Journal of Applied Developmental Psychology 52 (2017) 159–169

Contents lists available at ScienceDirect

Journal of Applied Developmental Psychology

journal homepage: www.elsevier.com/locate/jappdp

Spanish instruction in head start and dual language learners' academic achievement☆

Elizabeth B. Miller

New York University, United States

ARTICLE INFO

Keywords:
DLLs
Head Start
Spanish instruction

ABSTRACT

Data from the Head Start Impact Study (N = 1141) and the Head Start Family and Child Experiences Survey, 2009 Cohort (N = 825) were used to investigate whether Spanish instruction in Head Start differentially increased Spanish-speaking Dual Language Learners' (DLLs) academic achievement. Although hypothesized that Spanish instruction would be beneficial for DLLs' early literacy and math skills, results from residualized growth models showed there were no such positive associations. Somewhat surprisingly, DLL children instructed in Spanish had higher English receptive vocabulary skills at the end of the Head Start year than those not instructed, with children randomly assigned to Head Start and instructed in Spanish having the highest scores. Policy implications for Head Start-eligible Spanish-speaking DLLs are discussed.

1. Introduction

Quality early care and education (ECE) has been shown to help prepare young learners for future academic success (Karoly, Kilburn, & Cannon, 2005; Magnuson, Ruhm, & Waldfogel, 2007), and this may be especially true for low-income children (Barnett, 2011; Ramey & Ramey, 2006; Schweinhart, 2006). Prior research suggests that Spanish-speaking Dual Language Learners (DLLs) – young children learning two languages simultaneously, their home language and English (Espinosa, 2013) – differentially benefit from quality ECE compared with children of other subgroups and monolingual-English children (Buysse, Peisner-Feinberg, Páez, Hammer, & Knowles, 2014; Gormley, 2008; Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007). This same finding has also been shown with the Head Start program. The Final Report of the Head Start Impact Study (HSIS; U.S. DHHS, 2010a) found that Spanish-speaking DLL children benefitted more from random assignment to Head Start compared with monolingual-English children on some English outcomes, and Bloom and Weiland (2015) found that this was particularly the case for DLL children with low baseline levels of English receptive vocabulary skills.

Despite the positive impacts of ECE for Spanish-speaking DLLs, however, the literature has not sufficiently investigated the mechanisms underlying this finding, or what may account for these differential benefits. One idea that has garnered attention in recent years through developmental policy reports (Mancilla-Martinez & Lesaux, 2014; McCabe et al., 2013), research articles (Barnett, Yarosz, Thomas, Jung, & Blanco, 2007; Páez, Tabors, & López, 2007), and Head Start mandates (U.S. DHHS, 2008) is Spanish language instruction in the classroom. Such instruction may be an important pathway through which DLL children can maximize their English learning experiences in ECE, as developmentally, DLL children need a strong basis in one language before they can acquire another. When DLL children are exposed to the home language in the classroom, they may display faster rates of growth in English language skills than native English-speaking children (Mancilla-Martinez & Lesaux, 2011). As such, understanding whether the effects of programs like Head Start on DLL children's development differ by use of Spanish instruction in the classroom remains a key issue. The recent joint policy statement on DLLs in ECE (2016) by the U.S. Departments of Health and Human Services (DHHS) and Education (ED) unequivocally states that using the home language in the classroom is optimal for DLL children's language and literacy development, but whether such instruction contributes to English language academic skills is critical and largely unknown.

Using data from the two largest, nationally representative datasets on Head Start – the Head Start Impact Study (HSIS; UU.S. DHHS, 2002–2006) and the Head Start Family and Child Experiences Survey, 2009 Cohort (FACES-2009; U.S. DHHS, 2009–2013) – the purpose of the current study is to determine whether Spanish language instruction

☆ Research reported in this publication was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P01HD065704. Research was also supported by the U.S. Department of Health and Human Services, Administration for Children and Families under Grant Number 90YE0157. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health or U.S. Department of Health and Human Services.

E-mail address: ebmiller@nyu.edu.

http://dx.doi.org/10.1016/j.appdev.2017.07.008
Received 11 March 2016; Received in revised form 28 April 2017; Accepted 24 July 2017

is associated with school readiness skills for Head Start-eligible Spanish-speaking DLL children. Of particular importance is the examination of the relationship between Spanish instruction used by caregivers in Head Start settings and DLL children's subsequent English language academic achievement, as this may prove essential for their kindergarten readiness.

1.1. Theoretical framework

This study is grounded in bioecological theory, which posits that human development results from the interplay of enduring reciprocal and continuous interactions between an organism and their environment known as proximal processes (Bronfenbrenner & Morris, 2006). Through such interactions occurring on a regular basis over extended periods of time, children come to understand the world and their place in it. The effects of these proximal processes on developmental outcomes systematically vary based on the characteristics of the person and their environmental context. Consequently, children respond in varying ways to the environments they encounter. The current study considers the proximal processes of Spanish language interactions between teachers and DLL students in the context of Head Start classrooms.

This study is guided by bioecological theory such that as part of these classroom language interactions, children continuously and reciprocally converse with adults, which extend over many turns and utilize both English and Spanish. Consequently, children take part in the building of a complex linguistic structure where using the home language helps them to productively communicate in English. From early naming exchanges with adults, children eventually become more adept at responding to listener's cues and creating sentences with new information (Tabors, 2008). Therefore, early language interactions in one language can support and privilege later experiences in another (Uccelli, Hemphill, Pan, & Snow, 1999), and can be used to explore how DLL children fare in an ecological context like a Head Start classroom where they may be encountering formal academic English for the first time.

1.2. Spanish-speaking DLLs in the U.S.

As mentioned briefly above, this study uses the term “Dual Language Learner” to describe young children who are learning more than one language simultaneously – their home language and English (Espinosa, 2013). This term encompasses the diversity of this population, which includes children from a wide variety of language backgrounds. Young DLLs may be of limited English proficiency, completely bilingual, or may not speak their home language fluently (August & Hakuta, 1997). Regardless of their home language experiences, DLLs have less English language exposure and practice than monolingual English-speaking children – children from homes where English is the primary language – and do not perform on par with such children on various emergent English skills (August & Shanahan, 2006).

Moreover, the population of U.S. DLLs is growing rapidly. From 1994 to 95 to 2009–10, the number of school-aged DLLs increased by nearly 65% (National Clearinghouse for English Language Acquisition, 2011) – from 3.2 million students to over 5.2 million students, representing the fastest growing student segment in U.S. public schools (Calderón, Slavin, & Sánchez, 2011). Among younger children in Head Start and Early Head Start, DLLs now represent close to 40% of all participants (U.S. DHHS, 2014). Furthermore, the population of Latino children represents the largest group of children in poverty in the U.S. (López & Velasco, 2011), which further places Spanish-speaking DLLs at risk for delayed English language development (Hart & Risley, 1995; Hoff, 2013; Kieffer, 2010; Mancilla-Martinez & Vagh, 2013).

1.3. Spanish language instruction for DLL children's school readiness skills

Spanish language instruction in Head Start may be one way to boost DLL children's English school readiness skills. Unfortunately, at kindergarten entry, Spanish-speaking DLL children are already engaged in a game of “catch-up”, as they trail their monolingual English-speaking peers in important English language skills such as syntactic knowledge, phonological awareness, emergent literacy, and in particular, vocabulary (Hoff, 2013; Páez et al., 2007). Further, persistent K-12 reading achievement gaps between DLLs and monolingual-English speakers (National Assessment of Educational Progress, 2013) suggest that DLL students are not equipped with the English language skills to succeed academically.

In addition to these gaps, assessing DLL children only in English is problematic, as it provides an incomplete picture of their language skills. Rather, research syntheses on second language acquisition demonstrate that a sufficiently high quality match between the classroom language environment and children's language capabilities can help children successfully become bilingual (e.g., McCabe et al., 2013). This may be particularly true when accounting for DLL children's skills inclusive of their home language and English (Hoff, 2013; McCabe et al., 2013; Pearson, Fernández, & Oller, 1993). The revised hierarchical model for bilinguals suggests that DLLs represent their two languages with one conceptual system in which proficiency in one is thought to facilitate proficiency in another (Sunderman & Kroll, 2006). For instance, if a child learns the word *mesa*, they have a conceptual understanding of the word, enabling the acquisition of the English equivalent word of table. Further, DLL children's vocabulary knowledge has been shown to be distributed across languages, so accounting for both languages provides more accurate insight into DLLs' development and skills (Mancilla-Martinez & Vagh, 2013). Therefore, given children's conceptual understanding of language as well as the timing of sensitive periods for language development (Nelson & Sheridan, 2011), the preschool years may be an ideal time to learn two languages (Bialystok, 2001, 2011; Genesee, Paradis, & Crago, 2004; Kuhl, 2009; McCabe et al., 2013).

