Early Education and Development

ISSN: 1040-9289 (Print) 1556-6935 (Online) Journal homepage: https://www.tandfonline.com/loi/heed20

Highly Potent and Vastly Conditional Instructional Practices: Variations in Use and Utility of Language Interactions for Kindergarten

Sabina Neugebauer, Lia Sandilos, Michael Coyne, D. Betsy McCoach & Sharon Ware

EARLY EDUCATION AND DEVELOPMENT
2020, VOL. 31, NO. 4, 541–560
https://doi.org/10.1080/10409289.2019.1686928

Routledge
Taylor & Francis Group

Check for updates

Highly Potent and Vastly Conditional Instructional Practices: Variations in Use and Utility of Language Interactions for Kindergarten

Sabina Neugebauer[a], Lia Sandilos[b], Michael Coyne[c], D. Betsy McCoach[c], and Sharon Ware[d]

[a]Teaching and Learning, Temple University; [b]Psychological Studies in Education, Temple University; [c]Educational Psycholgy, University of Connecticut; [d]Special Education, University of Saint Joseph

ABSTRACT

Research Findings: This study of kindergarten monolingual and Dual Language Learners (DLLs) (N=263; mean age = 63.40 months) and their teachers (N=27) found significant mean differences in the observed quality of language interactions in classrooms with differing proportions of DLLs. Teachers instructing in classrooms with higher numbers of DLLs had lower levels of concept development and quality of feedback. Moderation findings revealed that interactions supporting concept development were particularly beneficial for DLLs' vocabulary growth. *Practice and Policy:* Study results underscore the importance of using language interaction practices that cultivate higher-order thinking with DLLs.

In the United States, the population of Dual Language Learners (DLLs), students who speak a language other than English at home (Gutiérrez, Zepeda, & Castro, 2010; Park, O'Toole, & Katsiaficas, 2017), has increased exponentially, with more than 4.6 million DLLs enrolled in public schools (National Center for Education Statistics [NCES], 2017). Many public school teachers will have at least one, and likely more, DLLs in their class and thus must support the academic development of these learners (Ballantyne, Sanderman, & Levy, 2008). A teacher's ability to tailor instruction to the needs of DLLs is critical given recent statistics that indicate that, on average, DLLs continue to fall behind their monolingual peers in their language and literacy performance (Donovan & Cross, 2002; Mulligan, McCarroll, Flanagan, & Potter, 2016), with significant consequences for their long-term academic trajectories (August & Hakuta, 1997; NCES, 2004).

Primary teachers, in particular, have an important role to play in supporting DLLs language development. DLL students, on average, begin formal schooling with lower vocabularies than their peers (NCES, 2017). Without instruction that can close this vocabulary gap these language differences can multiply over time with consequences for later literacy performance (Mancilla-Martinez & Lesaux, 2011). Given wide-spread recognition that early language development and vocabulary, specifically, are core predictors of later reading and academic success broadly (Ouellette & Beers, 2010), schools are increasingly adopting research-based language and literacy curricula intended to boost the vocabulary development of students, especially those students at elevated risk of language and literacy difficulties, including DLLs (Borman et al., 2005b). However, despite the promise of these programs, student outcomes in these programs are varied, particularly in urban schools serving students with lower language skills and high populations of DLLs (Chien et al., 2010; Justice, Mashburn, Hamre, & Pianta, 2008).

One explanation for the varied success of school-wide programs intended to boost language and literacy development is the variability in quality student-teacher interactions across these contexts

CONTACT Sabina Neugebauer Sabina.neugebauer@temple.edu College of Education, 1301 Cecil B. Moore Avenue, Philadelphia, PA 19122-6091

and student populations. These prescribed but unscripted programs provide a sequence of words, set of activities, and language-rich materials, yet language learning occurs not only as a function of the content covered but also based on the nature of student-teacher interactions about content (Connor, Morrison, & Slominski, 2006; Dickinson, Flushman, & Freiberg, 2009; Gibbons, 2010). Indeed, a wealth of research documents the essential role of teacher-student interactions for bolstering students' language skills and academic success (Pianta et al., 2005; Zaslow, Martinez-Beck, Tout, & Halle, 2011). Notably, these interactions are particularly challenging to script given their dynamic nature (leaving their use at the discretion of the teacher) and less is known about the efficacy of these interactions for linguistically diverse student populations (Downer et al., 2012). The majority of the existing research on language interactions in the primary grades focuses on monolingual English-speaking children and their interactions with teachers (Dickinson, Darrow, & Tinubu, 2008; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Justice et al., 2008).

Thus, the present study explored ways in which teachers interact with students during their planned language and literacy curriculum, and in particular, how these interactions may differ in classrooms with DLLs. In addition, this study examines if certain types of interactions are differentially associated with vocabulary development for DLLs and their monolingual peers. This investigation is essential for better understanding the language environments and interactions experienced by this population of learners at a crucial time in their language development.

Teaching DLLs Content and Language through Quality Language Interactions

In this work, we adopt a socio-cultural perspective of learning. Drawing from the work of Vygotsky (1978), this perspective posits that all learning is social in nature and that children learn language as well as content via language interactions with adults or knowledgeable peers. A core facet of this approach is that learning occurs as part of an apprenticeship model, in which a novice internalizes the skilled habits of an expert through interaction–often centred on linguistic communication. These interactions promote a gradual release of responsibility from the teacher to the student (Pearson & Gallagher, 1983), helping the student to gain mastery of a new skillset. In practice, this process often involves teacher language interactions that include modeling and open-ended questioning (Costa, 1985), providing opportunities for practice and feedback (Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009) and making connections between prior knowledge and new knowledge (Newton, 2000).

