

RESEARCH ARTICLE

Spanish-speaking Preschoolers' Conceptual Vocabulary Knowledge: Towards More Comprehensive Assessment

Jeannette Mancilla-Martinez

Vanderbilt University

Christa Mulker Greenfader

California State University, Fullerton

Wendy Ochoa

University of California, Irvine

This study examined Spanish-English preschoolers' ($n = 32$) vocabulary performance when using traditional English-only measures compared to measures that utilize conceptual scoring (i.e., vocabulary knowledge in terms of known concepts independent of whether the label for the concept is known in either Spanish or English) *designed for and normed on* Spanish-English bilinguals. Children's performance at the item level on the conceptually-scored measures was also examined. In English, receptive and expressive average scores were in the below-average ranges. However, on the conceptually-scored vocabulary measures, the average scores were within the average range, receptively and expressively. Examination of children's performance at the item level suggests there may be differences by language in bilingual children's receptive and expressive vocabulary performance. The findings underscore the value of utilizing conceptually-scored vocabulary measures, suggesting that their use may have potential to distinguish language difference from language disorder among young bilingual learners.

Keywords: bilingualism, vocabulary knowledge, vocabulary assessment, conceptual vocabulary, Spanish-speakers

Early vocabulary skills are key predictors of later achievement outcomes, among native and non-native English-speaking populations alike (Dickinson, 2011; Mancilla-Martinez & Lesaux, 2017; Morgan, Farkas, Hillemeier, Hammer, & Maczuga, 2015). But compared to their native English-speaking peers, children from Spanish-speaking homes in the United States – hereon referred to as Spanish-English bilinguals – disproportionately evidence English vocabulary skills, beginning prior to formal school entry, that raise concerns about their future academic prospects (García & Frede, 2010; García, Jensen, & Scribner, 2009; National Task Force on Early Childhood Education for Hispanics, 2007). Additionally, low-income status hinders child development,

including academic achievement (Brooks-Gunn & Duncan, 1997). Given that Spanish-English bilinguals in the U.S. disproportionately come from lower-income homes compared to their monolingual peers (López & Velasco, 2011), the recurrent findings of this group's lower vocabulary scores compared to their monolingual-English, higher-income peers are not surprising. However, the historical lack of reliable and valid vocabulary assessments designed for and normed on bilingual children continues to represent one of the greatest challenges in gaining insight into this population's vocabulary performance and development, and, by extension, on differentiating a language difference from a language disorder (Bedore & Peña, 2008; Klingner, 2014; Ortiz & Artiles, 2010; Paradis, Genesee, & Crago, 2011).

This study was designed with the goal of moving the field forward in gaining a better understanding of Spanish-English bilingual preschoolers' vocabulary knowledge. We sought to determine the utility of assessing receptive and expressive vocabulary skills using *standardized* measures designed for and normed on Spanish-English bilinguals that utilize conceptual scoring, compared to using receptive and expressive vocabulary measures designed for and normed on English monolinguals. Additionally, we examined the proportion of children potentially deemed "at-risk" for low academic achievement when English-only compared to conceptually-scored vocabulary measures are used. Finally, to gain further insight into Spanish-English bilingual students' patterns of performance in each language by domain (i.e., receptive and expressive vocabulary), we examined their performance at the item level on the conceptually-scored vocabulary measures. Results of this study have the potential to pave the way for a more inclusive notion of what it means to evaluate – with implications on how to support – the vocabulary development of young bilingual children by focusing on the use of standardized vocabulary measures that use conceptual scoring to tap understanding of concepts regardless of whether the label is known in Spanish or English.

Misrepresentation and Bilingual Assessment

As for all learners, misrepresentation of bilinguals as having a learning disability is consequential. For bilinguals in particular, English proficiency emerges as a key factor that contributes to misidentification. The "Exclusionary Clause" under the Individuals with Disabilities Education Act (IDEA) states that eligibility determination must not be primarily related to limited English proficiency (Individuals with Disabilities Education Act, 1997). Yet, many bilinguals, especially young children, are in the process of developing *both* their native language and English skills. If limited English proficiency should be excluded as a primary eligibility determination, bilingual children may be at risk for misidentification and this appears to be the case. For example, Morgan and colleagues (2012) report that, compared to their monolingual peers, bilingual children tend to be referred to early intervention and special education at lower rates. The need to ensure that bilingual children are properly identified cannot be overstated, and this process must begin prior to formal school entry (Reichow, Boyd, Barton, & Odom, 2016). The reality, however, is that the majority of bilingual children continue to be assessed only in English despite calls to ensure that both languages are accounted for (National Academies of Sciences, Engineering, and Medicine, 2017).

Like any single assessment, vocabulary measures should not be the only tool used to screen for potential academic difficulty in later years. However, speech pathologists and researchers commonly incorporate vocabulary measures into their battery of assessments when

evaluating the language development of young English monolinguals (Beitchman, Nair, Clegg, & Patel, 1986; Gray, 2004; Sheng & McGregor, 2010a; Tomblin et al., 1997; Walker, Greenwood, Hart, & Carta, 1994). Although less explored, language measures that include a vocabulary component have also been found to predict academic difficulties among young Spanish-English bilinguals (Lugo-Neris, Peña, Bedore, & Gilliam, 2015; Mancilla-Martinez & Lesaux, 2011, 2017; Peña, Bedore, & Kester, 2015a, 2015b; Sheng, Peña, Bedore, & Fiestas, 2012). As an example, Morgan and colleagues (2015) found that 2-year-olds with expressive vocabulary delays were at a high risk for later speech-language services, but limited English proficient children from low-income homes tended to receive speech-language services at significantly lower rates compared to their English-proficient peers. The authors hypothesize that an insufficient understanding by speech-language pathologists on how to differentiate a language difference from a language disorder might be at play.

Use of conceptually-scored vocabulary assessments – hereon defined as assessments that reflect vocabulary knowledge in terms of known concepts by crediting the item as correct independent of whether the label for the concept is known in either Spanish or English – may represent one mechanism by which to more accurately assess bilingual students' vocabulary skills, potentially minimizing misidentification and maximizing the match between their educational needs and the supports provided.

Dual Vocabulary Development: The Case for Conceptual Vocabulary

Bilingual children's vocabulary knowledge has been shown to be distributed across languages, severely limiting the extent to which monolingual vocabulary measures can provide accurate insight into performance and development in this domain (e.g., Bedore, Peña, Garcia, Cortez, 2005; Mancilla-Martinez & Vagh, 2013; Oller & Eilers, 2002; Pearson, Fernández, & Oller, 1994). Indeed, Grosjean's (1982, 1989) seminal work has long cautioned researchers against the monolingual view on bilingualism (i.e., that the bilingual is, or should be, two monolinguals in one person and should thus have "equal" proficiency in both languages). Yet, there continues to be a reliance on measures designed for monolinguals, and bilinguals' performance is appraised based on monolingual standards. The Revised Hierarchical Model (Kroll & Stewart, 1994), similar to Cummins' (1979) well-known linguistic interdependence hypothesis, posits that bilinguals represent their two languages in separate lexicons, but that they do so with one conceptual (language-free) system subserving both languages. Cummins' work further underscores that proficiency in one language facilitates proficiency in another, particularly if a certain threshold of proficiency is reached in the native language. For instance, if a child learns the Spanish word *semilla*, s/he has a general conceptual understanding of the word, facilitating the acquisition of the English equivalent (*seed*).

An additional consideration when attending to bilingual children's vocabulary development is that, unlike their monolingual peers, their language environments are fundamentally different; bilinguals receive input in two languages, though the amount of exposure to each language and children's own use of each language varies widely. Further, the setting (e.g., home vs. school) can play a role in priming bilinguals toward a more monolingual or bilingual language mode (Bialystok, Luk, Peets, & Yang, 2010; Grosjean, 1982, 1989, 2008), underscoring the need to attend to both languages when assessing bilingual children's vocabulary. Yet, standardized English language assessments in the U.S. rarely attend to whether

linguistically diverse speakers are included in the norming samples (Luk & Christodoulou, 2016).

It is beyond the scope of this study to investigate the extent to which conceptually-scored vocabulary measures better predict who goes on to experience academic challenges, but examining bilingual children's English-only compared to conceptually-scored vocabulary performance via standardized measures with established psychometric properties has the potential to provide *preliminary* insight into this complex issue. Predictive validity evidence of language and literacy assessments used with young bilinguals remain scarce (Bandel, Atkins-Burnett, Castro, Wulsin, & Putman, 2012), but Spanish-English bilinguals typically exhibit vocabulary scores within or above the average range when conceptual scoring is utilized. However, extant work has relied on the adaptation of monolingual measures to proxy conceptual vocabulary knowledge (Gross, Buac, & Kaushanskaya, 2014; Mancilla-Martinez, Pan, & Vagh, 2011; Sheng et al., 2012; Mancilla-Martinez & Vagh, 2013). An open question is whether use of conceptually-scored standardized vocabulary measures *designed for and normed on* Spanish-English bilinguals would mirror these findings. If so, these standardized measures could represent a promising mechanism for informing determinations, both in research and practice, regarding whether bilingual learners might require further language supports stemming from a genuine disability or whether they are simply still in the process of developing their language skills (Bedore & Peña, 2008).