Some empirical work supports the view that using the home language for classroom interactions and instruction may be critical to Spanish-speaking DLLs' overall development. In particular, the research on cross-linguistic transfer implies that continued rich language opportunities in children's home language may transfer and promote English language and literacy development (August & Shanahan, 2006), by helping children to integrate component skills in early literacy domains such as sound-symbol awareness, grammar, and decoding (Castro, Páez, Dickinson, & Frede, 2011). Rinaldi and Páez (2008), for example, found that Spanish-language word reading skills contributed to the development of such skills in English, while Dickinson et al., (2004) found that among Spanish-speaking DLL children, phonological awareness in one language was strongly related to phonological awareness in the other.

Furthermore, prior research with preschool-aged DLL children, some of which is experimental, demonstrates that classroom use of both the home language and English may lead to improvement in Spanish word reading skills and at least equivalent English emergent literacy skills compared with Spanish-speaking DLL children in all-English contexts (e.g., Barnett et al., 2007; Burchinal et al., 2016; Durán, Roseth, & Hoffman, 2010; Páez et al., 2007). Research on Spanish language instruction with older elementary-aged children suggests similar results for bilingual education compared with English-only programs by fourth grade (e.g., August & Shanahan, 2006; Goldenberg, 2012; Slavin & Cheung, 2005; Slavin, Madden, Calderón, Chamberlain, & Hennessy, 2011). Other work indicates that DLL children display better social skills and closer teacher-child relationships in classrooms where teachers use children's home language, and teacher ratings of DLL children's peer social skills and assertiveness may be positively associated with increased amounts of Spanish use (Chang

et al., 2007).

1.4. Oral language skills as a possible exception

Despite this encouraging body of studies, the research on the positive associations between Spanish language classroom instruction and Spanish-speaking DLLs' academic achievement is not definitive, particularly for English oral language skills of vocabulary. Spanish-speaking DLLs tend to lag behind monolingual-English speakers in English oral language skills, and especially among low-income children (Hoff, 2013; McCabe et al., 2013; Páez et al., 2007). Spanish-speaking DLL four- and five-year olds perform one to two standard deviations below monolingual-English norms, on average, on measures of English expressive and receptive vocabulary (Hoff, 2013). While an initial language gap is understandable given the limited amount of home exposure to English (compared with monolinguals), vocabulary gaps between Spanish-speaking and monolingual-English learners are still significant at age 11 (Mancilla-Martinez & Lesaux, 2011). Although these findings are troubling given the link between oral language proficiency and future literacy, academic success, and behavior (Spira, Bracken, & Fischel, 2005), because many Spanish-speaking DLLs are low-income (López & Velasco, 2011), it is not clear if these results are due to SES or language factors.

Limited work with higher-SES Spanish-speaking DLL children indicates that they can catch up to monolingual-English speakers in elementary grades (Umbel, Pearson, Fernandez, & Oller, 1992). Further, even among low-income DLL children, if their total vocabulary is measured, they tend to perform on par or score higher than monolingual children (Core, Hoff, Rumiche, & Señor, 2013; Hoff, 2013). It is unclear, however, if total vocabulary predicts future achievement.

Nonetheless, because of this lack of consensus on whether Spanish language instruction is beneficial for English achievement outcomes, a few researchers recommend that best practice for instructing DLL children should include more of a structured immersion framework within a protected Spanish language environment that is reduced after one to two years (e.g., Rossell & Kuder, 2005). Vitiello, Downer, and Williford (2011) for example found that found that although DLLs performed better on both measures of the home language as well as those in English than did their peers in groups where only English was used, more instruction in Spanish in preschool was associated with a lower chance of attaining English proficiency at the end of two years, particularly for children with very low English proficiency at preschool entry. Therefore, bilingual programs aim to provide adequate exposure and learning opportunities in English, particularly for oral language skills, in addition to fostering the home language (Goldenberg, Nemeth, Hicks, Zepeda, & Cardona, 2013).

2. Present study

Given the implications of ECE programs for improving the school readiness of all children (Karoly et al., 2005; Magnuson et al., 2007), and in particular for Spanish-speaking DLLs (Buysse et al., 2014; Gormley, 2008; Loeb et al., 2007; U.S. DHHS, 2010a), the present study views the early childhood years as a critical period to equip Spanish-speaking DLLs with the academic skills that they will need for future English literacy and school success (Hoff, 2013; Mancilla-Martinez & Vagh, 2013; National Task Force on Early Childhood Education for Hispanics, 2007). Despite the emerging body of research on the differential benefits of ECE for Spanish-speaking DLLs, whether Spanish language instruction contributes to these benefits in various ECE settings such as Head Start is still an empirical question. To address this gap, therefore, the present study tests whether Spanish language instruction in Head Start is associated with critical English academic school readiness skills, including oral language.

Prior studies examining the importance of Spanish language instruction for children's school readiness skills have relied on observational or survey data (e.g., Chang et al., 2007; Páez et al., 2007). Those that have taken advantage of random assignment experiments (e.g., Barnett et al., 2007; Durán et al., 2010; Slavin et al., 2011) used samples that were either small in size or not nationally representative. Further, no known studies to date have analyzed whether Head Start program impacts may differ by Spanish instruction in the classroom.

The present study contributes to the field by using the two largest nationally representative datasets on income-eligible Head Start children, families, and programs (HSIS and FACES-2009) to answer: 1) What is the association between Spanish language instruction and Spanish-speaking DLL children's English academic school readiness skills?; and 2) Do Head Start program impacts for Spanish-speaking DLL children differ by whether they are instructed in Spanish? Given Head Start's special focus on educating DLL children and its mandates to support their home language (U.S. DHHS, 2008) as well as DLLs representing an increasing share of Head Start and Early Head Start's participants (nearly 40%, U.S. DHHS, 2014), this study seeks to address important practical and policy questions on the associations between Spanish instruction and English early academic skills, and whether Head Start is differentially beneficial based on such use. Addressing these questions is key for Head Start programs and policy as they help to elucidate what promotes better academic outcomes for this growing population and whether Spanish language instruction in the classroom plays a role. Based on these results, Head Start and other ECE programs can better target their resources in ways that conform to national Head Start policy by supporting home language use and providing care in more culturally-responsive ways such as bilingual teacher and staff hiring, classroom language supports, and curriculum decisions that stress the importance of both languages (U.S. DHHS, 2008).

3. Method

3.1. Participants

This study is based on data from the two largest, nationally representative datasets on Head Start children, families, and programs. The first was the random-assignment Head Start Impact Study (HSIS; U.S. DHHS, 2002–2006), which was designed to estimate the causal impact of Head Start on children's school readiness skills and parenting practices, as well as determine the circumstances under which Head Start achieved its greatest impact and for which children (U.S. DHHS, 2010a, 2010b, Final Report). The second was the Head Start Child and Family Experiences Survey, 2009 Cohort (FACES-2009; U.S. DHHS, 2009–2013), a longitudinal study of program performance, and specifically the population served; staff qualifications, credentials, and opinions; Head Start classroom practices and quality measures; and child and family outcomes (U.S. DHHS, 2011).

3.1.1. HSIS

The HSIS was a nationally representative sample of 84 Head Start grantee and delegate agencies and nearly 4500 newly entering, eligible three and four-year-old children. Children were randomly assigned to either: (1) a Head Start group that had access to Head Start program services; or (2) a control group that was not eligible to enroll in the Head Start center to which they applied for the lottery, but could enroll in other early childhood programs or services selected by their parents (U.S. DHHS, 2010a, 2010b, Final Report).

The study employed a multi-stage sampling process to select a representative group of Head Start programs and children. It began with a list of 1715 grantee and delegate Head Start agencies that were operating in Fiscal Year (FY) 1998–99. This pool was then organized into 161 geographic clusters across 25 strata in order to ensure variation across region of the country, urban and rural location, race and ethnicity, and state pre-kindergarten and child care policies. One cluster was then randomly selected from each of the 25 strata yielding 261 grantee and delegate agencies. Agencies were eliminated if they had recently

closed, merged, or were serving all eligible children in their communities, and smaller agencies were grouped together. Approximately three grantee and delegate agencies were then randomly selected from each of the 25 strata, yielding a final pool of 84 grantee and delegate agencies.