Language-based interactions play a key role in apprenticing students in learning a new skill (Cazden, 1992; Dickinson et al., 2009; Gibbons, 2012). For example, the following three types of high quality language-based instructional practices are positively associated with students' academic and language skills and align with a socio-cultural lens for learning. These include *language modeling*, in which teachers demonstrate, scaffold, and provide rich vocabulary; *quality of feedback* in which teachers' expand on and extend student responses and contributions; and *concept development* in which teachers bolster higher order thinking skills through questioning, discussions, eliciting student responses, and connecting to students' background knowledge (Mashburn et al., 2008). These practices, found to be effective for monolingual students, are commonly captured with a measure that has demonstrated comparable psychometric properties in contexts with DLLs (Downer et al., 2012): The Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008).

The CLASS aligns with theory on high quality teacher practices in classroom settings and contains three primary domains for prekindergarten and elementary school classrooms: Emotional Support, Classroom Organization, and Instructional Support (Pianta et al., 2008). The two former domains capture teachers' ability to promote a positive, warm, academically and emotionally sensitive classroom climate as well as teachers' capacity to establish routines, manage behaviors, and integrate varied modalities in their instruction, respectively. We focus in the present study on the third domain, Instructional Support, which captures the three types of quality language interactions (i.e., Language Modeling, Quality of Feedback, and Concept Development) rooted in a socio-cultural understanding of learning. Emotional Support, Classroom Organization and Instructional Support all influence student learning and are related (Pianta et al., 2008), however, in the present study we were interested in the quality of language interactions, that is, the way and nature of the language teachers use with students that may promote language development. While teachers' emotional support and classroom organization may influence the content of what teachers say, we were most interested in the Instructional Support domain that specifically captures the quality of teachers' language interactions, a more proximal factor related to students' own language development.

These quality language interactions are particularly important for DLLs (Castro, Espinosa, & Paez, 2014; Sprachman, Caspe, & Atkins-Burnett, 2009; Walqui, 2006) who, on average, have less exposure to the academic discourse of the classroom (Bailey, 2007, 2010), also referred to as Cognitive Academic Language Proficiency (CALP; Cummins, 1980), and are likely to need additional supports to facilitate their accessing new language and content demands simultaneously (August & Hakuta, 1997). For DLLs, teacher scaffolding, or language modeling, is particularly critical and must incorporate comprehensible input (Krashen, 1981). For example, when modeling or asking open-ended questions teacher language must be appropriately adjusted so that DLLs can understand the language input while the language provided must also extend just beyond the learners' current proficiency to maintain the level of content rigor and learning (Echevarria, Vogt, & Short, 2008). Providing opportunity for practice and quality feedback is also essential for DLLs language growth, as they need to be exposed to sophisticated language and, equally important, they must have the opportunity to produce such language (Swain, 1995). Lastly, a wealth of research documents the importance of connecting new content learning to DLLs' own realities and experiences, and accessing students' knowledge of their own communities (August & Hakuta, 1997; Moll, 1994) in order to help students integrate new knowledge into their existing knowledge schemas (Ladson-Billings & Tate, 2006). Together these bodies of research with DLLs indicate that it is not only exposure to language and content that matters, but also the quality and nature of such exposure that leads to academic growth and language acquisition (Wong-Fillmore & Valadez, 1986).

To this end, our choice to focus on the Instructional Support domain and its corresponding dimensions (i.e., Language Modeling, Quality of Feedback, and Concept Development) reflects three core drivers of this investigation. First, as described above, the dimensions of Instructional Support align with areas recognized as particularly important for DLLs. Second, language interactions are central to all components of the Instructional Support domain while other classroom quality indicators such as Emotional Support and Classroom Organization include a focus on resources and physical behaviors. The focus on language interactions in the Instructional Support domain provides a rich opportunity to capture and describe the types of interactions DLLs are potentially experiencing in classrooms. Lastly, we focus on these dimensions because language interactions are the more difficult to script as part of existing curriculum and thus rely on teachers' ability to improvise in their interactions in the classroom, in the absence of built-in curricular directions (Neugebauer, Coyne, McCoach, & Ware, 2017).

This research is particularly important in the context of existing work that demonstrates that while the vast majority of primary teachers have DLLs in their classrooms, few teachers receive training on working with these learners and even fewer recognize the integral role of language interactions for supporting content (Heineke & Neugebauer, 2018; Fillmore, Snow, & Educational Resources Information Center, 2000). As a result, a common practice for teachers eager to support DLLs in accessing the curriculum is teachers reducing the language demands of their speech (Lucas, Villegas, & Freedson-Gonzalez, 2008). Such instruction can have two significant costs: This practice can reduce students' exposure to language that is more sophisticated and may inadvertently lead to the watering down of content (Walqui, 2006; White, 1987). Diaz and Flores (2001) refer to these negative consequences as examples of "negative zones" (p. 33). Drawing on Vygotsky's (1978) Zone of Proximal Development, which describes the distance between a learner's level of mastery without adult guidance and the potential level of mastery with the support of a knowledgeable adult, Diaz and Flores (2001) argue that teachers of DLLs create "positive" or "negative" zones based on their ability to integrate students' knowledge and experiences and connect them to new content (p. 34). By oversimplifying language and, equally important, failing to draw on DLLs' existing knowledge set, teachers may actually impede students' ability to learn more cognitively challenging content. Given the potentially positive and negative impact of teacher practices on DLLs' academic and linguistic growth, suggested by this framework, we were particularly interested in whether teachers' language interactions with students around content – modeling, feedback, and questions to develop higher-order thinking – might vary as a function of the student population.

Divergent Practices for Different Learners

In addition to exploring teachers' potentially divergent use of instructional practices when working with DLLs, we were also interested in examining whether certain instructional practices might be particularly valuable for DLLs compared with their monolingual peers. We review literature below that provides evidence to suggest that DLLs' exposure to language interactions, as well as the benefits they derive from such interactions, may be distinct. This research supports the likelihood that some language practices may be differentially supportive of vocabulary development for DLLs than their monolingual peers.