Vocabulary Assessments that Utilize Conceptual Scoring

While the call for vocabulary assessments designed for and normed on bilinguals is not new, historically, there has been a lack of such measures. In turn, a common practice in research and classrooms throughout the U.S. has been to *adapt* monolingual measures to index bilingual children's vocabulary knowledge (e.g., Pearson et al., 1993, 1995; Pearson & Fernandez, 1994). The most common adjustments to accommodate both languages for bilinguals are total and conceptual vocabulary scores. A total vocabulary score is derived by summing all the words in Spanish and English (e.g., children receive double credit for knowing both *puerta* and *door*) while a conceptual vocabulary score is derived by summing all the words in Spanish and English but then subtracting translation equivalent items (e.g., children do not receive double credit for knowing both *puerta* and *door*). Research to date generally demonstrates lower performance when bilinguals are tested in single-language vocabulary but comparable performance to their monolingual peers when adjusted total or conceptual vocabulary scores are used (Core, Hoff, Rumiche, & Señor, 2013; Gross et al., 2014; Mancilla-Martinez et al., 2011; Mancilla-Martinez & Vagh, 2013; Patterson, 1998; Pearson et al., 1994; Pearson et al., 1993; Peña, Bedore, & Kester, 2015a). While debate remains concerning the extent to which total or conceptual vocabulary is a better index of bilingual children's overall vocabulary knowledge, work focused on psycholinguistic aspects of bilingualism points to the existence of one combined lexico-semantic store that is similar to monolinguals (Dijkstra & Van Heuven, 2002; Kroll & Sunderman, 2003), providing theoretical support for the use of conceptually-scored measures. Yet, use of adapted measures proves challenging for numerous reasons, not the least of which is the lack of norms associated with the adaptation of monolingual measures.

We underscore that conceptually-scored vocabulary measurement is distinct from the concept of code-switching, though there are similarities. Code-switching refers to a natural,

distinctive occurrence among bilinguals, in which speakers shift between languages, at the word, phrase, sentence, or utterance levels (García, 2009; Grosjean, 1989; Wei & Martin, 2009). In other words, instead of strictly using one language in a single speech act (e.g., conversation), code-switching is characterized by the juxtaposition of two different linguistic systems (Gumperz, 1982; Poplack, 1980). While both code-switching and conceptual vocabulary scoring depend on bilingual children's linguistic reservoir, code-switching is used to compensate lexical gaps in a language that is less developed than the dominant language during conversations to accomplish their communicative goals (Genesee, Paradis, & Crago, 2004), while conceptual scoring is a sensitive way to accurately capture their lexical knowledge. As such, conceptual vocabulary scoring refers to the acceptance of responses in *either* Spanish or English and does not entail students' *mixing* of two linguistic systems (e.g., Spanish and English) during a conversational interaction. The focus of this study is on the assessment of vocabulary via use of standardized measures that apply conceptual scoring among children negotiating both Spanish and English.

To our knowledge, the co-normed vocabulary assessments, *Receptive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition (ROWPVT-4: SBE)*; Martin, 2013) and *Expressive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition (EOWPVT-4: SBE)*; Martin, 2013) represent the only set of standardized vocabulary assessments currently available to assess Spanish-English bilingual children's separate receptive and expressive vocabulary using conceptual scoring. These measures were designed for and normed on Spanish-English bilinguals, beginning at age two and extending through to adulthood, who speak Spanish and English with varying levels of proficiency.

Receptive-Expressive Vocabulary Patterns of Performance

As with English monolinguals (Benedict, 1979), Spanish-English bilinguals generally appear to exhibit better receptive than expressive skills (Gibson, Oller, Jarmulowicz, & Ethington, 2012; Gross et al., 2014; Miccio, Tabors, Páez, Hammer, & Wagstaff, 2005; Oller et al., 2007a; Oller, Pearson, & Cobo-Lewis, 2007b; Windsor & Kohnert, 2004). However, findings on the nature of the receptive-expressive gap are mixed for Spanish-English bilinguals. Some work reports that young Spanish-English bilinguals exhibit a receptive-expressive vocabulary gap in both languages (i.e., that receptive skills are higher than expressive skills), but that the gap is larger in Spanish (Gibson et al., 2012; Gross et al., 2014; Miccio et al., 2005; Oller et al., 2007a, 2007b; Swanson, Rosston, Gerber, & Solari, 2008). Yet, others have found the opposite pattern, even among children with and without language impairments (Gibson, Peña, & Bedore, 2014a, 2014b). Furthermore, other work reveals that young Spanish-English bilinguals indeed evidence higher receptive than expressive vocabulary in Spanish, but higher expressive than receptive vocabulary in English (Mancilla-Martinez & Vagh, 2013). More work is needed in this area, and examining preschoolers' receptive and expressive vocabulary performance on Spanish-English conceptually-scored measures at the item level could contribute to the open question of receptive-expressive vocabulary gaps. More practically, examining students' performance at the item level can help inform instructional efforts aimed at bolstering this population's vocabulary skills by providing guidance into the types of words bilinguals appear to already know or need support to develop.

Among others, considerations when evaluating bilingual children's receptive and expressive performance at the item level include context (e.g., home or school), cognates (i.e., words that share common or exact roots in both languages such as "elephant" and "elefante"), and categorization (e.g., broader semantic categories of words known such as "animals," "foods," etc.). Bialystok and colleagues (2010) developed their own system to classify known English words as either "home" or "school." They identified a word gap between bilingual and monolingual-English children in their English "home" vocabulary, such that monolingual children ostensibly knew more English "home" words than the bilingual children. However, no differences between the two groups were found in their knowledge of English "school" words. However, the authors acknowledge that categorizing words as being primarily from the "home" or "school" context is not absolute, and they thus call for additional research that uses more detailed categories. Another type of category may be cognates and studies have shown that bilingual children tend to perform better on cognate than non-cognate words (Kroll, Gerfen, & Dussias, 2008). Yet, an important consideration is that recognition of cognates does not appear to be automatic, even for older students, and instead appears to require explicit instruction (Nagy, García, Durgunoglu, & Hancin-Bhatt, 1993; Hancin-Bhatt & Nagy, 1994; Moss, 1992). The MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 2007) semantic categories offer a potentially useful framework by which to examine Spanish-English bilingual children's vocabulary performance at the item level. The CDI inventories are used to gauge the number of words children typically learn first in multiple languages, including Spanish and English, and the words are grouped in semantic categories (e.g., animals, outside things, people). Notably, previous research has shown that the English and Spanish CDI inventories can be used with preschool-aged Spanish-English bilingual children (Mancilla-Martinez et al., 2011). Exploring the utility of the CDI semantic categories for understanding Spanish-English bilingual children's conceptually-scored vocabulary knowledge is an area ripe for research as little is known about the types of words bilingual children know, limiting efforts to promote their oral language skills (García & Frede, 2010).

The Present Study

This study examines Spanish-English bilingual children's receptive and expressive vocabulary utilizing English-only and Spanish-English standardized measures that utilize conceptual scoring. Results of this study have the potential to advance our understanding of this population's *overall* vocabulary knowledge and to provide preliminary insight into the extent to which measures that utilize conceptual scoring may represent a promising mechanism for minimizing misrepresentation of bilinguals in special education. The results likewise have instructional implications for supporting this group's vocabulary development. We address two key research questions:

- (1) How do Spanish-English bilingual preschoolers' receptive and expressive vocabulary skills compare when English-only vs. Spanish-English standardized measures that utilize conceptual scoring are used? Relatedly, does, and if so to what extent, the proportion of bilingual children deemed "at risk" for low academic achievement change depending on the measures used?

- (2) What are Spanish-English bilingual children's patterns of performance at the item level on the Spanish-English receptive and expressive vocabulary measures that utilize conceptual scoring?

In line with previous research that has adapted monolingual measures to proxy Spanish-English bilingual children's conceptual vocabulary knowledge, we hypothesized that our sample of Spanish-English bilingual preschoolers would evidence higher scores on the receptive and expressive standardized vocabulary measures that use conceptual scoring compared to the English-only measures. Thus, we expected that the percentage of preschoolers identified as "at-risk" would be attenuated via use of the conceptually-scored measures. Further, and also in line with previous work, we hypothesized that children would evidence stronger receptive than expressive performance on the conceptually-scored measures. However, given the scarcity of work that has examined this population's vocabulary performance on conceptually-scored measures at the item level, we did not have a solid basis for making hypotheses about preschoolers' item level performance.

METHOD

Participants

The sample consists of 32 Spanish-English bilingual children, including one set of twins, enrolled in a state-funded preschool or transitional kindergarten program in Southern California (female $n = 20$; male $n = 12$). On average, children were 54.7 months old and the average child had been enrolled in preschool for 12.7 months (see Table 1). Instruction was provided in English to all children. Children were included in the study if parents reported Spanish use in the home, solely or with English.

Parent interviews, prepared in Spanish and English, were conducted with all of the families, except one due to scheduling conflicts (we obtained key demographic information on this family from a screening form at study entry). Most interviews (81%) were conducted with mothers. All children were reported by their families to be U.S.-born and Latino (see Table 1). Nearly all parents were foreign-born, and parents reported fewer than 10 years of education, on average. Following Brooks-Gunn, Duncan, and Britto (1999), we identified four income-to-needs ratio categories (see Appendix A), as follows: deep poverty (income-to-needs ratio less than .50), poverty (income-to-needs ratio greater than or equal to .50, but less than 1.0), near poverty (income-to-needs ratio greater than or equal to 1.0, but less than 1.5), and low income (income-to-needs ratio greater than or equal to 1.5, but less than 2.0). On average, families had an income-to-needs ratio at the poverty level (.90).