These 84 Head Start agencies generated lists of 1427 individual centers that were expected to be in operation for the 2002–03 school year. After individual programs were eliminated because they had recently closed, merged, or were serving all eligible children in their communities, and groups of centers were stratified along the same dimensions as the geographical agency clusters, 383 individual centers remained (U.S. DHHS, 2010a, 2010b, Final Report).

Once the centers were selected, a lottery process was used to determine which children were and were not assigned a place in Head Start. The goal was to randomly select 27 children from each center – 16 to be assigned to Head Start and 11 to the control condition. In total 4442 children were randomly selected – 2646 for Head Start and 1796 for the control condition. Data collection took place from fall 2002, at the time the treatment group entered Head Start, until spring 2006, at the end of first grade (U.S. DHHS, 2010a, 2010b, Final Report).

Approximately 25% of the total HSIS study sample (N = 1141) was classified as Spanish-speaking DLLs (see Measures, below) – 690 in Head Start and 451 in the control condition. Half of the children in this subgroup sample were male, and about 10% were classified as having a disability at baseline. Nearly 70% of the mothers of Spanish-speaking DLL children had less than a high school education, and about 60% had immigrated to the U.S. in the past ten years. The majority of Spanish-speaking DLL mothers were married (63%), and in 75% of Spanish-speaking DLL households, the biological parents lived together with the study child.

Complete descriptive statistics for Spanish-speaking DLL children and families are listed in Table 1, which also includes tests for treatment and control group differences. As shown in the table, balance was achieved on all covariates between the Head Start and control groups.

3.1.2. FACES-2009

Head Start FACES was launched in 1997 as a periodic, longitudinal study of program performance. Successive nationally representative samples of Head Start children, their families, classrooms, and programs were collected in 2000, 2003, 2006, and 2009 to provide descriptive information on the population served; staff qualifications, credentials, and opinions; Head Start classroom practices and quality measures; and child and family outcomes. This study used FACES-2009 due to its enhanced focus on children who speak a primary language other than English (U.S. DHHS, 2011).

The sample design for FACES-2009 was similar to that of the HSIS and included a multi-stage sampling process to select a representative group of Head Start 1) programs; 2) centers; 3) classrooms; and 4) newly-enrolled children. Sampling at the first three stages was done with probability proportional to size.

The sampling frame of eligible Head Start programs for FACES-2009 was constructed from the Head Start Program Information Report (PIR), and were stratified along 12 strata in order to ensure sufficient variation in census region, urbanicity, percentage of racial and ethnic minority enrollment, whether the program had at least 25% DLLs, program status as a public school district grantee, the percentage of children in the program whose primary home language was English, and the percentage of children with disabilities. From this frame, a sample of 60 programs was selected. In addition, approximately two centers per program and three classrooms per center were selected for participation. Within each classroom, a sample of newly-enrolled children was selected. In total, the FACES-2009 sample included 60 programs, 129 centers, 486 classrooms, and 3349 children. Data collection took place from fall 2009, at the time the children entered Head Start, until spring 2012, at the end of kindergarten (U.S. DHHS, 2013a).

Approximately 25% of the total FACES-2009 study sample (N = 825) was classified as Spanish-speaking DLLs (see Measures, below). Similar to the HSIS, half of the children in this subgroup sample in FACES-2009 were male, and about 5% were classified as having a disability at baseline. Sixty percent of the mothers of Spanish-speaking DLL children had less than a high school education, and almost 60% had immigrated to the U.S. in the past ten years. Close to half of Spanish-speaking DLL mothers were married (43%), and in two-thirds of Spanish-speaking DLL households, the biological parents lived together with the study child. Table 1 provides complete descriptive characteristics of these participants as well.

Table 1
Descriptive statistics of Spanish-speaking DLLs in HSIS and FACES-2009.

	HSIS (N = 1141)				FACES-2009 (N = 825)	
	Head start (N = 690)		Control (N = 451)			
	Mean/% of sample	SD	Mean/% of sample	SD	Mean/% of sample	SD
Key independent variable						
Teach children in Spanish	0.64		0.60		0.64	
Outcomes						
Baseline						
TVIP/PPVT	222.12	37.16	225.93	35.21	271.81	15.46
WM Letras y palabras/WJ letter-word	291.71	22.93	292.18	24.21	292.36	22.21
WM Problemas aplicados/WJ applied problems	362.15	24.25	362.72	22.30	362.66	26.72
End of head start						
PPVT	249.83	29.51	239.09	*** 35.52	280.68	14.29
WJ letter-word	306.99	26.45	302.65	* 27.85	319.11	25.37
WJ applied problems	375.69	30.54	367.15	*** 35.05	383.85	23.06
Child demographic characteristics - baseline						
Age at Spring assessment in weeks	238.05	31.43	241.30	25.98	217.06	25.54
Age-4 cohort	0.54		0.55		0.48	
Gender - male	0.48		0.47		0.51	
Disability	0.08		0.09		0.03	
Family demographic characteristics - baseline						
Caregiver age	30.02	6.61	29.94	5.06	30.13	6.12
Maternal education						
Less than high school	0.62		0.70		0.60	
High school diploma/GED	0.21		0.21		0.21	
Beyond high school	0.16		0.09		0.16	
Married mother	0.63		0.63		0.43	
Teenage mother	0.09		0.08		0.39	
Parents lived together	0.78		0.75		0.66	
Maternal depression	0.13		0.14		0.11	

Note. p level of treatment/control difference: * $p < 0.05$. *** $p < 0.001$. TVIP = Test de Imagenes Peabody. PPVT = Peabody Picture Vocabulary Test. WJ = Woodcock-Johnson.WM = Woodcock-Muñoz. DLL children in HSIS and DLL children in FACES-2009 who did not pass the language screener were assessed in Spanish at baseline. Children in FACES-2009 who passed the language screener were assessed in English at baseline. All DLL children in HSIS and FACES-2009 assessed in English at the end of the Head Start year. HSIS weight used = CHSPR2003WTCA. FACES-2009 weight used = P21RA2WT.

3.2. Measures

3.2.1. Spanish-speaking dual language learner (DLL) status

3.2.1.1. HSIS. Prior to program entry in the fall of 2002, treatment and control group children were administered a battery of assessments as a baseline measure of academic achievement. The language of this child assessment was chosen by HSIS as follows. At the start of the study in fall 2002, information was collected on each child's language ability. Assessors asked the child's primary caregiver three questions: (1) What language does the child speak most often at home?; (2) What language does the child speak most often at the child care setting?; and (3) What language does it appear the child prefers to speak? Children were tested in the language in which at least two of the three responses were the same (U.S. DHHS, 2010a, 2010b, Technical Report). Children were classified as a Spanish-speaking DLL if they required assessment in Spanish at baseline. About 25% of the overall study sample required baseline assessment in Spanish (N = 1141).

3.2.1.2. FACES-2009. In FACES-2009, all children at baseline were given two subtests – Simon Says and Art Show – from the language screener the Preschool Language Assessment Survey 2000 (preLAS 2000; Duncan & DeAvila, 1998). Children whose parents reported their primary home language was Spanish and who made five consecutive errors on both subtests of the preLAS were then routed to the Spanish-language baseline assessment. Using the language screener criteria, about 15% of the total FACES-2009 sample (N = 512) required assessment in Spanish at baseline (U.S. DHHS, 2013a).

However, many researchers and policymakers who specialize in Dual Language Learning criticize these language screeners as problematic and not indicative of children's true language abilities (U.S. DHHS, 2013b). Rather, triangulation of methods that use data from the primary caregiver, such as the one used in the HSIS, are better predictors of children's language abilities. Therefore, this study applied the same three questions from HSIS to FACES-2009 to classify a child as a DLL. Because both studies were conducted by the U.S. Department of Health and Human Services, Administration for Children and Families, the same questions were available in both datasets. If the answer to at least two out of the three questions was Spanish, the child was considered a Spanish-speaking DLL. Categorized this way, about 25% of the FACES-2009 sample (N = 825) was classified as Spanish-speaking DLLs. Fifty-three percent of DLL children using this broader, more triangulated classification method were assessed in Spanish at baseline (N = 436) and 47% were assessed in English (N = 389) according to this method. Thus, there was over 85% agreement between the two classification methods as to whom required baseline testing in Spanish. This study chose to use the broader, triangulated classification method, which is endorsed more by the child development field and more closely aligned with the HSIS (U.S. DHHS, 2013b), to ensure conceptually who was considered a Spanish-speaking DLL was the same across both samples rather than focus on language of baseline testing.