Existing scholarship on teachers' use of quality language interactions shows variability across schools with regard to the amount and type of language use as well as the implications for language development. For example, in a study conducted with 238 pre-K classrooms across six states, the authors found that schools with the lowest percentages of minority students, i.e., higher numbers of white students, had higher levels of quality interactions in the classroom (Valentino, 2018). Furthermore, a study by Chien et al. (2010) of approximately 700 classrooms found that classrooms with higher numbers of minority students had greater use of didactic instruction, as opposed to a more interactive approach to instruction, which was more common in schools with predominantly white students. Research documenting differences in pedagogy across student populations using the CLASS measure of high-quality teacher practices, found that in 4,000 pre-K classrooms in one state, teachers in schools composed of large proportions of students of color and from low-income households, had the lowest scores on the CLASS (Bassok & Galdo, 2016). These studies provide supporting evidence that teachers' use of high-quality interactions vary as a function of whether they teach higher numbers of minority students from low-income backgrounds, many of whom were likely DLLs. Notably, DLLs have not been the specific focus of these investigations (Valencia, Wixson, Ackerman, & Sanders, 2017). We hypothesized that this same pattern might hold in classrooms with higher numbers of DLLs in particular.

The varied use of these three quality interactions across different populations of learners may also reflect different perspectives on what are the most beneficial practices for different student groups. Some scholars argue that more specialized practices for DLLs are necessary for closing the vocabulary gap (August & Shanahan, 2006; Garcia & Jensen, 2009; Genesee, Lindholm-Leary, Saunders, & Christian, 2006; Goldenberg, 2008; Slavin & Cheung, 2005). DLLs may require additional supports that are not equally suited for monolingual students (Goldenberg, 2008). To elaborate, DLLs' unique linguistic and cultural differences from their Anglo monolingual peers may interact in distinctive ways with different types of quality language interactions. Indeed, some studies argue that certain linguistic and ethnic groups are better served by more didactic forms of instruction (Delpit, 1988; Stipek, 2004), instruction that incorporates different student funds of knowledge beyond school (i.e., experiences from their daily life; Moll, Amanti, Neff, & González, 1992) or instruction that incorporates their native language (Barnett, Yarosz, Thomas, Jung, & Blanco, 2007; Durán, Roseth, & Hoffman, 2010). As such, an open question is whether high-leverage language interactions are differentially associated with growth in vocabulary knowledge for DLLs and their monolingual peers.

In addition, teachers may under or over utilize some language interactions (modeling, quality feedback, and concept development) when working with DLLs. That is, teachers may privilege certain types of language interactions over others based on what they see as most valuable for DLLs' learning, or based on the instructional materials available to them. Indeed, research indicates that teachers who are tasked with working with DLLs tend to focus on modeling and increasing students' vocabulary, but by and large do not make the connection between language demands and understanding content (Bailey, 2010; Heineke & Neugebauer, 2018). These understandings of students' needs may lead teachers to provide sufficient language modeling and quality feedback, as they may see these as necessary linguistic supports. Yet, teachers may fall short of providing DLLs with rich opportunities for content development, with consequences for students' development of more sophisticated vocabulary associated with the disciplines (Nagy & Townsend, 2012).

Curricular programs can provide guidance on important language interactions. For example, language and literacy programs commonly script practices for modeling new vocabulary with a word list and friendly definitions or example sentences (Beck & McKowen, 2004). However, such programs do not address all types of quality language interactions. They cannot provide teachers with a dynamic script including how to address student comments, questions and misunderstandings (i.e., teachers' quality of feedback), and ways to make connections to their respective students' background knowledge (i.e., instructional moves to support content development). Indeed, a core facet of content development rests on the familiarity of the teacher with her students' experiences and knowledge set, information not included in scripted programs. Thus, we hypothesized that Concept Development, in particular, might be an area in which teachers provide less support and from which ELLs might derive particular benefit for their language development.

Pedagogy that Matters for DLLs

Kindergarten, the start of formal schooling, represents a crucial moment for bolstering students' language skills, closing linguistic gaps between learners, and preventing later reading difficulties. By age three, there are already staunch differences between students with high and low levels of English vocabulary knowledge (Fernald, Marchman, & Weisleder, 2013). These differences grow larger over time and can contribute to significant differences in students' English reading comprehension, with students with higher vocabularies having better comprehension later in their schooling (Cunningham & Stanovich, 1997; Hemphill & Tivnan, 2008; Ouellette, 2006).

The robust relationship between vocabulary and later reading outcomes raises concerns for students who begin schooling with lower initial vocabularies and who are therefore at risk for language and literacy difficulties. DLLs often fall into this group, performing on average considerably below their monolingual peers at kindergarten entry in English vocabulary (Mancilla-Martinez & Lesaux, 2011). These students are less likely to improve their language knowledge through classroom-based instruction than their peers with higher vocabularies (Penno, Wilkinson, & Moore, 2002). As such, essential is research that can identify best practices for increasing the vocabulary knowledge of these learners to prevent further linguistic inequities between these groups of students. Identifying kindergarten practices that improve DLLs' language skills has the potential to help close the existing English vocabulary gap between students who start formal schooling with more or less language exposure, but also can bolster a variety of language and literacy related outcomes essential for long-term academic success (Coyne et al., 2009).

The Present Study

The present study set out to understand the language interactions experienced by DLLs in kindergarten and the association between these interactions and students' developing general vocabulary knowledge with two central research questions:

(1) Do teachers' observed Instructional Support practices (i.e., Language Modeling, Quality of Feedback, or Concept Development) differ in classrooms with higher numbers of DLLs versus lower numbers of DLLs?
(2) Are Instructional Support dimensions differentially supportive of vocabulary development for DLLs than monolingual students?