Procedure

Participants were recruited from a large collaborating school district in Southern California. The preschool Director disseminated flyers for parent recruitment to all classrooms, and teachers distributed the flyers to all children's parents. Parents who agreed to participate returned the screening form, permitting trained research assistants (RAs) to directly contact families and

invite them into the study. All families who reported using Spanish at home, solely or with English, were invited to participate. RAs gained formal consent and conducted telephone or in-person interviews (per parent preference) and administered the direct child vocabulary assessments. All RAs were Spanish-English bilingual, biliterate undergraduate and graduate students. Parent interviews provided relevant demographic data (see Table 1), as well as detailed information on home language use (described below). Children's receptive and expressive vocabulary skills were assessed using direct standardized assessments in English (using monolingual English measures) and conceptually (using Spanish-English conceptual measures). Testing was conducted over two days, with RAs administering one set of vocabulary assessments (i.e., English-only or Spanish-English conceptual) per day; the order of administration was counterbalanced. English-only assessments took about 35 minutes and Spanish-English conceptual assessments about 45 minutes. For the Spanish-English conceptual measures, all testing began in Spanish per parent report of Spanish as the dominant home language.

Parent Interview

With two exceptions, all parents completed the roughly 30-minute interview in Spanish. The demographic information is summarized above (see Table 1). We also collected information from parents on home language use (see Table 2). Parents were asked to report on the language(s) that the mother, father, other adults and other children in the home used when speaking to their child, and on the language(s) their child used when speaking to each of these members, as applicable. The following scale was used: 1 = Only Spanish, 2 = Mostly Spanish, 3 = Equal Amounts of English and Spanish, 4 = Mostly English, and 5 = Only English.

Measures

English Receptive Vocabulary. English-only receptive vocabulary was assessed with the *Receptive One Word Picture Vocabulary Test- 4 (ROWPVT-4;* Martin & Brownell, 2011). Children identified pictured objects, actions, and concepts that were ordered by increasing difficulty, beginning with the easiest concepts and ending with less frequently encountered items. The task was discontinued when the child made 6 errors within 8 consecutive responses. The publisher reports the median internal consistency reliability coefficient as .97.

Conceptual Receptive Vocabulary. Spanish-English conceptually-scored receptive vocabulary was assessed with the *Receptive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition (ROWPVT-4: SBE;* Martin, 2013). Children identified pictured objects, actions, and concepts that were ordered by increasing difficulty, arranged in a developmental sequence specific to the bilingual normative sample using Classical Test Theory and Item Response Theory, beginning with the easiest concepts and ending with less frequently encountered items. The task was discontinued when the child made 4 errors within 6 consecutive responses. As a Spanish-English conceptually-scored measure, children were presented with the target item first in Spanish (per parent report of Spanish being the dominant home language). If a child indicated the incorrect answer or did not respond at all, the examiner repeated the question in English, allowing for the assessment of receptive knowledge in *either* language. If the child missed the

item once both languages were targeted, the item was scored as incorrect. The publisher reports the median internal consistency reliability coefficient as .95.

TABLE 1
Demographic characteristics for Spanish-English bilingual children and their families

Demographic Characteristics	Average Values & Percentages	
Child age in months, average (n = 32)	54.7 (5.6)	
Child age at preschool entry in months, average (n = 26)	42.0 (5.6)	
Child's country of birth (n = 32)	100%	
United States	100%	
Child's ethnicity (n = 32)	100%	
Latino	100%	
Parents' country of birth	Mother (n=32)	Father (n=29)
United States	3%	10%
Mexico	94%	86%
Peru	3%	0%
Guatemala	0%	3%
Parents' age at immigration	Mother (n=29)	Father (n=25)
Under 10 years old	17%	16%
10-18 years old	31%	32%
Over 18 years old	52%	52%
Parents' Education	Mother (n=30)	Father (n=28)
Total years, average	9.6 (4.2)	9.9 (3.3)
Family economic condition		
Deep poverty (<.50)	19%	
Poverty (.50 to <1.0)	39%	
Near Poverty (1.0 to <1.5)	26%	
Low-income (1.5 to <2.0)	13%	
Middle-income (\geq 2.0)	3%	
Average	0.90 (0.5)	

Note. Standard deviation in parentheses. Some questions were not answered by interviewees, as reflected by the sample sizes noted.

TABLE 2
Patterns of home language use for the Spanish-English bilingual students

Pattern	Mostly Spanish		Equal Amounts		Mostly English		<i>r</i>
	%	<i>N</i>	%	<i>N</i>	%	<i>N</i>	
Input by mother to child (n = 31)	90	28	6	2	3	1	
Output by child to mother (n = 31)	58	18	23	7	19	6	.4*
Input by father to child (n = 29)	72	21	7	2	21	6	
Output by child to father (n = 29)	52	15	21	6	28	8	.7***
Input by other adults to child (n = 29)	93	27	3	1	7	2	
Output by child to other adults (n = 30)	57	17	13	4	30	9	.4~
Input by other children to child (n = 29)	41	12	17	5	41	12	
Output by child to other children (n = 29)	38	11	21	6	41	12	1.0***
<i>Overall input: family to child (n = 31)</i>	<i>71</i>	<i>22</i>	<i>16</i>	<i>5</i>	<i>13</i>	<i>4</i>	
<i>Overall output: child to family (n = 31)</i>	<i>45</i>	<i>14</i>	<i>10</i>	<i>3</i>	<i>45</i>	<i>14</i>	<i>.7***</i>

Note. ~ = $p = .05$, * $p < .05$, *** $p < .001$

English Expressive Vocabulary. English-only expressive vocabulary was assessed with the *Expressive One Word Picture Vocabulary Test- 4 (EOWPVT-4; Martin & Brownell 2011)*. Children named pictured objects, actions, and concepts that were ordered by increasing difficulty, beginning with the easiest concepts and ending with less frequently encountered items. The task was discontinued when the child failed six consecutive items. The publisher reports the median internal consistency reliability coefficient as .95.

Conceptual Expressive Vocabulary. Spanish-English conceptually-scored expressive vocabulary was assessed with the *Expressive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition (EOWPVT-4: SBE; Martin, 2013)*. Children named pictured objects, actions, and concepts that were ordered by increasing difficulty, arranged in a developmental sequence specific to the bilingual normative sample using Classical Test Theory and Item Response Theory, beginning with the easiest concepts and ending with less frequently encountered items. The task was discontinued when the child failed six consecutive items. As a Spanish-English conceptually-scored measure, children were first prompted in Spanish (per parent report of Spanish being the dominant home language). If the child did not respond or answered incorrectly, the examiner repeated the prompt in English, allowing for the assessment of expressive knowledge in *either* language. If the child missed the item once both languages were targeted, the item was scored as incorrect. The publisher reports the median internal consistency reliability coefficient as .95.

The bilingual normative sample for both the *ROWPVT-4: SBE* and *EOWPVT-4: SBE* (Martin, 2013) approximates the U.S. Hispanic population; roughly 63% of the normative sample was of Mexican origin, roughly 41% resided in the West Coast region of the U.S., and about one-third reported less than a high school education. If under age 18, accounting for roughly 73% of the sample, parent education is utilized. The present sample is similar to this normative sample, although the present sample is predominantly of Mexican origin, is entirely from the West Coast region of the U.S., and reported lower levels of education.

Table 3 illustrates the distribution of items for both the *ROWPVT-4-SBE* and the

EOWPVT-4-SBE (Martin, 2013) in semantic categories, modeled after the MacArthur-Bates CDI Toddler Long Forms (Fenson et al., 2007). We utilized the CDI semantic categories because they represent the types of words children typically learn first, in both Spanish and English. Each item was entered once, with no overlap. The semantic category, “Other nouns,” included nouns that did not fit into any of the other categories.

TABLE 3
Spanish-English bilinguals children’s (n=32) receptive and expressive vocabulary performance on the Spanish-English Conceptual vocabulary measures (using either Spanish or English)

MacArthur-Bates CDI	Receptive		Expressive	
	Number of items	% Correct	Number of items	% Correct
Animals	9	91	18	50
Vehicles	3	74	8	63
Toys	1	100	2	33
Food and Drink	6	67	10	56
Clothing	6	89	2	63
Body Parts	3	96	6	80
Small Household Items	6	72	7	79
Furniture and Rooms	3	100	5	26
Outside Things	9	74	7	45
Places to Go	2	86	2	17
People	4	77	2	38
Games and Routines	1	53	1	0
Action Words	13	66	7	50
Nouns	17	63	28	35
Adjectives/Descriptive Words	11	68	0	--
Average Percentage		78		45

Note. MacArthur-Bates CDI was used to categorize items from the receptive and expressive conceptual vocabulary assessments. Items were classified once; there were no overlapping categories containing the same item. Items on the receptive and expressive tests are not exactly equivalent.

While not exactly similar, the majority of the items between the conceptual and English assessments were the same; there was over 90% overlap between the two language versions of both the expressive and receptive assessments. Additionally, it is interesting to note the percentages of cognates for each assessment. The percentages of the total number of words administered that were cognates are as follows: English Receptive – 40%; Conceptual Receptive – 38%; English Expressive – 36%; and Conceptual Expressive – 46%.

Analytic Approach

Research Question 1a: Mean Score Comparisons. The first step was to compare Spanish-English bilingual children's English-only and Spanish-English conceptually-scored receptive and expressive vocabulary scores. We thus first conducted paired t-tests *between* the English-only and Spanish-English conceptually-scored vocabulary scores to examine whether a difference exists across languages, receptively and expressively (i.e., between-language comparisons). We then also conducted paired t-tests *within*-language to examine whether a difference exists within language, that is between English receptive and English expressive vocabulary, as well as between Spanish-English receptive conceptually-scored vocabulary and Spanish-English expressive conceptually-scored vocabulary (i.e., within-language comparison). We underscore that the standardized assessments we utilized provide normative information for determining whether standard score differences exist between receptive and expressive vocabulary (i.e., within-language differences in receptive and expressive vocabulary). For this reason, we conducted paired-t-tests using both the raw and standard scores. Our rationale for using raw scores was that absolute differences that may exist are masked by standard scores. However, and as noted, we felt compelled to also examine standard score differences given that the assessments rely on standard score differences when determining whether discrepancies exist between students' within-language receptive and expressive vocabulary. We also calculated effect sizes for both the between- and within-language comparisons and interpret differences using Cohen's (1992) conventions for effect sizes (i.e., $\sim .2$ a small effect, $\sim .5$ a medium effect, and $.8+$ a large effect).