3.2.2. Spanish language classroom instruction

In both HSIS and FACES-2009, each study child's primary teacher or caregiver was asked if they taught the child in their care in Spanish. This variable was coded as “0” if the teacher/caregiver interviewed indicated they did not teach the children at all in Spanish and “1” if they did. For HSIS children who attended Head Start and FACES-2009 children, this question was asked of the Lead Teacher in Head Start. For HSIS children in the control group, this question was asked of the child's primary teacher/caregiver – either the lead teacher in another program, the family day care provider, or the parent if they did not attend any kind of center care, as Head Start values parents as children first and primary teachers (Zigler & Styfco, 2010). Although this variable masks potential variability in actual Spanish use, it is useful for capturing whether teachers nationwide report using Spanish at all in the classroom. Table 1 provides information on the prevalence of Spanish instruction in HSIS and FACES-2009 and indicates there is sufficient variation among Spanish-speaking DLL children.

3.2.3. Academic achievement outcomes

3.2.3.1. HSIS. At baseline in the fall of 2002, Spanish-speaking DLL children were administered a complete Spanish language assessment battery as a measure of academic achievement that included the *Test de Vocabulario en Imágenes Peabody* (TVIP; Dunn, Lugo, Padilla, & Dunn, 1986), and two subtests of the *Bateriá Woodcock-Muñoz* (WM) *Pruebas de Aprovechamiento-Revisada* (Woodcock & Muñoz-Sandoval, 1996), *Identificación de letras y palabras* and *problemas aplicados.* The TVIP measures a child's Spanish receptive vocabulary (α = 0.93), the WM *letras y palabras* measures a child's reading identification skills of Spanish letters and words (α = 0.97), and the WM *problemas aplicados* measures a child's ability to analyze and solve math problems (α = 0.90; U.S. DHHS, 2010a, 2010b, Technical Report). These academic domains are critically important for later academic success (Duncan et al., 2007; Whitehurst & Lonigan, 2003; Yesil-Dagli, 2011), as together they form the building blocks of academic competence.

In spring 2003 and in all subsequent data collection periods, children were given only the complete English assessment battery (U.S. DHHS, 2010a, 2010b, Technical Report). Thus, the English language versions of the Peabody Picture Vocabulary Test, Third Edition (PPVT; Dunn & Dunn, 1997) and the Woodcock-Johnson (WJ) III Letter-Word Identification and Applied Problems tests (Woodcock, McGrew, & Mather, 2001) were used as outcomes. A complete list of descriptive statistics for these baselines scores and outcomes one academic year later is reported in Table 1. Although there may be benefits to instructing children in Spanish on Spanish assessment measures (e.g., Barnett et al., 2007; Burchinal et al., 2016; Durán et al., 2010; Páez et al., 2007), because DLL children were only given the English assessments as outcomes, such a focus on Spanish results was not possible and this study was limited to English-only outcomes.

3.2.3.2. FACES-2009. In FACES-2009, Spanish-speaking DLLs who required baseline assessment in Spanish based on the results of the language screener were given a complete assessment battery in the fall of 2009 that included the same three assessments as HSIS – the TVIP, and the WM *Letras y palabras* and *problemas applicados* tests. Those DLLs who passed the language screener were given the complete English language baseline battery that included the PPVT, Fourth Edition (Dunn & Dunn, 2007) and the WJ III Letter Word and Applied Problems tests (U.S. DHHS, 2013a). Similar to other experimental studies of DLL children (e.g., Barnett et al., 2007), whichever language version a child was assessed in fall 2009 was considered their baseline score, as the scores from the WJ and WM tests are directly equated (Pontón & León-Carrión, 2001; Woodcock et al., 2001). In spring 2010, similar to the HSIS, children classified as Spanish-speaking DLLs according to the triangulated classification method were given the complete English version of the assessment battery and those scores were used as outcomes. Complete descriptive statistics for these scores are presented as well in Table 1.

3.2.4. Covariates

In order to increase the precision of the estimates of the associations between language of instruction and child achievement, several child and family covariates were included in all analyses. Child covariates included: gender; whether the child was classified as having a disability at baseline; and age in weeks at the spring assessment. Family covariates included: caregiver age in years; an indicator of caregiver depression; highest level of maternal education; and three family structure variables including whether both biological parents lived with the child, whether the child's mother was married, and whether the mother was a teenage mother at the child's birth. For purposes of analyses, all the covariates were centered at their mean. Descriptive statistics for all

covariates are displayed in Table 1.

3.2.5. Non-response

As with any longitudinal dataset, there was non-response in both HSIS and FACES-2009. To control for this potential bias, this study weighted all analyses, including descriptives and estimation models, using the appropriate weights, which were based on the probability of sample selection at every stage multiplied by adjustment for the probability of non-response. The weights included in the analyses are listed at the bottom of every table and helped control potential non-response bias by compensating for different data collection response rates across demographic groups of children. Weights are important in complicated multi-stage sampling studies such as the HSIS and FACES-2009 because they allow researchers to make inferences to the relevant general population, and they account for differential selection probabilities and differential non-response. The weights used in HSIS and FACES-2009 in particular ensured that the study samples were representative of the population of all newly-enrolled three- and four-year old children in Head Start at the time of their respective data collections (U.S. DHHS, 2010a, 2010b, Technical Report).

3.3. Analysis plan

The analysis for this study's research questions was a three-step process in which the analytic sample was restricted to Spanish-speaking DLL children in all cases. First, using FACES-2009, residualized growth models were used to estimate associations between Spanish language instruction and Spanish-speaking DLLs' English academic achievement. Specifically, estimated regressions tested for associations between Spanish language classroom instruction in Head Start and English receptive vocabulary, early literacy, and early math skills. A parallel non-experimental analysis was then performed using the HSIS to estimate the same associations for children who attended Head Start.

Thus, the final regression model for both FACES-2009 and the HSIS was:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 (SPANISH\ INSTRUCTION)_t + \gamma COVARIATES_{t-1} + e_t,$$

where Y_t was the outcome variable of interest after one year; Y_{t-1} was the lagged outcome variable at baseline; $SPANISH\ INSTRUCTION_t$ was a dummy variable indicating whether a child attended a Head Start center where they were taught in Spanish; $COVARIATES_{t-1}$ was a vector of additional demographic covariates; and e_t was an error term. Although this analysis was non-experimental, the use of both datasets enabled us to explore the robustness of the associations between Spanish language instruction and Spanish-speaking DLLs' English language academic school readiness skills.

To answer the second research question, this study used the entire Spanish-speaking DLL sample to estimate whether Head Start program impacts were differentially beneficial for Spanish-speaking DLLs instructed in Spanish. This analysis compared Spanish-speaking DLLs who attended Spanish instruction Head Start classrooms with DLL children who attended non-Spanish instruction Head Start classrooms as well as DLLs in the control condition. It thus estimated regressions including Spanish language instruction, random assignment to Head Start, and an interaction term between these variables on English receptive vocabulary, early literacy, and early math skills. Thus, the final interacted regression model was:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 (SPANISH\ INSTRUCTION)_t + \beta_3 TX + \beta_4 (SPANISH\ INSTRUCTION_t\ x\ TX) + \gamma COVARIATES_{t-1} + e_t,$$

where Y_t was the outcome variable of interest after one year; Y_{t-1} was the lagged outcome variable at baseline; $SPANISH\ INSTRUCTION_t$ was a dummy variable indicating whether a child attended an ECE center where they were taught in Spanish; TX was the dummy variable for random assignment to Head Start; $SPANISH\ INSTRUCTION_t\ x\ TX$ was the interaction of Spanish instruction and assignment to Head Start; $COVARIATES_{t-1}$ was the vector of additional demographic covariates; and e_t was an error term.

Given the strong potential for Head Start center-level variation, all of the models included Head Start center-level econometric fixed effects combined with standard errors calculated from replication methods, which properly adjust the point estimates and standard errors for weighting and clustering.

3.4. Hypothesized results

Prior research suggests that the language of classroom interaction and instruction may be important for Spanish-speaking DLL children's school readiness skills (Barnett et al., 2007; Burchinal et al., 2016; Chang et al., 2007; Durán et al., 2010; Páez et al., 2007). Moreover, Head Start mandates support promoting the linguistic context in which DLL children are served (U.S. DHHS, 2008). It was therefore hypothesized that Spanish language instruction would be positively associated with DLL children's English academic skills of early literacy and early math in HSIS and FACES-2009 due to the better match between the classroom language environment and children's language capabilities. That is, because children's home language is supported in the classroom as opposed to being linguistically isolated and due to some evidence of cross linguistic transfer in emergent literacy (August & Shanahan, 2006; Castro et al., 2011; Dickinson, McCabe, & Clark-Chiarelli, & Wolf, A., 2004; Rinaldi & Páez, 2008), Spanish language instruction would help DLL children have higher English outcomes compared with DLL children educated in English-only contexts. Because of the ambiguity in the research on oral language, it was unclear what the expected direction of the results should be for receptive vocabulary. Using the HSIS data, it was further hypothesized that DLL children who attended Head Start would also benefit differentially on these academic domains from Spanish language classroom instruction compared with DLLs in Head Start where Spanish instruction was not used as well as DLLs in the control condition. It was hypothesized that Spanish instruction would increase the effectiveness of Head Start on English outcomes for this growing population.