Method

All teachers in the present study used the Elements of Reading Vocabulary Program (EOR-V; Beck & McKeown, 2004) as their core literacy curriculum and received a day of training and professional development at the beginning of the year on how to implement the program practices. This program includes a sequence of stories and five weekly vocabulary words, taught across 24 weeks for 20 minutes each day of the week. Each week incorporates a read aloud of a story that contains the target words followed by word activities, picture cards, and graphic organizers provided in the teacher's manual. EOR-V is not a scripted program, but it does include friendly definitions as well as example sentences of target words. This program is similar to other commonly used research-based language and literacy programs in that it provides an evidence-based sequence of activities, materials, and teacher prompts/models that facilitate productive conversations about words and has been found to be effective in other studies (Apthorp et al., 2012). This study draws from a larger study on cohorts of students that participated in two different programs of research investigating the efficacy of an experimental small-group supplemental vocabulary intervention. As part of these studies, kindergarten teachers implemented the EOR-V curriculum as the standardized universal classroom vocabulary instruction for all students (Cuticelli, 2016). The present sample includes students with a range of vocabulary scores based on the Peabody Picture Vocabulary Test (PPVT) with fall standard scores ranging from *extremely low* to *extremely high*, as described within the PPVT scoring manual (Dunn & Dunn, 2007).

Participants

This study included 263 kindergarten students and 27 kindergarten teachers across 12 elementary schools in the Northeastern United States. Participating students were 50% female and were 5 years, 3.4 months old on average (Mean = 63.40 months, Range = 57–80 months, SD = 3.53). The majority of students were Latino (34%), African American (26%), and White (34%). The remaining 6% of students were those identified as Asian (1%), Native American (1%), Multiracial (2%), or "Other" race (2%). Nearly a quarter (23%) of students in the sample were identified as DLLs because they spoke a language other than English at home.

Teachers were largely female (96%), and the ethnic composition of teachers was White (83%) and African American (17%). Almost all participating teachers had a master's degree (96%) and taught in full-day kindergarten classrooms with an average class size of 20 students (Range = 18–23 students, SD = 1.43). On average, teachers in this sample had 14.02 years of teaching experience (range <1 to 35, SD = 9.56). The number of DLL students in a given classroom ranged widely from 0–20 (Mean = 4, SD = 5.18). No particular racial group of teachers was assigned to classes with higher numbers of DLLs. That is, classes with higher numbers of DLLs were distributed across both White and African American teachers. In this study we define high levels as classrooms with 20% or more DLL students, a cutoff that has been used in prior studies examining linguistically diverse classrooms (Cervetti, Kulikowich, & Bravo, 2015; Sandilos et al., 2015). Across the schools, the percentage of students receiving free or reduced-price lunch (FRPL) ranged from 16% to 87%. Classrooms classified as having 20% or more DLL students did not statistically significantly differ from classrooms with less than 20% DLL students on the school-wide percentage of students receiving free and reduce priced lunch.

Procedures and Measures

At the start of the academic year, kindergarten students' receptive standardized vocabulary was assessed using the PPVT-IV and was assessed a second time in the spring semester to determine students' growth in general vocabulary knowledge over the course of their kindergarten year. To determine the nature of teachers' language interactions teachers were video recorded during their scheduled whole-class vocabulary time during which they were implementing EOR-V. These observations were conducted mid-way through the academic year when teachers were familiar and comfortable with the standardized vocabulary program. Teacher pacing of the program varied and so teachers were observed using different – although similarly leveled – books from the program anthology and with some differences in whether they were on day 1 through 4 of the program for that week. All program days used similar activities and materials. These video-observations were then coded using the CLASS to determine teachers' scores on the Instructional Support domain. It is worth noting that research staff observed teachers on several occasions during the course of the year, and so they were accustomed to having research staff present during their EOR-V instructional time. Other observations were conducted to capture teachers' fidelity to the vocabulary curriculum. While initial analyses included this fidelity measure composed of a checklist of materials and practices included in vocabulary program, this indicator of fidelity did not significantly alter the study findings; for this reason we have not included it here.

CLASS

The CLASS K-3 (Pianta et al., 2008) is a systematic observation tool that assesses the quality of teacher-child interactions in early elementary classrooms (i.e., kindergarten through third grade). Trained observers score teachers on 10 dimensions using a 7-point scale ranging from *Low* (1–2), *Middle* (3–5), to *High* (6–7). One observation cycle consists of 20 minutes of observation followed by 10 minutes of scoring (Pianta et al., 2008). In the present study, a trained observer coded one 20-minute videotaped vocabulary lesson for each teacher. While additional observation cycles would have been more consistent with recommendations for use of the CLASS instrument, we were only able to capture one cycle due to the prescribed curriculum of 20-minute daily lessons.

The 10 dimensions form three primary domains: Emotional Support, Classroom Organization, and Instructional Support. The present study examined the Instructional Support domain, which is comprised of Concept Development, Quality of Feedback, and Language Modeling dimensions. Concept Development represents the teachers' ability to cultivate high-order thinking by connecting concepts, activating prior knowledge, and using real-world examples. Quality of Feedback is exemplified by the teachers' use of scaffolding techniques and follow-up questions to strengthen skill development. Language Modeling represents the teachers' efforts to increase language development, such as using and defining new vocabulary words.

Within the current sample, internal consistency for the Instructional Support domain was high ($\alpha = .93$), and the three dimensions exhibited high inter-correlations ($r = .67$-$.87$). All classroom observations were conducted by a certified CLASS trainer. The CLASS trainer certification requires a rigorous and extensive training process that includes 32 hours of coding-based tasks and discussion of observations, and a 5-hour exam, as well as routine recertification tests (with a minimum of 80% reliability) to offset rater drift. Consistent with the CLASS certification requirements, the CLASS trainer in the present study maintained 80% reliability across all routine checks for rater drift.