Research Question 1b: At-risk Classifications. In line with previous research (e.g., Gibson et al., 2012; Gross et al., 2014; Miccio et al., 2005; Oller et al., 2007a, 2007b; Windsor & Kohnert, 2004), we next compared the proportion of students who fell within each of the following three categories on the English-only vs. the Spanish-English conceptually-scored receptive and expressive vocabulary measures: 1) -1 SD *below* the mean, 2) within the *average range*, and 3) +1 SD *above* the mean.

Research Question 2: Performance at the Item-level. We examined the extent to which children demonstrated knowledge in Spanish, compared to having to resort to English or to not knowing the item. As noted, all children were administered the Spanish-English conceptually-scored vocabulary measures in Spanish first (per parent report of Spanish being the dominant home language, not because administration has to begin in Spanish), and then in English for any missed Spanish items. Thus, we calculated the percentage of items known in Spanish, then in English if missed in Spanish, or in neither language if missed in both languages.

RESULTS

Preliminary descriptive analyses

As Table 2 shows, there was a positive, moderate association between patterns of language exposure and language use for mother/child, father/child, and other adults/child pairs. Further, the language association between other children in the home and the target child was virtually

perfect ($r = .97, p < .001$). We thus combined the language exposure (i.e., all questions related to the language(s) children heard from household members) and language use (i.e., all question related to the language(s) children spoke to household members) data. The correlation between the language exposure and language use variables was positive and high ($r = .72, p < .001$). We then proceeded to create a composite home language variable by averaging the overall home language exposure and overall home language use to reflect Spanish-English bilingual students' home language. The home language composite average was 2.1 (1.0), reflecting predominant Spanish use at home.

Table 4 displays the correlations among all measures. There is a positive, moderate association between income and home language and between income and the English-only receptive and expressive vocabulary measures. However, there is no association between income and Spanish-English conceptually-scored vocabulary. Further, the magnitude and significance of association between students' English-only receptive and expressive vocabulary performance and students' Spanish-English conceptually-scored receptive and expressive vocabulary performance is similar. Finally, home language is not associated with children's vocabulary scores.

TABLE 4
Correlations among all variables for Spanish-English bilingual children ($n = 32^a$)

	Income	Home Language	English-only receptive	English-only expressive	Spanish-English Conceptual receptive	Spanish-English Conceptual expressive
Income	1.0	.51**	.39*	.46**	.03	-.10
Home Language	---	1.0	.04	.10	.07	-.09
English-only receptive	---	---	1.0	.75***	.52**	.43*
English-only expressive	---	---	---	1.0	.33	.41*
Spanish-English Conceptual receptive	---	---	---	---	1.0	.63***
Spanish-English Conceptual expressive	---	---	---	---	---	1.0

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; $N = 32$, except for income and home language as 1 parent interview was missing

Key Results

Research Question 1a: Mean Score Comparisons. Table 5 displays Spanish-English bilingual students' English-only and Spanish-English conceptually-scored receptive and expressive vocabulary raw and standard scores.

Between-language comparisons. Comparison of the English-only and Spanish-English conceptually-scored receptive vocabulary raw and standard scores revealed a significant difference, such that students performed lower on English-only compared to Spanish-English conceptually-scored vocabulary (9.4 raw score point difference; 18.6 standard score point difference). Indeed, the effect size for the raw scores was moderate (-0.58) and it was large when comparing the standard scores (-1.34). However, when comparing students' English-only and Spanish-English conceptually-scored expressive vocabulary, results differed depending on whether raw or standard scores were considered. Although Spanish-English expressive conceptually-scored vocabulary raw scores were higher than English-only expressive raw scores, on average (4.9 raw score point difference), this raw score difference did not reach statistical significance and the effect size was small (-0.31). However, the expressive standard score difference of 23.8 points was statistically significant, indicating that students evidenced higher Spanish-English conceptually-scored than English-only expressive vocabulary standard scores, yielding a large effect size (-1.40).

Within-language comparisons. Comparison of children's raw and standard scores on English-only receptive compared to English-only expressive vocabulary reveal a significant difference; students performed lower on expressive compared to receptive vocabulary (7.2 raw score point difference; 6.2 standard score point difference). Here, the effect size was moderate (0.43). Furthermore, when utilizing the difference scores provided by the publisher, 50% of our sample evidenced a significant difference between their English receptive and expressive scores. Of these students, 88% evidenced higher receptive than expressive scores. The publisher reports that more than 25% of the standardization sample also evidence such a standard score difference. Finally, when comparing students' Spanish-English conceptually-scored receptive vocabulary to their Spanish-English conceptually-scored expressive vocabulary, results differed depending on whether raw or standard scores were considered. While there was a significant within-language difference in children's Spanish-English conceptually-scored receptive and expressive raw scores, with students performing lower on expressive compared to receptive vocabulary (11.7 raw score difference and large effect size of 0.77), the standard score difference was not significant (1 standard score point difference and negligible effect size of 0.06). Additionally, when utilizing the difference scores provided by the publisher, 56% of our sample evidenced a significant difference between their Spanish-English conceptually-scored receptive and expressive scores. Of these students, 61% evidenced higher receptive than expressive scores. The publisher reports that more than 25% of the standardization sample also evidence such a standard score difference.

Research Question 1b: At-risk Classifications. Table 6 shows the percentage of Spanish-English bilingual children across each category (i.e., -1SD below the mean, average range, and +1SD above the mean) on the English-only and Spanish-English conceptually-scored receptive and expressive vocabulary measures. When English-only vocabulary measures were

considered, compared to the Spanish-English conceptually-scored receptive and expressive vocabulary measures, there were notable differences in the percentage of Spanish-English bilingual children across each category. Specifically, nearly half and over half of the Spanish-English bilinguals were classified as -1SD below the mean on receptive and expressive English-only vocabulary, respectively. But when Spanish-English conceptually-scored vocabulary was accounted for, the landscape changes considerably, with Spanish-English bilingual children's classifications as being -1SD below the mean reduced to 12.5% receptively and 6% expressively.

TABLE 5
Between-language (top of table) and within-language (bottom of table) comparisons of Spanish-English bilingual children's (n = 32) receptive and expressive vocabulary raw and standard score means (standard deviations in parentheses)

Between Language Comparisons		
	Raw	Standard Scores
Receptive English	39.1 (16.0)	84.8 (14.7)
Receptive Conceptual	48.5 (16.4)	103.4 (12.9)
English-Conceptual Difference	-9.4 (t = -3.36, p = .0021)	-18.6 (t = -6.92, p < .0001)
Cohen's <i>d</i>	-0.58	-1.34
Expressive English	31.9 (17.7)	78.6 (17.8)
Expressive Conceptual	36.8 (13.8)	102.5 (16.2)
English-Conceptual Difference	-4.9 (t = -1.60, p = .1195)	-23.9 (t = -6.69, p < .0001)
Cohen's <i>d</i>	-0.31	-1.40
Within Language Comparisons		
	Raw	Standard Scores
Receptive English	39.1 (16.0)	84.8 (14.7)
Expressive English	31.9 (17.7)	78.6 (17.8)
English Difference	7.2 (t = -3.39, p = .0019)	6.2 (t = -2.84, p = .0080)
Cohen's <i>d</i>	0.43	0.38
Receptive Conceptual	48.5 (16.4)	103.4 (12.9)
Expressive Conceptual	36.8 (13.8)	102.5 (16.2)
Conceptual Difference	11.7 (t = -5.40, p < .0001)	0.9 (t = -0.43, p = 0.6727)
Cohen's <i>d</i>	0.77	0.06

Note. Receptive English refers to the *Receptive One-Word Picture Vocabulary Test*, Receptive Conceptual refers *Receptive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition*, Expressive English refers *Expressive One-Word Picture Vocabulary Test*, and Expressive Conceptual refers to the *Receptive One-Word Picture Vocabulary Test- 4: Spanish-Bilingual Edition*. All of the assessments are based on a population distribution having a mean of 100 and standard deviation of 15.

Research Question 2: Performance at the Item-level. Table 3, shown previously, shows the percentage of overall correct items answered by children in *either* Spanish or English for each MacArthur-Bates CDI (Fenson et al., 2007) semantic category on the receptive and expressive Spanish-English conceptually-scored vocabulary measures. On average, 78% of

answers were correct in either language receptively and 45% were correct in either language expressively.

We next examined the percentage of correct answers that the children gave in Spanish, receptively and expressively, before potentially resorting to English, and the percentage of correct answers in English (of the total answers). On average across all semantic categories, the majority (62%) of the children answered correctly receptively when initially prompted in Spanish; 16% of children gave correct answers on the receptive assessment when prompted in English. On expressive vocabulary, however, only 24% of the children gave the correct answer when initially prompted in Spanish; 21% of children provided the correct answers on the expressive assessment in English. Figure 1 illustrates the breakdown by CDI category, for both receptive and expressive language, of the percentages of children who answered: (1) correctly in Spanish (dark grey shading); (2) correctly in English (medium gray shading); and (3) incorrectly in both languages (light grey shading).