4. Results

4.1. Research question 1

Results for the first research question on whether there were associations between Spanish language instruction and Spanish-speaking DLL children's English academic school readiness skills are displayed in Table 2. In HSIS, results indicate that Spanish-speaking DLL children instructed in Spanish in Head Start scored about 1/6 *SD* higher on English receptive vocabulary ($\beta = 0.16$, $p < 0.001$) compared with DLL children not instructed in Spanish. Interestingly, Spanish instruction in Head Start was not associated with the other English assessment outcomes of early literacy or early math – in fact, estimated associations were negative although statistically insignificant. Similarly, in FACES-2009, Spanish-speaking DLL children instructed in Spanish in Head Start scored 1/5 *SD* higher on English receptive vocabulary ($\beta = 0.20$, $p < 0.05$) compared with DLL children not instructed in Spanish, and there were also null findings for the English outcomes of early literacy and early math.

Because this study analyzed two independent samples of Head Start children collected seven years apart with varying sample sizes, it supported a test of whether the coefficient estimates from HSIS and FACES-2009 were significantly different from each other using a two-sample test of independent means to see if the results from each dataset individually replicated across both. The *p*-value of the HSIS/FACES-2009 difference for English receptive vocabulary skills was 0.71. Because this value was far from statistically significant, results for the first research

Table 2
Regressions relating Spanish classroom instruction with academic outcomes in HSIS and FACES-2009.

	Outcomes							
	HSIS			FACES-2009			*p*-Value of HSIS/FACES-2009 difference	
	Receptive vocabulary:	Early literacy:	Early math:	Receptive vocabulary:	Early literacy:	Early math:		
	PPVT	WJ letter-word	WJ applied problems	PPVT	WJ letter-word	WJ applied problems		
Teach children in Spanish	0.16***	− 0.06	− 0.04	0.20*	− 0.20	− 0.05	0.71	WJ-LW: 0.16
	(0.06)	(0.05)	(0.03)	(0.09)	(0.11)	(0.04)		WJ-AP: 0.84
N	590	590	590	472	472	472		
R2	0.58	0.52	0.51	0.31	0.43	0.48		

Note. Outcomes are standardized. Weight-adjusted standard errors in parentheses.
PPVT = Peabody Picture Vocabulary Test. WJ = Woodcock-Johnson. Teach in SPanish is teacher report. Head Start center-level fixed effects included in all models. Demographic covariates (centered at mean): baseline achievement, child cohort, child gender, child disability status, maternal education, maternal marital status, caregiver depression, teenage mother status, caregiver age, and child age at spring assessment. HSIS weight used = CHSPR2003WTCA; FACES-2009 weight used = P21RA2WT.
* $p < 0.05$.
*** $p < 0.001$.

question appear to be robust across the two independent Head Start samples of HSIS and FACES-2009.

4.2. Research question 2

Results for the second research question on whether Head Start program impacts were differentially beneficial for Spanish-speaking DLL children instructed in Spanish are displayed in Table 3. When interaction variables are included in the models, the main effects of Spanish language instruction are those of control group children only. Consequently, for each outcome it is necessary to discuss the results for control group children first and then the Head Start children.

Table 3
Regressions relating Spanish classroom instruction, assignment to Head Start, and their interaction in HSIS.

	Outcomes		
	HSIS		
	Receptive vocabulary:	Early literacy:	Early math:
	PPVT	WJ Letter-Word	WJ applied problems
Teach children in Spanish (dummy)	0.11	0.03	− 0.05
	(0.07)	(0.08)	(0.06)
Treatment (dummy assignment to H.S.)	0.50**	0.26	0.32***
	(0.15)	(0.16)	(0.10)
Teach children in Spanish x treatment	0.17***	− 0.04	0.02
	(0.04)	(0.03)	(0.02)
N	935	935	935
R2	0.58	0.51	0.50

PPVT = Peabody Picture Vocabulary Test. WJ = Woodcock-Johnson. Teach in Spanish is teacher report. Head Start center-level fixed effects included in all models. Demographic covariates (centered at mean): baseline achievement, child cohort, child gender, child disability status, maternal education, maternal marital status, caregiver depression, teenage mother status, caregiver age, child age at spring 2003 assessment, number of weeks elapsed between 09/01/02 and spring 2003 child assessment. HSIS weight used = CHSPR2003WTCA.
Note. Outcomes are standardized. Standard errors in parentheses calculated using jackknife replicate weights.
** $p < 0.01$.
*** $p < 0.001$.

For English receptive vocabulary, there was a positive but statistically insignificant main effect of Spanish instruction for Spanish-speaking DLL control group children. Further, Spanish-speaking DLL children who attended Head Start and were not taught in Spanish scored ½ *SD* higher on English receptive vocabulary skills compared with control group children ($\beta = 0.50$, $p < 0.01$). Additionally, there was a significant interaction between Spanish language and instruction and random assignment to Head Start on English receptive vocabulary skills as well ($\beta = 0.17$, $p < 0.001$) in that Spanish-speaking DLL children who attended Head Start and were instructed in Spanish scored higher on English receptive vocabulary skills than both Spanish-speaking DLL children who attended Head Start but were *not* instructed in Spanish and Spanish-speaking DLL children in the control condition who did not attend Head Start. As with the first research question, Spanish instruction in Head Start was not associated with the other English assessment outcomes of early literacy or early math.

To better understand this interaction on English receptive vocabulary, Fig. 1 shows predicted scores computed from the coefficient estimates in Table 3. As indicated by the three different groups, Spanish-speaking DLL children who attended Head Start and were instructed in Spanish had the highest English receptive vocabulary scores of all the

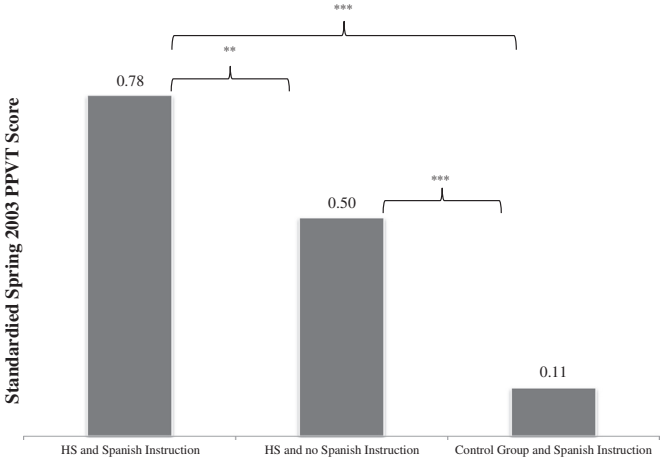

Fig. 1. Predicted standardized spring 2003 PPVT scores by Spanish language instruction and random assignment to Head Start. The fourth group, Control Group and no Spanish Instruction, is the omitted category. (** $p < 0.01$. *** $p < 0.001$.)

groups. Their scores were significantly higher than those of DLL children who attended Head Start but were not instructed in Spanish as well as those of DLL children in the control condition who did not attend Head Start.

4.3. Robustness checks

It was possible that the teachers in HSIS and FACES-2009 who reported using Spanish for instruction in Head Start classrooms had higher markers of structural quality than those who did not report such classroom use of Spanish, such as more education in general or certification to teach specifically in early childhood (Howes et al., 2008). Though such markers were not a priori thought to be directly associated with the proximal processes of language interactions between teachers and children, because the findings from this study offered a somewhat novel result than prior research on the benefits of Spanish language instruction for Spanish-speaking DLL children's oral language skills, robustness checking was necessary.