Language Survey

Teachers received a demographic questionnaire to determine students' home language, and information on students' proficiency based on the state's language proficiency test, when applicable or available. If teachers responded "yes" to the question "Does this child speak a language other than English at home?" the student was considered a DLL. If the teacher indicated that no other language was spoken at home, the student was considered a monolingual English speaker. Teacher responses were corroborated with data provided by the school on whether students were DLLs or not. There were no discrepancies between teacher and school reports. This measure and its corresponding items are aligned with other language surveys collected in the majority of states across the country (Wolf et al., 2008).

PPVT-IV

The PPVT (PPVT-IV, Dunn & Dunn, 2007) is an individually administered standardized norm-referenced test of receptive language and vocabulary that was collected by trained graduate assistants on the research team. For each test item, a student is shown four pictures and is asked to point to the picture that best represents the meaning of the word said aloud by the examiner. Test-retest reliability for the PPVT-IV is .77 and alternate form reliability is .82. The PPVT-IV demonstrates adequate reliability and validity (Salvia & Ysseldyke, 1998).

Data Analysis

Descriptive statistics and *t*-tests of mean differences were analyzed using SPSS version 25. The distribution of the dimension scores approximated normality with acceptable ranges for skew (−.78 to −1.24) and kurtosis (−.06 to 1.22). To assess mean differences in Instructional Support dimensions across classrooms, a dichotomous classroom composition variable was created indicating if classrooms had 20% or more DLL students or less than 20% DLL students (1 = ≥ 20%, 0 = < 20%), which is a cutoff that has been used in prior studies examining linguistically diverse classrooms (Cervetti et al., 2015; Sandilos et al., 2015). The Hedges' *g* effect size, which accounts for uneven and/or small sample sizes, was calculated to determine substantive significance of the findings (small effect = 0.2, medium effect = 0.5, large effect = 0.8; Ellis, 2010).

Multilevel modeling with cross-level interactions was conducted in M*plus* version 8 (Muthén & Muthén, 1998–2012). A total of 48 children were missing spring PPVT outcome data. However, using chi-square difference tests, no statistically significant demographic differences (i.e., race variables) and no differences in the expressive language risk covariate (described in more detail below) were identified between students with missing data and those with complete data. One participating teacher was missing data on CLASS scores, and five teachers were missing data on one or more demographic variables. No statistically significant differences were identified on CLASS dimensions between teachers with and without demographic data. Thus, all multilevel models were estimated using Full Information Maximum Likelihood (FIML) to account for the missing student and teacher data (Enders & Bandalos, 2001).

The intra-class correlation (ICC) for the PPVT spring outcome measure was 0.35 indicating that 35% of the variance in students' vocabulary scores was attributable to the nesting of students within classrooms (Finch & Bolin, 2017). Additionally, the design effect (DEFF), which is a function of the size of the intraclass correlation and average cluster size, was greater than 2.00 (DEFF = 6.97) further supporting the need to account for students being nested within classrooms (Satorra & Muthen, 1995). To examine growth in PPVT scores over the course of the year, students' baseline PPVT standard score was group mean centred (Enders & Tofighi, 2007) and included as a level 1 covariate in all models. The use of standard scores also accounted for children's age.

Language and Demographic Covariates

Several language and demographic covariates were included in the multilevel models. In addition to PPVT baseline scores, students' baseline expressive vocabulary, as measured by the Expressive Vocabulary Test-2 (EVT-2; Williams, 2007) was included in the models at level 1. Participating schools were implementing a multi-tiered system of support model for vocabulary, as such vocabulary data were collected to inform tier 2, small-group vocabulary interventions. Thus a dichotomous expressive vocabulary variable was included representing students who were and were not at risk of vocabulary difficulties and eligible for supplementary vocabulary support. Specifically, students with EVT standard scores (SS) falling within the *below expected* or *well below expected* ranges were deemed as "at risk" (SS ≤ 84 = 1, SS > 84 = 0). Students' race (i.e., White [reference group], Latino, African American, Other race [Asian, Native American, Multiracial, "Other"]) was also controlled for at level 1. A dichotomous variable indicating DLL status (DLL = 1, monolingual = 0) was added to each model at level 1 as a predictor of interest.

At level 2, one of the CLASS Instructional Support dimensions (Quality of Feedback, Concept Development, or Language Modeling) was grand mean centred and entered into the model. Additional level 2 covariates consisted of teachers' years of experience and teacher race (1 = African American, 0 = White). Given that our sample was underpowered at the school level, a third level could not be added to the multilevel models (Aguinis, Gottfredson, & Culpepper, 2013). Thus, two school-level covariates were also included at level 2. The first covariate was an indicator of whether or not the school was considered "high poverty" (1 = yes, 0 = no). In accordance with data from the National Centre for Education Statistics (NCES, 2018), schools were coded as high poverty if more than 70% of students in the school received free or reduced price lunch. The second covariate represented the proportion of students in a given school who were non-fluent English speakers.

To determine the moderating influence of each Instructional Support dimension on the relation between DLL status and vocabulary growth, cross-level interactions were estimated between one Instructional Support dimension at level 2 and DLL status at level 1. Simple slopes for statistically significant cross-level interactions were probed using procedures recommended by Preacher, Curran, and Bauer (2006).

Results

Students' PPVT standard scores varied widely, with scores falling in the *extremely low* (<70) to *extremely high* (>130) ranges at baseline and spring data collection. Students' DLL status was significantly and negatively associated with fall and spring PPVT scores ($r = -.19$ to $-.20$, $p < .01$). Teachers' scores on the Instructional Support dimensions ranged from *low* (<3) to *high* (>5), with scores falling in the *middle* range (3–5) on average. Table 1 displays the descriptive statistics for students and teachers. Appendix shows teacher mean scores on all three CLASS dimensions, i.e., Emotional Support, Classroom Organization, and Instructional Support to descriptively capture the pedagogical contexts in this study.