As a final step, to glean further insight into the children’s receptive and expressive knowledge in English given that the assessments began in Spanish per parents’ report of Spanish being the dominant language, we calculated the percentages of correct answers given in English of those prompted in English (Figure 2). On the receptive assessment, on average, 38% of the questions were prompted in English (after being answered incorrectly in Spanish; see dark grey shading) and, of those, 55% were answered correctly (see light grey shading). On the expressive assessment, on average, 71% of the questions were prompted in English (after being answered incorrectly in Spanish; see dark grey shading) and, of those, 35% were answered correctly (see light grey shading).

TABLE 6
Comparison of at-risk classifications for Spanish-English bilinguals, based on English-only and Spanish-English conceptual receptive and expressive vocabulary assessments

	-1SD	Average Range	+1SD
Receptive			
<i>English-Only</i>	47%	53%	0%
<i>Spanish-English Conceptual</i>	12.5%	75%	12.5%
Expressive			
<i>English-Only</i>	59%	38%	3%
<i>Spanish-English Conceptual</i>	6%	75%	19%

DISCUSSION

To our knowledge, this is the first study to assess Spanish-English bilingual preschoolers’ vocabulary knowledge using *standardized* receptive and expressive measures that utilize conceptual scoring and are specifically designed for and normed on this population, and also the first study to describe students’ patterns of conceptually-scored receptive and expressive vocabulary performance at the item level. Two key findings emerged from this study. First, in line with previous work (e.g., Core et al., 2013; Gross et al., 2014; Mancilla-Martinez et al., 2011; Pearson et al., 1993), vocabulary measures that utilize conceptual scoring appear to be a step in the right direction to gain a more comprehensive understanding of young Spanish-English

bilingual children's vocabulary knowledge, although there are caveats to consider. Second, the description of students' performance on the conceptually-scored Spanish-English vocabulary measures revealed that receptive and expressive vocabulary knowledge may differ. We discuss the findings in the sections that follow, focusing on assessment implications but also offering preliminary implications for practice and directions for future work in this area.

If bilingual students' vocabulary had only been assessed in English — a common practice in U.S. classrooms — a different profile would have emerged, namely one in which young Spanish-English bilinguals are found to evidence a substantial receptive and expressive English vocabulary gap and are identified as “at-risk” at high rates. In a normal distribution, roughly 15-16% of the population is expected to fall -1SD below the mean. Yet, 47% and 59% of children in our sample fell -1SD below the mean receptively and expressively, respectively, on English-only vocabulary (see Table 6). Furthermore, the within-language comparison of students' raw and standard scores revealed a significant receptive-expressive English vocabulary gap, such that their performance was significantly higher receptively than expressively. As a reference point, the publisher reports that more than 25% of the standardization sample evidence a difference between their English receptive and expressive scores for this age bracket, but they do not specify the direction of the gap (i.e., whether receptive or expressive is higher); in our sample, 50% evidenced this pattern, with nearly all (88%) evidencing higher receptive than expressive scores.

In contrast to the English-only vocabulary results, when measures that utilize Spanish-English conceptual scoring were considered, only 12.5% and 6%, receptively and expressively, respectively, of bilingual preschoolers fell -1SD below the mean (see Table 6). Furthermore, the within-language conceptually-scored vocabulary comparisons revealed that, as was the case with English-only vocabulary, bilingual students evidenced significantly higher receptive than expressive scores. However, this was only the case when raw scores were considered. That is, when comparing standard scores, the 1 standard score point difference was not significant. Again as a reference point, the publisher reports that 25% of the standardization sample evidence a difference between their conceptually-scored receptive and expressive scores for this age bracket, but they do not specify the direction of the gap (i.e., whether receptive or expressive is higher); in our sample, 56% evidenced this pattern, with over half (61%) evidencing higher receptive than expressive scores. This means that, as a sample, Spanish-English bilingual children generally evidence higher receptive than expressive scores, whether English-only or Spanish-English conceptually-scored performance is considered. The between-language comparisons provide further insight.

When directly comparing English-only and Spanish-English conceptually-scored receptive vocabulary, students performed significantly higher on Spanish-English conceptually-scored vocabulary compared to English-only vocabulary, whether raw or standard scores were considered. However, when comparing English-only and Spanish-English conceptually-scored expressive vocabulary, findings were mixed. Even though, on average, bilingual preschoolers evidenced higher expressive raw scores on the measures that utilized Spanish-English conceptual scoring compared to the English-only measures, the difference did not reach significance. But when standard scores were compared, the difference was significant. The bulk of work to date has focused on comparing standard scores (Gibson et al., 2012; Gross et al., 2014) and our results align with those of previous findings, revealing that receptive and expressive scores that utilize Spanish-English conceptual scoring tend to effectively be higher than English-only scores.

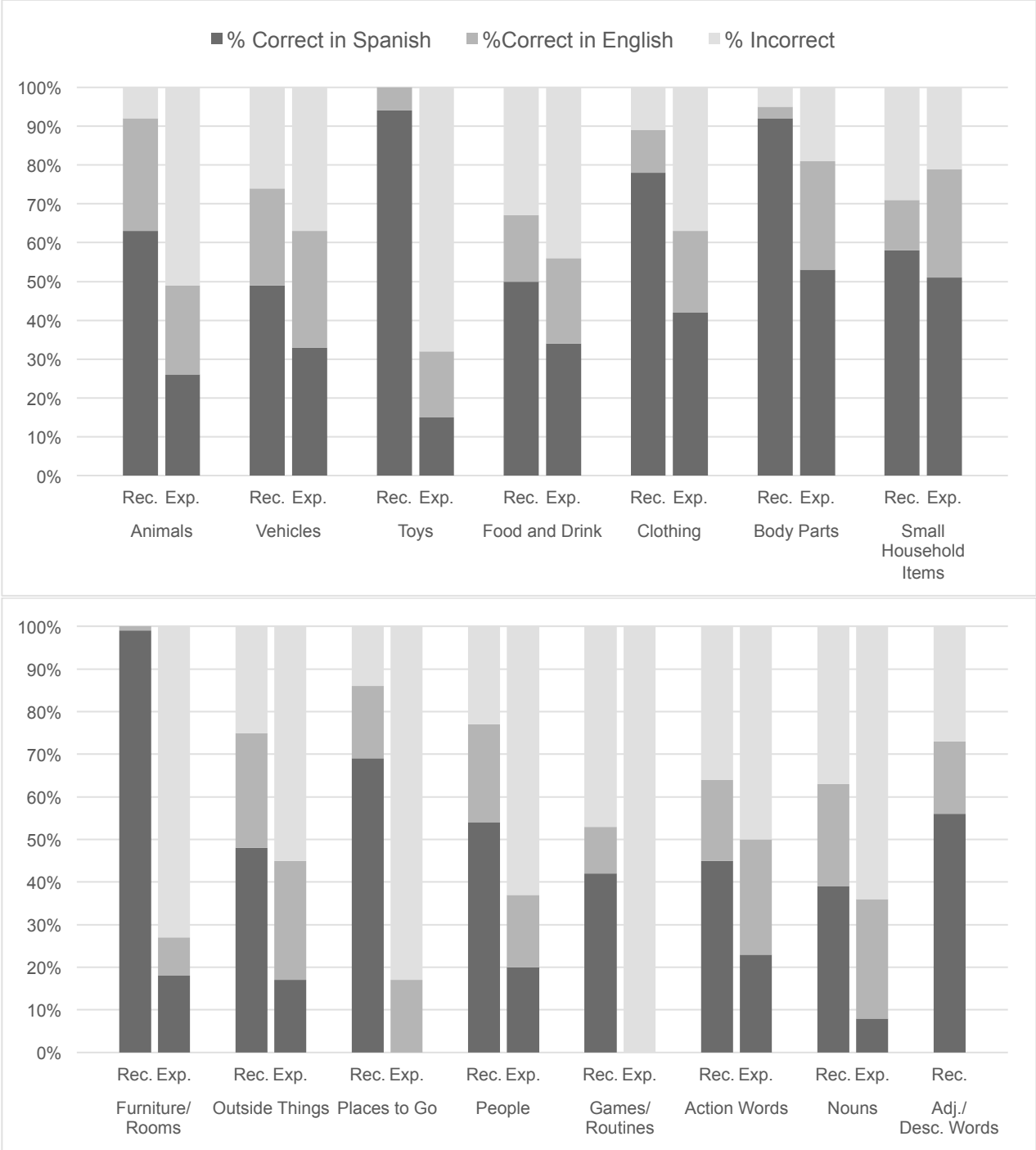


Figure 1. Comparison between receptive and expressive Spanish-English conceptual vocabulary, illustrating the percentage of correct answers in Spanish, percentage of correct answers in English (of the total answers given), and the percentage of incorrect answers in both languages. There were no expressive items in Adjectives/Descriptive Words.