Therefore, as a key robustness check, this study tested several teacher qualifications relating to structural quality that might be confounding the results of this study to examine if they differed by whether the teacher reported instructing children in Spanish in both Head Start datasets. The qualifications tested included: total years teaching, total years teaching in Head Start, highest level of teacher education, whether the teacher was certified or licensed to teach, whether the teacher had a degree in early childhood or child development or any coursework in these respective fields, and whether the teacher had a Child Development Associate (CDA) credential. Results are displayed in Appendix A and indicate that in both HSIS and FACES-2009, there were no significant differences either within each dataset or across datasets on any of these qualifications as to whether the teacher reported instructing DLL children in Spanish. As a further robustness check, we included all of the teacher characteristics in the main analytic models and interacted them with treatment in the HSIS sample in order to ensure that the language interaction with Head Start was not picking up the effects of these predictors. Our primary interaction variable of Spanish language instruction retained its significance in this fully-interacted model: $\beta = 0.15$, $p < 0.01$. Therefore, these robustness checks as well as the replication across both Head Start datasets provided confidence that these observable characteristics of teachers were not biasing the findings of this study, though it cannot preclude other unmeasured or unobserved skills.

5. Discussion

This study used the HSIS and FACES-2009 datasets to determine the association between Spanish language instruction and Spanish-speaking DLLs' English academic achievement skills. It examined this question within each of the two datasets individually, and then additional analyses were conducted to determine if the results replicated across both datasets of HSIS and FACES-2009. It further sought to understand whether Head Start program impacts differed based on such Spanish language instruction. Given the positive impacts of ECE for Spanish-speaking DLLs (Buysse et al., 2014; Gormley, 2008; Loeb et al., 2007; U.S. DHHS, 2010a), instruction in Spanish may be an important pathway through which DLL children can maximize their learning experiences in child care.

Although it was hypothesized that Spanish language instruction in Head Start would be beneficial for DLLs' English early literacy and early math skills, in fact there was no such positive association in either dataset; there appeared to be no added benefit to instructing DLL children in Spanish on these skills though it also did not compromise their performance. Somewhat surprisingly, however, such Spanish language instruction in Head Start was positively associated with English receptive vocabulary skills. That is, Spanish-speaking DLL children instructed in Spanish had significantly higher English PPVT scores than those DLL children not instructed in Spanish in both HSIS and FACES-2009. Further, in the regressions that took advantage of the experimental variation in random assignment to Head Start, Spanish-speaking DLL children instructed in Spanish in Head Start also had significantly higher English receptive vocabulary scores than either DLL children in Head Start who were not instructed in Spanish or DLL children in the control condition who did not attend Head Start.

The results of this study are rather unexpected in light of previous research that Spanish-speaking DLLs tend to lag behind monolingual-English speakers particularly in English oral language skills such as vocabulary (Hoff, 2013; McCabe et al., 2013; Páez et al., 2007), particularly in a low-income sample of Head Start-eligible children, although there is other recent empirical work that converges with the findings of this study. Burchinal et al. (2016) for example, also found some evidence of cross-linguistic transfer for low-income Spanish-speaking DLL children's English receptive vocabulary skills. This study therefore suggests that there may be something fundamentally different about the way Spanish instruction impacts DLLs' receptive vocabulary skills in Head Start compared with other domains, including early literacy. Perhaps when teachers use Spanish for instruction and particularly for translation purposes, it helps children better understand the meanings of English words, thus expanding their English vocabulary. In addition, hearing more language generally and in both languages may promote better oral language skills overall through conceptual understanding independent of the language in which it occurs (Pearson, 1998). This is consistent with prior research as Core et al. (2013) and Hoff (2013) who found that when total language was considered, DLL children performed on par or higher than monolingual-English speakers on measures of oral language, as well as consistent with theory in that early language experiences in one language support and privilege later literacy experiences in another (Uccelli et al., 1999).

Conversely, with other early skills such as emergent literacy, perhaps instructing children in Spanish was not necessary to promote these skills due to the high degree of cross-linguistic transfer in areas such as phonological awareness and letter-word knowledge (August & Shanahan, 2006; Castro et al., 2011; Dickinson et al., 2004; Rinaldi & Páez, 2008). That is, because of the likelihood of cross-linguistic transfer in these key literacy areas, and given that many DLL children likely have a high degree of exposure to English before coming to Head Start from sources like television and media even if they use Spanish exclusively in the home, there was no added benefit for the teacher instructing the children in Spanish on these English outcome domains. In fact, Spanish-speaking DLL children perform on par with monolingual English-speaking children on measures of emergent literacy (Páez et al., 2007; Rinaldi & Páez, 2008).

Some observational qualitative evidence supports this. In a study of four local Head Start classrooms in a large urban, predominantly Latino county in the Southwest U.S., observations indicated that both when introducing children to new vocabulary as well as using key established vocabulary in the classroom, teachers always said the words in one language and then deliberately translated into the other. For example, the teacher would ask, 'who has to go to the bathroom? *Baño* anyone?' or when reading *The Very Hungry Caterpillar*, 'who knows what sausage is? Sausage is *chorizo*.' Thus, such explicit translation provided Spanish-speaking DLL children with more classroom opportunities to make connections to the vocabulary and to increase their overall oral language skills through conceptual understanding independent of the language in which in occurred (Miller, 2017a, 2017b).

On the other hand, when teachers focused on letter-word names and sounds as well as early math, these activities mostly took place only in English. On the rare occasion that the teacher conducted these activities in Spanish, the children responded only in English (Miller, 2017a, 2017b). Future work should continue to study how classroom language use differs across language and literacy domains, as this appears to be an important source of variation for Spanish-speaking DLL children's English school readiness skills. Future research should also include an

analysis of the effects of Spanish language instruction on Spanish outcomes, which was not possible in this study, given that there may be benefits to instructing children in Spanish on Spanish assessment measures (e.g., Barnett et al., 2007; Burchinal et al., 2016; Durán et al., 2010; Páez et al., 2007).

5.1. Limitations and future directions

Some study limitations should be noted. First, although the second study research question was able to take advantage of the experimental nature of the HSIS, this was not possible for the first research question and the results reported for this question are not causal. Nonetheless, because the findings replicated across two large, independent samples of Head Start children, these results are likely to be more robust than would be possible from either sample alone. In addition, the supplementary robustness checks helped ensure that the findings were not attributable to other observable teacher characteristics. Thus, these data are the best currently available to address the research questions asked in this study.

Second, the purpose of this paper was to determine whether there were associations between Spanish language instruction and DLL children's English academic school readiness, and whether Head Start was differentially beneficial for children based on such use. The variable used to ascertain these results was crude and potentially masked tremendous variability in actual Spanish use in the classroom. However, its dichotomous nature likely resulted in measurement error biasing the coefficient toward zero, and as such this variable was the best available, albeit limited and conservative, estimate for understanding these classroom language dynamics. Nonetheless, the findings from this study therefore raise an important issue of how much Spanish instruction actually occurs in the classroom and under which conditions. This is a tremendous gap in the literature. Research studies building capacity to answer questions on DLL children go together with programmatic efforts to improve services for these children (U.S. DHHS, 2013b). Therefore, future work can answer more detailed questions about how much Spanish occurs in classrooms and when, particularly if national data sets collect more detailed information on classroom Spanish use, and new qualitative studies (e.g., Miller, 2017a, 2017b) continue to describe these language dynamics in more detail.

Given that Spanish-speaking DLL children are the fastest growing demographic in the U.S. as well as an increasing share of Head Start and Early Head Start participants (nearly 40%; U.S. DHHS, 2014), the results of this paper offer implications for policy and practice. The findings can help centers respond to Head Start mandates (U.S. DHHS and ED, 2016; U.S. DHHS, 2008) to support DLL children's home language and provide services to families in culturally responsive ways such as bilingual teacher and staff hiring, classroom language supports, and curriculum decisions that stress the importance of both languages.

In sum, this study investigated whether Spanish language classroom instruction was beneficial for Spanish-speaking DLL children's English school readiness. It extended prior research that has examined this relationship for various language and literacy outcomes and is the first to use the two largest nationally representative samples of Head Start – HSIS and FACES-2009 – to gain an understanding of how using Spanish language instruction relates to English academic achievement. Consistent with bioecological theory (Bronfenbrenner & Morris, 2006), the findings from this study reveal that there are different associations depending on the outcome domain of interest. As a somewhat novel finding, it appears that Head Start with Spanish instruction does a particularly good job at promoting DLL children's oral language skills, as it explicitly uses both the home language and English to target new and existing key vocabulary. Given the link between oral language proficiency and future literacy, academic success, and behavior (Spira et al., 2005), Head Start appears to be responding well to the needs of this fast-growing and important demographic of children and families.

Appendix Table A1. Teacher qualifications by Spanish language classroom instruction for Spanish-speaking DLL children within each dataset.