T-tests were conducted to determine if there were significant mean differences in Instructional Support dimensions across classrooms with 20% or more DLLs students or less than 20% DLLs (Table 2). Findings revealed statistically significant differences in teachers' Quality of Feedback ($t = -2.72$, $p < .05$, mean difference = – 0.95) and Concept Development ($t = -2.13$, $p < .05$, mean difference = – 0.86) scores, with higher scores occurring in classrooms with less than 20% DLLs. The large effect size for both Quality of Feedback (Hedges' g = 1.07) and Concept Development (Hedges' g = 0.82) indicated that the magnitude of the difference was also substantively significant. Across classrooms with 20% or more DLL students or less

Table 1. Descriptive statistics for students and teachers.

Measures	*M*	*SD*	Range	Percentage
Students				
PPVT Fall Raw Score	89.90	21.47	14–142	-
PPVT Fall Standard Score	102.56	15.40	42–141	-
PPVT Spring Raw Score	99.05	19.40	39–149	-
PPVT Spring Standard Score	101.71	13.86	60–137	-
DLL	-	-	-	23%
Expressive Vocabulary Risk				16%
Female				50%
White	-	-	-	34%
African American	-	-	-	26%
Latino	-	-	-	34%
Other Race	-	-	-	6%
Teachers/Schools				
Quality of Feedback	3.88	0.99	2–5	-
Concept Development	3.73	1.12	1–5	-
Language Modeling	4.81	1.10	2–6	-
Years of Experience	14.02	9.56	<1–35	-
Female				96%
White	-	-	-	83%
African American	-	-	-	17%
High Poverty School	-	-	-	78%
Proportion not fluent in English	.17	15.87	.05– 0.62	-

Table 2. T-test of differences in instructional support dimensions across classrooms with 20% or more DLL students versus less than 20% DLL students.

	T-Test		Classrooms < 20% DLLs (N = 16)		Classrooms ≥ 20% DLLs (N = 10)	
	t	Mean Difference	Range	Mean (SD)	Range	Mean (SD)
Quality of Feedback	−2.72*	−0.95	1–5	4.25 (0.93)	2–4	3.30 (0.82)
Concept Development	−2.13*	−0.86	2–5	4.06 (1.12)	2–4	3.20 (0.92)
Language Modeling	−1.14	−0.50	2–6	5.00 (1.03)	3–6	4.50 (1.18)

*$p < .05$. **$p < .01$. ***$p < .001$.

than 20% DLLs there were no significant differences in teachers' Language Modeling scores. To elaborate on these findings, teachers in classrooms with fewer DLL students often provided scaffolding for struggling students, frequently engaged in feedback loops and prompts for student reasoning, and frequently expanded on student thinking and encouraged student involvement while teachers with 20% or more DLLs with scores closer to a 3 on this dimension of the CLASS engaged in these practices with less frequency and instead in some cases dismissed or ignored a struggling students misunderstandings or gave perfunctory feedback. Similarly, teachers in classrooms with 20% or more DLLs in their classroom who scored in the 3 range on average on Concept Development only occasionally encouraged analysis and reasoning, related concepts to students actual lives and only sometimes provided opportunity for students to generate ideas, link concepts and activities to things they had already learned, by contrast teachers with fewer DLLs, who scored in the 4 range on average more frequently and with more consistency engaged in these teacher moves. Teachers across classrooms with high and low numbers of DLLs in their classrooms engaged in conversations in the classroom, used open-ended questions, repeated and extended student responses, used advanced language and mapper his/her own actions and the students actions through language occasionally or sometimes. Notably, participating teachers' midrange scores (i.e., scores of 3,4, and 5) are consistent with the existing literature (National Center on Quality Teaching and Learning, 2013; Pakarinen et al., 2010; Sandilos et al., 2015), with scholarship also substantiating slightly lower average scores for teachers instructing DLLs (Downer et al., 2012).

Multilevel Models

Separate multilevel models were estimated for each instructional support dimension. Main effects models and models with cross-level interactions are reported in Table 3. Model 1 examined the moderating influence of DLL status on the relation between teachers' observed Quality of Feedback and students' spring PPVT scores. Although the main effect of Quality of Feedback trended toward significance ($p < .10$), the interaction with DLL status did not reach statistical significance.

Model 2 explored the moderating influence of DLL status on the relation between Concept Development and PPVT scores (Table 3). The cross-level interaction between DLL status and Concept Development was statistically significant ($b = 3.19$, $p < .05$). An analysis of simple slopes indicated that DLLs exhibited statistically significantly more growth on PPVT scores ($b = 3.57$, $p < .05$) when they were in classrooms with high Concept Development (+1 standard deviation above the mean), as compared to DLLs in classrooms with low Concept Development (−1 standard deviation below the mean). The slope was not significant for monolingual students indicating that their vocabulary growth did not statistically significantly differ in classrooms with high versus low Concept Development (Figure 1).

Model 3 assessed the cross-level interaction between the Language Modeling dimension and DLL status to predict students' vocabulary growth (Table 3). There were no significant main effects of DLL status or Language Modeling, and there was no statistically significant cross-level interaction between Language Modeling and DLL status.

Table 3. Cross-level moderation analyses examining the interaction between DLL status and CLASS instructional support dimensions.