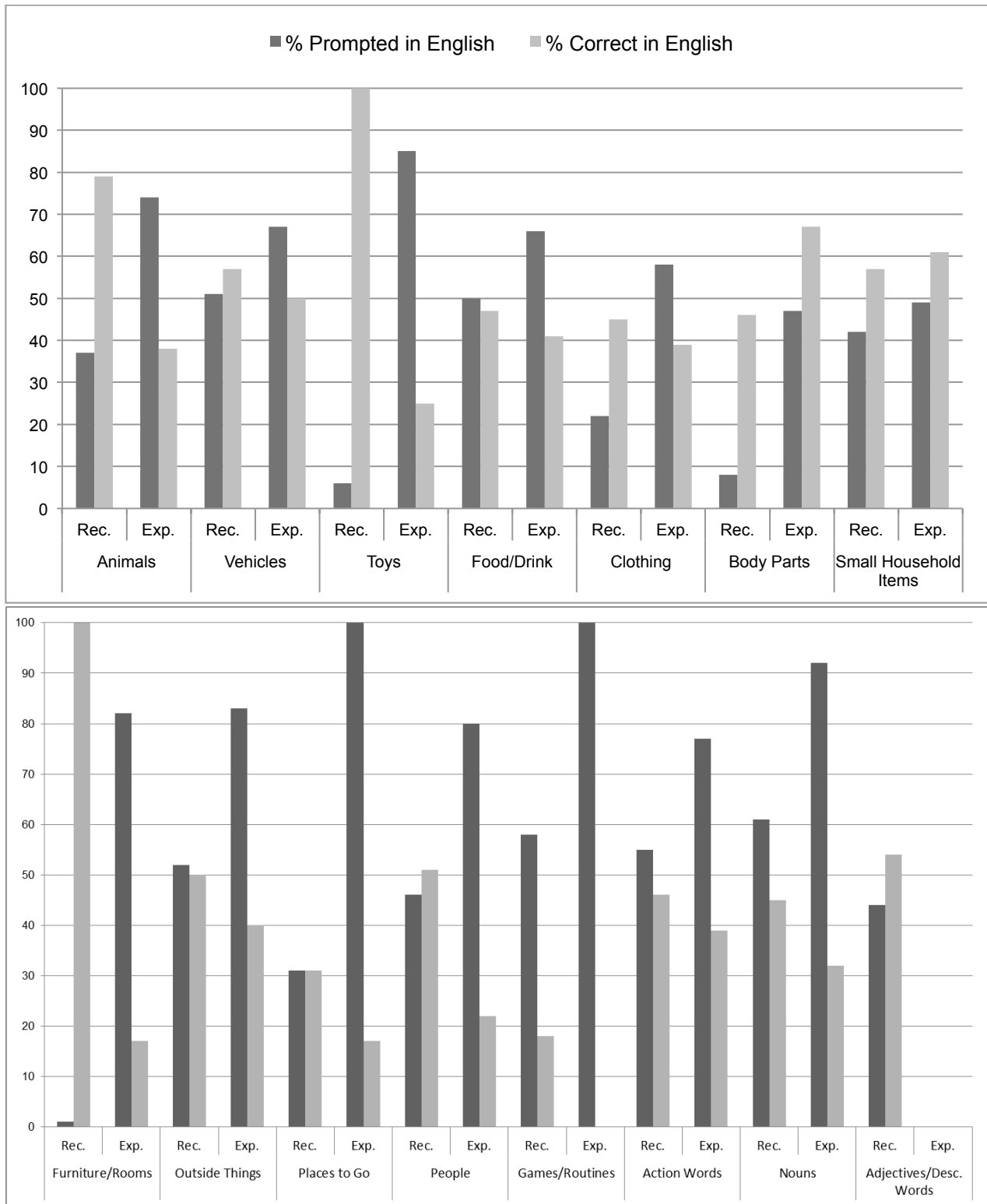


Figure 2. Comparison between receptive and expressive Spanish-English conceptual vocabulary, illustrating the percentage of questions prompted in English, and, of those, the percentage of correct answers given in English. There were no expressive items in Adjectives/Descriptive Words. Note that the English percentages are not an overall representation of students' English receptive vocabulary, but instead represent the percentage of words known conceptually in English if the item was missed in Spanish.

Taken together, our results comparing patterns of receptive and expressive vocabulary performance diverge with those of others who have pointed to the possibility of expressive language delays, and particularly of Spanish expressive language delays (Gibson et al., 2012; Gross et al., 2014; Miccio et al., 2005; Oller et al., 2007a, 2007b; Swanson et al., 2008). While our results do reveal that Spanish-English bilingual preschoolers' tend to evidence higher receptive than expressive vocabulary scores, in English and in Spanish-English conceptually-scored measures, the gap is larger in English, when standard scores (the typical metric used in previous studies) are considered. These mixed findings likely relate to the nature of measures used; while previous work has compared English-only versus Spanish-only vocabulary, we focused on comparisons of English-only versus Spanish-English conceptually-scored vocabulary. Our second key finding, targeting performance on the conceptually-scored measures, provides more insight and nuance on the nature of the receptive-expressive gap.

At the surface level, Spanish-English bilingual children's receptive and expressive Spanish-English conceptually-scored vocabulary performance did not differ; the average standard score in both domains was 103 (see Table 5). However, comparison of students' raw scores on the conceptually-scored measures revealed a significant difference, such that bilingual preschoolers had higher receptive compared to expressive Spanish-English conceptually-scored vocabulary scores. Thus, similar to English monolinguals (Benedict, 1979), Spanish-English bilinguals indeed appear to exhibit better receptive than expressive skills (Gibson et al., 2012; Gross et al., 2014; Miccio et al., 2005; Oller et al., 2007a, 2007b; Windsor & Kohnert, 2004). But examination of students' performance at the item-level lends unique insight into conceptually-scored vocabulary performance.

As noted, the conceptually-scored vocabulary measures were administered to children in Spanish first and missed items were re-administered in English. An examination of the extent to which children obtained the correct response in Spanish or resorted to English revealed that children generally evidenced Spanish receptive vocabulary knowledge without resorting to English. In contrast, the percentage of correct responses on expressive vocabulary in Spanish was notably lower, resulting in a higher percentage of questions prompted in both languages, and over half of the expressive questions were answered incorrectly in both languages. Indeed, only 38% of the items were re-administered in English on the receptive conceptually-scored assessment, whereas a majority (71%) of the items had to be re-administered in English on the expressive conceptually-scored assessment. Once assessed in English, children were more likely to produce the correct response receptively than expressively (55% vs. 35%). While these results give us insight into children's Spanish-English conceptually-scored vocabularies, it is plausible that children's English skills were under-represented; children who initially answered correctly in Spanish were not given the opportunity to respond in English and thus their English receptive and expressive skills might have been higher than our study was able to determine.

Notwithstanding the noted limitations of our results given the standardized test administration requirements, our findings do suggest different patterns between receptive and expressive vocabulary and indicate there might be differences between types of words (based on the CDI semantic categories) children know. To this end, it is important to consider the context (i.e., such as the home or preschool) in which children might encounter different semantic categories of words (Bialystok et al., 2010). Children performed highest in Spanish on the Spanish-English conceptually-scored receptive and expressive categories for "body parts" and "clothing," on the receptive categories for "furniture and rooms" and "toys," and on the expressive categories for "small household items" – vocabulary likely heard and/or used at home. On the other hand, children resorted to English at the highest rate on the Spanish-English conceptually-scored receptive and expressive categories for "nouns" and "games and routines,"

on the receptive categories for “action words” and on the Spanish-English conceptually-scored expressive categories for “places to go” – all likely topics encountered with greater frequency in a school context. Further, the context in which children were assessed likely influenced their patterns of performance as the setting can play a role in priming bilinguals toward a more monolingual or bilingual language mode (see Grosjean, 1982, 1989, 2008). In this case, children attended English-only preschools; they may have responded in English expressively on account of the instructional context. That is, they may have purposely switched expressively to English. By re-administering missed Spanish items in English, children may have gathered that either language was acceptable, which is true, and may have preferred to respond in English. Although the finding that the majority of children exhibited strong Spanish receptive skills suggests a fair amount of exposure to and familiarity with Spanish vocabulary, it does not allow for deep insight into their English receptive vocabulary; because they were not prompted in English first, we do not know the extent of their English vocabulary comprehension.

Assessment Implications

Our sample of Spanish-English bilinguals evidenced Spanish-English conceptually-scored vocabulary knowledge that does not raise concerns, based on standard scores. But when raw scores were considered, students tended to evidence stronger receptive compared to expressive skills, which is not unexpected given previous findings and the children’s developmental stage. Together, these findings have assessment implications for preschool classrooms serving children from Spanish-speaking homes, particularly in how language disabilities in Spanish-English bilinguals are diagnosed. While we know little about how best to predict language impairment among bilingual populations (Lugo-Neris et al., 2015), we do know that roughly 7% of monolingual English speakers have language impairments (Tomblin et al., 1997); the percentage of bilingual children flagged as potentially at-risk for language impairments should be roughly similar (Gilliam, Peña, Bedore, Bohman, & Mendez-Perez, 2013). Yet, roughly half of our sample fell -1SD below the mean on English-only receptive and expressive vocabulary, but only 12.5% and 6% did so when Spanish-English conceptually-scored receptive and expressive vocabulary, respectively, were assessed. Our findings suggest that using Spanish-English conceptually-scored measures could potentially improve the diagnostic accuracy of a language disability for young Spanish-English bilinguals.

Relatedly, assessments that utilize conceptual scoring and provide normative data, such as the standardized assessments used in this study, appear to hold potential for raising educators’ awareness of the strengths that their Spanish-English bilingual students bring to the language-learning task. By having a more comprehensive understanding of the words bilingual children know, educators may have higher expectations, resulting in higher quality instruction and potentially helping to minimize misrepresentation. Specifically, use of conceptually-scored measures to assess bilingual students’ vocabulary could enable educators to focus their efforts on language learning tasks that *build* upon students’ conceptual skills. For example, if a Spanish-English bilingual preschooler does not appear to understand instructions or a story, it may well be due to a mere language difference, not a disorder. That is, the child may indeed understand the concept being presented in the home language versus not understanding the concept in any language. Attention to the semantic categories of words bilinguals did and did not encounter difficulties with, by language, can serve as one tool for teachers to select target words/concepts

for instruction, thereby supporting students' language learning process and reducing the possibility of misidentification.

