, $R^2_{adj} = .47$, $F(4, 33) = 9.15$, $p < .0001$

R^2 change = .09, $F(6, 31) = 8.23$, $p < .05$

Sensitivity and Specificity Classifications Between Kindergarten Predictor Measures and First Grade Reading Criterion Measures

Table 8

<u>First Grade Reading Criterion Measures</u>							
<u>NWF</u>		<u>ORF</u>		<u>Word ID</u>		<u>Reading Comprehension</u>	
Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Kindergarten Predictor Measures		.86	.90	.67	.82	.75	.91
						.67	.93

Note. NWF = first grade nonsense word fluency; ORF = first grade oral reading fluency; Word ID = first grade word identification.