	Outcomes							
	HSIS: classroom instruction in Spanish				FACES-2009: classroom instruction in Spanish			
	Yes (N = 711)		No (N = 430)		Yes (N = 463)		No (N = 362)	
	Mean/% of sample	SD	Mean/% of sample	SD	Mean/% of sample	SD	Mean/% of sample	SD
Total years teaching	13.01	9.81	11.75	7.62	11.93	7.78	12.18	8.39
Total years teaching in head start	7.74	6.89	6.70	5.88	8.10	6.33	8.52	7.35
Highest level of education completed								
Less than college	0.29		0.25		0.22		0.26	
College graduate	0.64		0.65		0.69		0.67	
graduate degree	0.06		0.03		0.09		0.08	
Teaching certificate/license	0.76		0.69		0.54		0.48	
Degree in early childhood	0.91		0.90		0.60		0.65	
Coursework in early childhood	0.96		0.96		0.91		0.92	
Child development associate credential	0.58		0.61		0.50		0.44	

Note. HSIS weight used = S03TRWTCA_PI_TS; FACES-2009 weight used = T12TCHWT.

Appendix Table A2. Teacher qualifications by Spanish language classroom instruction for Spanish-speaking DLL children comparing across datasets.

	Outcomes

	Classroom instruction in Spanish					No classroom instruction in Spanish				
	HSIS (N = 711)		FACES-2009 (N = 463)		*p*-Value of HSIS/ FACES-2009 difference	HSIS (N = 430)		FACES-2009 (N = 362)		*p*-Value of HSIS/ FACES-2009 difference
	Mean/% of sample	SD	Mean/% of sample	SD		Mean/% of sample	SD	Mean/% of sample	SD	
Total years teaching	13.01	9.81	11.93	7.78	0.55	11.75	7.62	12.18	8.39	0.48
Total years teaching in head Start	7.74	6.89	8.10	6.33	0.48	6.70	5.88	8.52	7.35	0.39
Highest Level of education completed										
Less than college	0.29		0.22		0.57	0.25		0.26		0.54
College graduate	0.64		0.69		0.46	0.65		0.67		0.48
Graduate degree	0.06		0.09		0.45	0.03		0.08		0.47
Teaching certificate/ license	0.76		0.54		0.32	0.69		0.48		0.38
Degree in early childhood	0.91		0.60		0.20	0.90		0.65		0.22
Coursework in early childhood	0.96		0.91		0.39	0.96		0.92		0.38
Child development associate credential	0.58		0.50		0.57	0.61		0.44		0.36

Note. HSIS weight used = S03TRWTCA_PI_TS; FACES-2009 weight used = T12TCHWT.

References

August, D., & Shanahan, T. (Eds.), (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth.* Mahwah, NJ: Erlbaum.

August, D. L., & Hakuta, K. (1997). *Improving schooling for language-minority children.* Washington, DC: National Academy Press.

Barnett, W. S. (2011). Effectiveness of educational intervention. *Science*, *333*(6045), 975–978. http://dx.doi.org/10.1126/science.1204534.

Barnett, W. S., Yarosz, D. J., Thomas, J., Jung, K., & Blanco, D. (2007). Two-way and monolingual English immersion in preschool education: An experimental comparison. *Early Child Research Quarterly*, *22*(3), 277–293. http://dx.doi.org/10.1016/j.ecresq.2007.03.003.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition.* New York, NY: Cambridge University Press.

Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology*, *65*(4), 229–235. http://dx.doi.org/10.1037/a0025406.

Bloom, H., & Weiland, C. (2015). Quantifying variation in head start effects on young children's cognitive and socio-emotional skills using data from the National Head Start Impact Study. *MDRC working paper*. Retrieved from http://www.mdrc.org/sites/default/files/quantifying_variation_in_head_start.pdf.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner, & W. Damon (Vol. Eds.), *Handbook of child psychology: Theoretical models of human development*(6th ed.). *Vol. 1. Handbook of child psychology: Theoretical models of human development* (pp. 793–828). New York: Wiley.

Burchinal, M., Soliday Hong, S., LaForett, D., Hammer, C., Páez, M., ... Castro, D. (2016). *Change in Language and Literacy Knowledge for Spanish-English Dual Language Learners at School Entry: Analyses From Three Studies. Manuscript Submitted for Publication*.

Buysse, V., Peisner-Feinberg, E., Páez, M., Hammer, C. S., & Knowles, M. (2014). Effects of early education programs and practices on the development and learning of dual language learners: A review of the literature. *Early Child Research Quarterly*, *29*(4), 765–785. http://dx.doi.org/10.1016/j.ecresq.2013.08.004.

Calderón, M., Slavin, R., & Sánchez, M. (2011). Effective instruction for English learners. *The Future of Children*, *21*(1), 103–127. http://dx.doi.org/10.1353/foc.2011.0007.

Castro, D. C., Páez, M. M., Dickinson, D. K., & Frede, E. (2011). Promoting language and literacy in young dual language learners: Research, practice and policy. *Child Development Perspectives*, *5*(1), 15–21. http://dx.doi.org/10.1111/j.1750-8602.2010.00142.x.

Chang, F., Crawford, G., Early, D., Bryant, D., Howes, C., Burchinal, M., ... Pianta, R. (2007). Spanish-speaking children's social and language development in pre-kindergarten classrooms. *Early Education and Development*, *18*(2), 243–269. http://dx.doi.org/10.1080/10409280701282959.

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*(5), 1637–1649. http://dx.doi.org/10.1044/1092-4388(2013/11-0044).

Dickinson, D. K., McCabe, A., & Clark-Chiarelli, & Wolf, A. (2004). Cross-language transfer of phonological awareness in low-income Spanish and English bilingual preschool children. *Applied PsychoLinguistics*, *25*(3), 323–347. http://dx.doi.org/10.1017/S0142716404001158.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... Japel, C. (2007). School readiness and later achievement. *Developmental Psychology*, *43*(6), 1428–1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428.

Duncan, S. E., & DeAvila, E. (1998). *Preschool language assessment survey 2000.* Monterey, CA: CTB/McGraw-Hill.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test, fourth edition (PPVT-4).* Minneapolis, MN: Pearson Assessments.

Dunn, L. M., & Dunn, L. L. (1997). *Peabody picture vocabulary test, third edition (PPVT-3).* Circle Pines, MN: American Guidance Services.

Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *TVIP – Test Vocabulario en Imagenes Peabody.* Circle Pines, MN: American Guidance Service.

Durán, L. K., Roseth, C. J., & Hoffman, P. (2010). An experimental study comparing English-only and transitional bilingual education on Spanish-speaking preschoolers' early literacy development. *Early Child Research Quarterly*, *25*(2), 207–217. http://dx.doi.org/10.1080/15235882.2012.735213.

Espinosa, L. M. (2013). *Early education for dual language learners: Promoting school readiness and early school success.* Washington, DC: Migration Policy Institute.

Goldenberg, C. (2012). Research on English learner instruction. In M. Calderón (Ed.), *Breaking through: Effective instruction & assessment for reaching English learners* (pp. 39–61). Bloomington, IN: Solution Tree.

Goldenberg, C., Nemeth, K., Hicks, J., Zepeda, M., & Cardona, L. M. (2013). Program elements and teaching practices for young dual language learners. *California's best practices for young dual language learners: Research overview papers*California Department of Education (CDE): Child Development Division. Retrieved from http://www.cde.ca.gov/sp/cd/ce/documents/dllresearchpapers.pdf.

Gormley, W. T. (2008). The effects of Oklahoma's pre-k program on Hispanic children. *Social Science Quarterly*, *89*(4), 916–936. http://dx.doi.org/10.1111/j.1540-6237.2008.00591.x.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H. Brookes.

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4–14. http://dx.doi.org/10.1037/a0027238.

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Child Research Quarterly*, *23*(1), 27–50. http://dx.doi.org/10.1016/j.ecresq.2007.05.002.

Karoly, L. A., Kilburn, M. R., & Cannon, J. S. (2005). *Early childhood interventions: Proven results, future promise.* Santa Monica, CA: Rand Corporation.

Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Research*, *39*(6), 484–486 (doi:10.1.1.172.7069).

Kuhl, P. K. (2009). Early language acquisition: Neural substrates and theoretical models. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 837–854). (4th ed.). Cambridge, MA: Massachusetts Institute of Technology.

Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too

much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, *26*(1), 52–66. http://dx.doi.org/10.1016/j.econedurev.2005.11.005.

López, M. H., & Velasco, G. (2011). *Childhood poverty among Hispanics sets record, leads nation*. Washington, D.C.: Pew Hispanic Center, Pew Research Center. Retrieved from http://www.pewhispanic.org/2011/09/28/childhood-poverty-among-hispanics-setsrecord-leads-nation.

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, *26*(1), 33–51. http://dx.doi.org/10.1016/j.econedurev.2005.09.008.

Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development*, *82*(5), 1544–1560. http://dx.doi.org/10.1111/j.1467-8624.2011.01633.x.

Mancilla-Martinez, J., & Lesaux, N. K. (2014). Bilingual teachers in early head start/head start. *NHSA Dialog: A Research-to Practice Journal for the Early Childhood Field*, *17*(1), 108–112.

Mancilla-Martinez, J., & Vagh, S. B. (2013). Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Child Research Quarterly*, *28*(3), 555–567. http://dx.doi.org/10.1016/j.ecresq.2013.03.004.

McCabe, A., Tamis-LeMonda, C. S., Bornstein, M. H., Cates, C. B., Golinkoff, R., Guerra, A. W., ... Song, L. (2013). Multilingual children: Beyond myths and towards best practices. *SRCD Social Policy Report*, *27*(4), 1–21.

Miller, E. B. (2017a). How Spanish is used in head start: Observational evidence from four classrooms. *NHSA Dialog: The Research-to- Practice Journal for the Early Childhood Field*, *19*(3), 1–26.

Miller, E. B. (2017b). Four classrooms model how teachers use Spanish in head start. *NHSA Dialog: The Research-to- Practice Journal for the Early Childhood Field*, *19*(3), 83–93.

National Assessment of Educational Progress (2013). The nation's report card. Retrieved from http://nces.ed.gov/nationsreportcard/subject/publications/main2013/pdf/2014451.pdf.

National Clearinghouse for English Language Acquisition (2011). The growing numbers of English Learner Students, 1998/99-2008/09. Retrieved from http://www.ncela.gwu.edu/files/uploads/9/growing_EL_0910.pdf.

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.

Nelson, C. A., & Sheridan, M. A. (2011). Lessons from neuroscience research for understanding causal links between family and neighborhood characteristics and educational outcomes. In G. Duncan, & R. Murnane (Eds.), *Whither opportunity: Rising inequality, schools, and children's life chances* (pp. 27–46). New York: Russell Sage.

Páez, M. M., Tabors, P. O., & López, L. M. (2007). Dual language and literacy development of Spanish-speaking preschool children. *Journal of Applied Developmental Psychology*, *28*(2), 85–102. http://dx.doi.org/10.1016/j.appdev.2006.12.007.

Pearson, B. Z. (1998). Assessing lexical development in bilingual babies and toddlers. *International Journal of Bilingualism*, *2*(3), 347–372 (doi:101177/136700699800200305).

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. http://dx.doi.org/10.1111/j.1467-1770.1993.tb00174.x.

Pontón, M. O., & León-Carrión, J. (2001). *Neuropsychology and the Hispanic patient: A clinical handbook*. Mahwah, NJ: Lawrence Erlbaum Associates.

Ramey, C. T., & Ramey, S. L. (2006). Early learning and school readiness: Can early intervention make a difference? In N. F. Watt, C. Ayoub, R. H. Bradley, J. E. Puma, & A. LeBoeuf (Eds.), *The crisis in youth mental health* (pp. 291–318). Westport, CT: Praeger.

Rinaldi, C., & Páez, M. (2008). Preschool matters: Predicting reading difficulties for Spanish-speaking students in first grade. *Learning Disabilities: A Contemporary Journal*, *6*(1), 71–84.

Rossell, C. H., & Kuder, J. (2005). Meta-murky: A rebuttal to recent meta-analyses of bilingual education. In J. Söhn (Ed.), *The effectiveness of bilingual school programs for immigrant children* (pp. 43–76). Berlin, Germany: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Schweinhart, L. J. (2006). The high/scope approach: Evidence that participatory learning in early childhood contributes to human development. In N. F. Watt, C. Ayoub, R. H. Bradley, J. E. Puma, & W. A. LeBoeuf (Eds.), *The crisis in youth mental health* (pp. 207–227). Westport, CT: Praeger.

Slavin, R. E., & Cheung, A. (2005). A synthesis of research on language of reading instruction for English language learners. *Review of Educational Research*, *75*(2), 247–281. http://dx.doi.org/10.3102/00346543075002247.

Slavin, R. E., Madden, N., Calderón, M., Chamberlain, A., & Hennessy, M. (2011). Reading and language outcomes of a multiyear randomized evaluation of transitional bilingual education. *Educational Evaluation and Policy Analysis*, *33*(1), 47–58 (doi:http://dx.doi.org/10.3102/0162373711398127).

Spira, E. G., Bracken, S. S., & Fischel, J. E. (2005). Predicting improvement after first-grade reading difficulties: The effects of oral language, emergent literacy, and behavior skills. *Developmental Psychology*, *41*(1), 225–234. http://dx.doi.org/10.1037/0012-1649.41.1.225.

Sunderman, G., & Kroll, J. F. (2006). First language activation during second language lexical processing: An investigation of lexical form, meaning, and grammatical class. *Studies in Second Language Acquisition*, *28*(3), 387–422. http://dx.doi.org/10.1017/S0272263106060177.

Tabors, P. O. (2008). *One child, two languages*. Baltimore, MD: Paul H. Brookes.

U.S. Department of Health and Human Services, & Administration for Children and Families (2010a). *Head start impact study*. Washington, DC: Author (Final report).

U.S. Department of Health and Human Services, & Administration for Children and Families (2010b). *Head start impact study*. Washington, DC: Author (Technical report).

U.S. Department of Health and Human Services, & Administration for Children and Families (2011). *FACES 2009 study design. (OPRE Report 2011-9)*. Washington, DC: Author.

U.S. Department of Health and Human Services, & Administration for Children and Families (2013a). *Head start family and child experiences survey (FACES)*. Washington, DC: Author (2009 Cohort. User guide).

U.S. Department of Health and Human Services, & Administration for Children and Families (2013b). *Report to Congress on dual language learners in Head Start and Early Head Start programs*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start (2008). *Dual language learning: What does it take? Head start dual language report*. Washington, DC: Author.

U.S. Department of Health and Human Services, Administration for Children and Families, & Office of Head Start (2014). Head start program fact sheet, Fiscal year 2014 (Retrieved from) http://eclkc.ohs.acf.hhs.gov/hslc/data/factsheets/docs/hs-program-fact-sheet-2014.pdf.

U.S. Department of Health and Human Services, Administration for Children and Families, & Office of Planning, Research and Evaluation (2002–2006). *Head start impact study*. (Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2012-06-15. doi:http://dx.doi.org/10.3886/ICPSR29462.v4 ICPSR29462-v4.).

U.S. Department of Health and Human Services, Administration for Children and Families, & Office of Planning, Research and Evaluation (2009). *Head start family and child experiences survey (FACES)*. (Cohort. (2009–2013). ICPSR34558-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. doi:http://dx.doi.org/10.3886/ICPSR34558.v1).

U.S. Department of Health and Human Services, & U.S. Department of Education (2016). *Policy statement on supporting the development of children who are Dual Language Learners in early childhood programs*. Washington, DC: Authors.

Uccelli, P., Hemphill, L., Pan, B. A., & Snow, C. (1999). Telling two kinds of stories: Sources of narrative skill. In L. Butler, & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 215–233). Philadelphia, PA: Psychology Press.

Umbel, V. M., Pearson, B. Z., Fernandez, M. C., & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, *63*(4), 1012–1020. http://dx.doi.org/10.1111/j.1467-8624.1992.tb01678.

Vitiello, V. E., Downer, J. T., & Williford, A. P. (2011). Preschool classroom experiences of dual language learners: Summary of findings from publicly funded programs in 11 states. In C. Howes, J. T. Downer, & R. C. Pianta (Eds.), *Dual language learners in the early childhood classroom* (pp. 69–92). Baltimore, MD: Brookes.

Whitehurst, G. J., & Lonigan, C. J. (2003). Emergent literacy: Development from pre-readers to readers. In S. B. Neuman, & D. K. Dickinson (Vol. Eds.), *Handbook of early literacy research. Vol. 1. Handbook of early literacy research* (pp. 11–29). New York, NY: Guilford.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Woodcock, R. W., & Muñoz-Sandoval, A. F. (1996). *Bateriá Woodcock-Muñoz Pruebas de Aprovechamiento-Revisada*. Itasca, IL: Riverside Publishing.

Yesil-Dagli, U. (2011). Predicting ELL students' beginning first grade English oral reading fluency from initial kindergarten vocabulary, letter naming, and phonological awareness skills. *Early Child Research Quarterly*, *26*(1), 15–29. http://dx.doi.org/10.1016/j.ecresq.2010.06.001.

Zigler, E., & Styfco, S. J. (2010). *The hidden history of head start*. New York, NY: Oxford University Press.