Finally, our sample demonstrated strong receptive skills in Spanish, yet exhibited more reliance on English expressively. While young Spanish-English bilinguals may exhibit a strong comfort in Spanish receptive vocabulary, the fact that they are fairly balanced between the two languages on expressive vocabulary makes them prime candidates for verbal interactions in the classroom. Encouraging talk across multiple contexts may be a useful strategy.

Limitations and Future Research

Findings from this study raise questions and issues to be addressed in future research. First, this study's sample consisted of Spanish-English bilingual children from Spanish-speaking, predominantly Mexican immigrant, low-income homes enrolled in a state-funded, English-medium preschool program in Southern California, and the majority of children were female. Thus, any generalizations from our findings must consider the demographics of the children studied. For example, we do not have data on the children's length of exposure to English, or sources of English for the children, limiting our understanding of whether they are simultaneous or sequential bilinguals. Further, the lack of association between children's home language use and their vocabulary was expected given our small sample size and the predominant use of Spanish in the home; both factors likely placed a limit on the detection of a relationship. Yet, home language use and income positively and significantly correlated, as did income and students' English vocabulary. These patterns of association point to the need for further research in this area. The fact that income only related to English vocabulary, but not Spanish-English conceptually-scored vocabulary, raises the possibility that the English measures proxy experiential differences between monolinguals and bilinguals, namely children from low-income homes.

Second, our comparison of Spanish-English bilingual students' performance on the English-only and Spanish-English conceptually-scored receptive and expressive vocabulary measures must not be taken to mean that the items on the tests in the two languages are equivalent. There were a handful of different words between the English-only and Spanish-English conceptually-scored versions. Additionally, the overlapping words (i.e., the words that were on both versions) were presented in different orders, reflecting the different levels of difficulty of words in English and Spanish. However, these differences were slight, and we were thus able to make a comparison between the two types of assessments, offering insight into the competencies of young Spanish-English bilinguals in both languages. In order to minimize concerns about testing effects, the two types of assessments (English-only and Spanish-English conceptually-scored) were administered on different test days. Third, our description of students' performance at the item level is limited in its scope. While it provides a good indication of children's knowledge receptively and expressively in Spanish, as this was the language in which all questions were initially asked per parent report of Spanish being the dominant language, it does not serve as an indicator of children's English knowledge. For example, a child could have known words in English, but followed directions/answered in Spanish because that was the language in which the prompt was given. Further research differentiating between Spanish *and* English expressive and receptive knowledge is warranted.

An additional consideration is that analysis of raw compared to standard scored yielded somewhat different results. Specifically, we found stable results in the receptive domain such that significant differences emerged between and within languages independent of whether raw or standard scores were considered. However, results for the expressive domain revealed that significant between-language differences emerge only when standard scores are considered, and we underscore that previous work has relied on standard score comparisons. Yet the opposite was found for the within-language comparison such that a significant difference only emerged on the receptive conceptually-scored measures when raw scores were considered. Because standard scores do not measure absolute growth, the argument can be made that use of raw scores may be more appropriate. However, the standardized conceptually-scored measures used rely on standard score differences to flag students with a receptive-expressive gap, compelling us to compare students' performance using both metrics.

Finally, it was beyond the scope of this study to investigate the extent to which use of the Spanish-English conceptually-scored vocabulary measures predict Spanish-English bilingual students' later academic outcomes and/or possible language impairments compared to single-language measures. Goodrich, Lonigan, Kleuver and Farver (2015) recently investigated the predictive validity of Spanish-English conceptual scoring among preschoolers from Spanish-speaking homes, finding that Spanish-English conceptual scores did not predict later single-language vocabulary scores. However, the authors relied on the *adaptation* of monolingual measures to index Spanish-English conceptual knowledge, and they underscored that studies that investigate the predictive validity of Spanish-English conceptual scoring in diagnosing language impairment are needed and may prove useful. At the same time, it is also possible that the associated norms for the standardized conceptual vocabulary measures mask the need for vocabulary supports among children from low-income homes; such a possibility further underscores the need for studies that investigate the predictive validity of these measures. We believe that longitudinal studies are a natural next step to address this critically important issue, and utilizing a larger sample size and a wider developmental range will allow for more concrete, generalizable findings. Relatedly, inclusion of measures that tap Spanish-only receptive and expressive vocabulary knowledge would lend additional insight into this population's vocabulary knowledge. Given the district's practical concern with additional testing time from preschoolers' instructional day, we limited testing to English-only and Spanish-English conceptually-scored vocabulary measures.

Our findings demonstrate that Spanish-English bilingual children from immigrant, low-income homes evidence receptive and expressive Spanish-English conceptually-scored vocabulary knowledge that is on par with that of their monolingual English peers, when standard scores are considered. Consequently, Spanish-English bilingual children's at-risk classifications mirror those of their monolingual English peers. However, we found that a significant between-language difference only holds for receptive, but not expressive, vocabulary, when raw scores are compared. Finally, our analysis of children's performance at the item-level on the conceptually-scored vocabulary measures provides unique insight into receptive and expressive patterns of performance in each language and reveals that Spanish-English bilinguals evidence rather strong Spanish receptive vocabulary knowledge and more distributed Spanish and English expressive vocabulary knowledge. This work underscores the need to ensure that we recognize and appropriately assess the knowledge base that Spanish-English bilingual preschoolers bring to the vocabulary-learning task and suggests that reliance on monolingual measures and monolingual norms will provide only partial insight and a misrepresentation of bilingual

students' knowledge, contributing to misidentification and thus placing a limit on appropriate instructional supports.

REFERENCES

- Bandel, E., Atkins-Burnett, S., Castro, D. C., Wulsin, C. S., & Putman, M. (2012). *Examining the Use of Language and Literacy Assessments with Young Dual Language Learners*. Research report #1. Center for Early Care and Education Research-Dual Language Learners (CECER-DLL). Chapel Hill: The University of North Carolina, Frank Porter Graham Child Development Institute.
- Bedore, L. M., & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism, 11*, 1-29. doi:10.2167/beb392.0
- Bedore, L. M., Peña, E. D., Garcia, M., & Cortez, C. (2005). Conceptual versus monolingual scoring: When does it make a difference. *Language, Speech, and Hearing Services in Schools, 36*, 188-200. doi:10.1044/0161-1461(2005/020)
- Beitchman, J. H., Nair, R., Clegg, M., & Patel, P. G. (1986). Prevalence of speech and language disorders in 5-year-old kindergarten children in the Ottawa-Carleton region. *Journal of Speech and Hearing Disorders, 51*, 98-110.
- Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6*, 183-200. doi:10.1017/S0305000900002245
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*, 525-531. <https://doi.org/10.1017/S1366728909990423>
- Brooks-Gunn, J., & Duncan, G. J. (1997). The effects of poverty on children. *The Future of Children, 7*, 55-71.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155-159.
- Core, C., Hoff, E., Rumiche, R., & Señor, M. (2013). Total and conceptual vocabulary in Spanish-English bilinguals from 22 to 30 months: Implications for assessment. *Journal of Speech, Language, and Hearing Research, 56*, 1637-1650. doi:10.1044/1092-4388(2013/11-0044)
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49*, 222-251. doi:10.3102/00346543049002222
- Dickinson, D. K. (2011). Teachers' language practices and academic outcomes of preschool children. *Science, 333*, 964-967. doi:10.1126/science.1204526
- Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175-197. doi:10.1017/S1366728902003012
- Fenson, L., Marchman, V. A., Thai, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur communicative development inventories: User's guide and technical manual, 2nd ed.* Baltimore, MD: Brookes.
- García, E., & Frede, E. (2010). *Developing the research agenda for young English language learners*. New York, NY: Teachers College Press.
- García, E. E., Jensen, B. T., & Scribner, K. P. (2009). The demographic imperative. *Educational Leadership, 66*, 8-13.
- García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Malden, MA: Wiley-Blackwell.
- Genesee F., Paradis J., & Crago, M. B. (2004). *Dual language development and disorders. A handbook on bilingualism and second language learning*. Baltimore, MD: Paul H. Brookes Publishing.
- Gibson, T. A., Oller, K. D., Jarmulowicz, L., & Ethington, C. A. (2012). The receptive-expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism (Cambridge, England), 15*, 102-116. doi:10.1017/S1366728910000490
- Gibson, T. A., Peña, E. D., & Bedore, L. M. (2014a). The relation between language experience and receptive-expressive semantic gaps in bilingual children. *International Journal of Bilingual Education and Bilingualism, 17*, 90-110. doi:10.1080/13670050.2012.743960
- Gibson, T. A., Peña, E. D., & Bedore, L. M. (2014b). The receptive-expressive gap in bilingual children with and without primary language impairment. *American Journal of Speech-Language Pathology, 23*, 655-667. doi:10.1044/2014_AJSLP-12-0119