	Spring PPVT Standard Scores (SS)					
	Model 1 Main Effects *b* (*SE*)	Model 1 Interaction *b* (*SE*)	Model 2 Main effects *b* (*SE*)	Model 2 Interaction *b* (*SE*)	Model 3 Main effects *b* (*SE*)	Model 3 Interaction *b* (*SE*)
Level 1 (N = 263)						
Fall PPVT SS	0.70 (0.04)***	0.71 (0.05)***	0.70 (0.04)***	0.71 (0.04)***	0.70 (0.04)***	0.71 (0.04)***
Expressive vocabulary risk	−4.69 (1.91)*	−4.67 (1.89)*	−4.64 (1.87)*	−4.80 (1.85)**	−4.57 (1.84)***	−4.58 (1.86)*
African American	−1.58 (2.22)	−1.69 (2.26)	−1.73 (2.21)	−1.83 (2.27)	−1.66 (2.24)	−1.79 (2.30)
Latino	−3.87 (2.06) $^{+}$	−3.52 (2.14) $^{+}$	−4.07 (2.02)*	−3.72 (2.11)$^{+}$	−3.98 (2.06) $^{+}$	−3.91 (2.10) $^{+}$
Other Race	−0.01 (2.01)	−0.03 (1.96)	0.12 (1.96)	−0.09 (1.97)	0.03 (2.06)	0.03 (2.08)
Dual Language Learner	1.80 (1.66)	1.23 (1.58)	1.82 (1.71)	1.02 (1.61)	1.68 (1.75)	1.09 (1.66)
Level 2 (N = 27)						
Years of Experience	0.14 (0.10)	0.14 (0.11)	0.14 (0.11)	0.15 (0.11)	0.13 (0.09)	0.13 (0.10)
African American	−4.26 (1.65)**	−3.63 (1.60)*	−4.52 (1.84)*	−3.85 (1.67)*	−4.02 (1.94)*	−3.61 (1.77)*
High Poverty School	−12.78 (2.24)***	−12.68 (1.97)***	−13.08 (2.19)***	−12.95 (2.10)***	−12.95 (2.35)***	−12.96 (2.35)***
Proportion not fluent in English	−0.06 (0.04)	−0.07 (0.05)	−0.05 (0.05)	−0.06 (0.05)	−0.05 (0.04)	−0.05 (0.04)
Quality of Feedback	1.73 (0.91) $^{+}$	1.22 (1.03)	-			-
Concept Development	-	-	0.86 (1.07)	0.39 (1.16)		-
Language Modeling	-	-	-	-	1.10 (1.31)	0.78 (1.41)
DLL x Quality of Feedback	-	2.78 (1.64) $^{+}$	-	-		-
DLL x Concept Development	-	-	-	3.19 (1.55)*		-
DLL x Language Modeling	-	-	-	-		1.74 (1.66)

Model 1 includes the Quality of Feedback dimension; Model 2 includes the Concept Development dimension; Model 3 includes the Language Modeling dimension. Unstandardized estimates (*b*) are reported with the standard error (*SE*) in parenthesis. $^{+}p < .10$. $^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

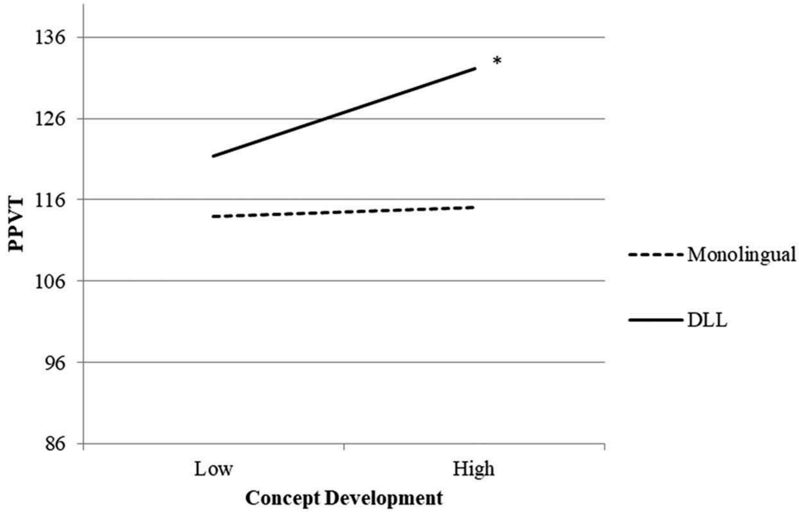

Figure 1. Cross-level interaction between concept development and DLL status to predict spring PPVT scores.
$^{*}p < .05$.

Discussion

The present study contributes to the growing literature on quality language interactions to support DLL language development in several important ways. First, this study found significant differences in the observed quality of language interactions in classrooms with higher numbers of DLLs (i.e., ≥20%) versus classrooms with fewer numbers of DLLs (i.e., <.20%). In particular, teachers instructing in classrooms with higher numbers of DLLs were less likely to engage in language interactions to support higher-order thinking skills associated with Concept Development and were less likely to provide quality feedback to expand and extend student responses. The less frequent use of such practices in classrooms with more DLLs is particularly notable given our second finding in the present investigation: teachers' use of practices to support students' Concept Development was associated with greater gains in vocabulary knowledge at the end of kindergarten for DLLs. As such, this study provides preliminary evidence that the practices that DLLs may differentially benefit from may be less frequently provided and when provided may be in contexts where there are fewer DLLs. Therefore, this study highlights a critical area for addressing a potential gap in practices that may be particularly useful for promoting DLLs' language development and closing achievement differences between groups. We discuss these findings in more detail below.