- Goodrich, J. M., Lonigan, C. J., Kleuver, C. G., & Farver, J. M. (2015). Development and transfer of vocabulary knowledge in Spanish-Speaking language minority preschool children. *Journal of Child Language, 43*, 969-992. doi:10.1017/S030500009150002X
- Gray, S. (2004). Word learning by preschoolers with specific language impairment: Predictors and poor learners. *Journal of Speech, Language, and Hearing Research, 47*, 1117-1132. doi:10.1044/1092-4388(2004/083)
- Grosjean, F. (1982). *Life with two languages: An introduction to bilingualism*. Cambridge, MA: Harvard University Press. doi:10.2307/414002
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language, 36*, 3-15. doi:10.1016/0093-934X(89)90048-5
- Grosjean, F. (2008). *Studying bilinguals*. Oxford, England: Oxford University Press.
- Gross, M., Buac, M., & Kaushanskaya, M. (2014). Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology, 23*, 574-586. doi:10.1044/2014_AJSLP-13-0026
- Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Hancin-Bhatt, B., & Nagy, W. (1994). Lexical transfer and second language morphological development. *Applied Psycholinguistics, 15*, 289-310.
- Klingner, J. K. (2014). *Distinguishing language acquisition from learning disabilities*. New York: Division of English Language Learners and Student Support, New York City Department of Education.
- Kroll, J. F., Gerfen, C., & Dussias, P. E. (2008). Laboratory designs and paradigms: Words, sounds, and sentences. In L. Wei & M. G. Moyer (Eds.), *The Blackwell guide to research methods in bilingualism and multilingualism* (pp. 108-131). Malden, MA: Blackwell.
- Kroll, J., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149-174. doi:10.1006/jmla.1994.1008
- Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp.104-129). Oxford, UK: Blackwell Publishing Ltd.
- Lopez, M. H., & Velasco, G. (2011). *Childhood poverty among Hispanics sets record, leads nation*. Washington DC: Pew Research Center/ Pew Hispanic Center. Retrieved from <http://www.pewhispanic.org/2011/09/28/childhood-poverty-among-hispanics-sets-record-leads-nation/>
- Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., & Gillam, R. B. (2015). Utility of a language screening measure for predicting risk for language impairment in bilinguals. *American Journal of Speech-Language Pathology, 24*, 426-437. doi:10.1044/2015_AJSLP-14-0061
- Luk, G., & Christodoulou, J. A. (2016). Assessing and understanding the needs of dual-language learners. In N. Lesaux & S. Jones, *The Leading Edge of Early Childhood Education: Linking Science To Policy for A New Generation of Pre-K* (pp.67-90). Cambridge, MA: Harvard Education Press.
- Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish-speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*, 1544-1560. doi:10.1111/j.1467-8624.2011.01633.x
- Mancilla-Martinez, J., & Lesaux, N. K. (2017). Early indicators of later reading comprehension outcomes among Spanish-speaking language minority learners. *Scientific Studies of Reading, 5*, 428-448. doi:10.1080/10888438.2017.1320402
- Mancilla-Martinez, J., Pan, B. A., & Vagh, S. B. (2011). Assessing the productive vocabulary of Spanish-English bilingual toddlers from low-income families. *Applied Psycholinguistics, 32*, 333-357. doi:10.1017/S0142716410000433
- Mancilla-Martinez, J., & Vagh, S. B. (2013). Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly, 28*, 555-567. doi:10.1016/j.ecresq.2013.03.004
- Martin, N., & Brownell, R. (2011). *Expressive one-word picture vocabulary test-4 (EOWPVT-4)*. Academic Therapy Publications. Novato, CA.
- Martin, N., & Brownell, R. (2011). *Receptive one-word picture vocabulary test-4 (ROWPVT-4)*. Academic Therapy Publications. Novato, CA.
- Martin, N. (2013). *Expressive one-word picture vocabulary test-4: Spanish-bilingual edition (EOWPVT-4:SBE)*. Academic Therapy Publications. Novato, CA.
- Martin, N. (2013). *Receptive one-word picture vocabulary test- 4: Spanish-bilingual edition (ROWPVT-4:SBE)*. Academic Therapy Publications. Novato, CA.

- Miccio, A. W., Tabors, P. O., Pérez, M. M., Hammer, C. S., & Wagstaff, D. A. (2005). Vocabulary development in Spanish-speaking head start children of Puerto Rican descent. In J. Cohen, K. McAlister, K. Rolstad, & J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 1614-1617). Somerville, MA: Cascadilla Press.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development, 86*, 1351-1370. doi:10.1111/cdev.12398
- Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2012). Are minority children disproportionately represented in early intervention and early childhood special education? *Educational Researcher, 41*(9), 339-351.
- Moss, G. (1992). Cognate recognition: Its importance in the teaching of ESP reading courses to Spanish speakers. *English for Specific Purposes, 11*, 141-158.
- Nagy, W. E., Garcia, G. E., Durgunoglu, A. Y., & Hancin-Bhatt, B. (1993). Spanish-English bilingual students' use of cognates in English reading. *Journal of Reading Behavior, 25*, 241-259.
- National Academies of Sciences, Engineering, and Medicine (2017). *Promoting the educational success of children and youth learning English: Promising futures*. Washington DC: The National Academies Press. <https://doi.org/10.17226/24677>
- Oller, D. K., & Eilers, R. E. (Eds.). (2002). *Language and literacy in bilingual children (Vol. 2)*. Clevedon, UK: Multilingual Matters.
- Oller, D. K., Jarmulowicz, L., Gibson, T., Hoff, E., Caunt-Milton, H., Kulatilake, S., & Woo, I. (2007a). First language vocabulary loss in early bilinguals during language immersion: A possible role for suppression. *Proceedings of the 31st Annual Boston University Conference on Language Development, 2*, 474-484.
- Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007b). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics, 28*, 191-230. doi:10.1017/S0142716407070117
- Ortiz, A., & Artiles, A.J. (2010). Meeting the needs of English language learners with disabilities: A linguistically and culturally responsive model. In G. Li and P. Edwards (Eds.), *Best Practices in ELL Instruction* (pp. 247-272). New York: Guilford Press.
- National Task Force on Early Childhood Education for Hispanics. (2007). *Para nuestros niños: Expanding and improving early education for Hispanics [Main report]*. Tempe, AZ: Arizona State University.
- Paradis, J., Genesee, F., & Crago, M. B. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning*. Baltimore, MD: Paul H. Brookes Pub. Co.
- Patterson, J. L. (1998). Expressive vocabulary development and word combinations of Spanish-English bilingual toddlers. *American Journal of Speech and Language Pathology, 7*(4), 46-56. doi:10.1044/1058-0360.0704.46
- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*, 93-120. doi:10.1111/j.1467-1770.1993.tb00174.x
- Pearson, B. Z., Fernández, S. C., & Oller, D. K. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning, 44*, 617-653. doi:10.1111/j.1467-1770.1994.tb00633.x
- Peña, E. D., Bedore, L. M., & Kester, E. S. (2015a). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders, 51*, 192-202. doi:10.1111/1460-6984.12199
- Peña, E. D., Bedore, L. M., & Kester, E. S. (2015b). Discriminant accuracy of a semantics measure with Latino English-speaking, Spanish-speaking, and English-Spanish bilingual children. *Journal of Communication Disorders, 53*, 30-41. doi:10.1016/j.jcomdis.2014.11.001
- Poplack, S. (1980). "Sometimes I'll start a sentence in Spanish y termino en Español": Toward a typology of code-switching. *Linguistics, 18*, 581-618.
- Reichow, B., Boyd, B. A., Barton, E. E., & Odom, S. L. (2016). *Handbook of early childhood special education*. Switzerland: Springer International Publishing.
- Sheng, L., & McGregor, K. K. (2010a). Lexical-semantic organization in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 53*, 146-159. doi:10.1044/1092-4388(2009/08-0160)
- Sheng, L., Peña, E. D., Bedore, L. M., & Fiestas, C. E. (2012). Semantic deficits in Spanish English bilingual children with language impairment. *Journal of Speech, Language & Hearing Research, 55*, 1-15. doi:10.1044/1092-4388(2011/10-0254)

- Swanson, H. L., Rosston, K., Gerber, M., & Solari, E. (2008). Influence of oral language and phonological awareness on children's bilingual reading. *Journal of School Psychology, 46*, 413-429.
- Tomblin, J., Records, N., Buckwalter, P., Zhang, X., Smith, E., & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech Language, and Hearing Research, 40*, 1245-1260. doi:10.1044/jslhr.4006.1245
- U.S. Census Bureau. (2014). *Preliminary estimates of weighted average poverty thresholds for 2014*. Retrieved from <http://www.census.gov/hhes/www/poverty/data/threshld/index.html>
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development, 65*, 606-621. doi:10.1111/j.1467-8624.1994.tb00771.x
- Wei, L., & Martin, P. (2009). Conflicts and tensions in classroom codeswitching: An introduction. *International Journal of Bilingual Education and Bilingualism 12*(2), 117-122.
- Windsor, J., & Kohnert, K. (2004). The search for common ground part I. Lexical performance by linguistically diverse learners. *Journal of Speech, Language, and Hearing Research, 47*, 877-890. doi:10.1044/1092-4388(2004/065)

APPENDIX A

Computing the Income to Needs Ratio

The official U.S. definition of poverty is based on a comparison of a household's income to an income threshold level that varies by family size and composition. The threshold levels are updated annually for inflation with the Consumer Price Index. If a family's total income is less than the threshold level, the family is considered to be living in poverty. For example, in 2001, a family of five (two adults and three children under age 18) with an annual income of \$25,000 had a poverty threshold of \$21,135 and thus an income-to-needs ratio of 1.18, which is above the poverty level (but considered to be near poverty). The 2014 poverty thresholds issued by the U.S. Census Bureau were used to calculate the income-to-needs ratio for this study. See <http://www.census.gov/hhes/www/poverty/data/threshld/index.html> for threshold levels by year. Following Brooks-Gunn, Duncan, and Britto (1999), we identified five income-to-needs ratio categories as our sample was predominantly living at or near poverty.

Five categories were identified: deep poverty (income-to-needs ratio less than .50), poverty (income-to-needs ratio greater than or equal to .50, but less than 1.0), near poverty (income-to-needs ratio between 1.0 and 1.5), low income (income-to-needs ratio between 1.5 and 2.0), and middle income (income-to-needs ratio greater than or equal to 2.0).