Our results indicate that in classrooms with higher percentages of DLLs, teachers were less likely to engage in practices associated with higher-order thinking skills (i.e., Concept Development) and responsive feedback (i.e., Quality of Feedback), than teachers in classrooms with fewer numbers of DLLs. This finding is consistent with the literature on the use of interactive instructional practices in schools with high minority populations, which also uncovered higher instances of didactic practices and fewer quality language interactions around content (Chien et al., 2010; Valentino, 2018). This study further bolsters research indicating that teachers may vary their instructional practices as a function of the student population. In the case of the present study, where Concept Development was associated with vocabulary growth, such variations may not necessarily serve all learners. Notably, it is important to acknowledge research indicating that students with low language scores may not benefit from the linguistically complex syntactic utterances that are often present in talk that promotes high-order thinking (e.g., Bowers & Vasilyeva, 2011). Therefore, in some classrooms, variation may have been appropriate. However, existing research also indicates that it is low language scores, not simply language status (i.e., being labeled a DLL) that supports complex syntax as punitive for vocabulary growth in certain learners (Gamez et al., 2017) and thus we would encourage future work that can disentangle the effect of teachers' practices that support higher-order thinking for DLLs with varied high and low levels of language and literacy skills in English.

In the present study we did not find differences in teachers' use of all types of quality language interactions, as has been found in previous studies with high minority school populations (Bassok & Galdo, 2016). Teachers' use of Language Modeling practices was not statistically different across classrooms with high and low percentages of DLL students, indicating that teachers are engaging in some quality language interactions equally, irrespective of the student population. However, quality Language Modeling differs from the two types of interactions that did demonstrate differences between classrooms with higher and lower numbers of DLLs. Language Modeling interactions do not require substantial deviation from a vocabulary program curricular script whereas Concept Development and Quality Feedback require a certain level of improvisation in response to student background knowledge and comments, respectively. To elaborate, Language Modeling, where teachers model and provide rich vocabulary words, is commonly a crucial aspect of teacher manuals (Beck & McKeown, 2004). Indeed, teachers in this study, for example, used definitions provided by the school-wide vocabulary program, in some cases 14 times during a given 20-minute period.

By contrast, Concept Development, when teachers promote higher-order thinking skills via questions, discussions and connecting to students' background knowledge, requires tailored instruction to the learners in the classroom. Yet, a wealth of scholarship indicates that general education teachers may struggle with recognizing the abilities and needs of linguistically diverse students. Specifically, research indicates that teachers are prone to misidentifying DLLs' language abilities (Limbos & Geva, 2001) and struggle with connecting to the varied cultural backgrounds of the DLLs in their class (Rueda, Monzo, & Higareda, 2004). Teachers' inability to interpret and anticipate student abilities may lead to a failure to efficaciously connect to students' background knowledge in the course of a lesson, or their reluctance to attempt to do so. Similarly, quality feedback, in which teachers expand on and extend student responses and contributions, requires that teachers engage in more impromptu interactions with their students – again a practice that cannot adequately be scripted in a program curriculum and must be flexible to varying student ability levels. Teachers with higher numbers of DLLs may feel less equipped to extend and expand on the responses of students who are at a distinctively different level of language knowledge than their monolingual peers, decreasing the frequency with which teachers may engage in productive extensions of student talk. A recent study found that teacher self-reported beliefs about the importance of engaging in high-quality language interactions with DLLs did not necessarily translate to their observed practice (Sawyer et al., 2016), which may indicate a lack of knowledge or practical skills regarding how to engage in these improvised language interactions that can be used in classrooms with higher numbers of DLLs.

Our finding that Language Modeling, and not Concept Development and Quality Feedback, was equally represented across classrooms with varying percentages of DLLs may also reflect teachers' understandings of the needs of DLLs. Indeed, research by Heineke and Neugebauer (2018) found that teachers' understanding of the language demands of DLLs were largely focused on the need to model and use academic vocabulary with DLLs; however, teachers infrequently recognized the intertwined nature of content development and language demands.

An important finding from the present investigation is that teacher-student interactions that support Concept Development seem to be particularly important for DLLs. That is, DLLs had greater vocabulary gains in classrooms with higher use of interactions focused on Concept Development, compared with their monolingual peers for whom statistically significant gains were not observed. Concept Development captures teachers' ability to support higher-order thinking skills through questioning and discussions, as well as by connecting to students' background knowledge. We propose below several potential explanations that may shed light on why these teacher practices may be so potent for DLLs.

First, existing research supports the association between teaching environments that enhance critical thinking with debates as well as discussions, and developing English language proficiency, in studies with older learners (Michaels, O'Connor, & Resnick, 2008). Many DLLs entering kindergarten may be proficient in Basic Interpersonal Communication (BICS); however, school contexts must provide the opportunity to learn language they may not be exposed to outside of school, specifically CALP (Cummins, 1980). Teacher practices that ask students to provide academic language to support their reasoning in academic discussions present a unique opportunity for DLLs to hear and participate in talk essential for developing their academic language use. By contrast, many of the monolingual students in the class may have access to such academic language outside of the classroom (Bailey, 2007). In the present study, at kindergarten entry participating monolingual students had higher general vocabulary scores than did their DLLs peers (standard score on the PPVT-IV of 104.67 for monolingual students, and a score of 95.40 for DLLs). For students who already have a sophisticated academic language knowledge base these practices may be less necessary, but for DLLs these conversations and discussions present an opportunity to be immersed in a rich language environment they may not have access to elsewhere.

Second, the differential benefit of Concept Development for DLLs on vocabulary gains may reflect the importance of real-world examples and explicit connections to prior knowledge for this group of learners in particular. Indeed, research shows that DLLs often struggle with a lack of content relevant background knowledge and an inability to match what they already know with new material (Francis, Rivera, Lesaux, Kieffer, & Rivera, 2006; Klingner & Soltero-González, 2009). Teachers who do not make explicit connections between new learning and prior knowledge may be inadvertently disadvantaging DLLs who, unlike their monolingual peers may not be able to easily map a new concept with their existing knowledge set. Diaz and Flores (2001) argue that teachers who do not integrate students' knowledge and experiences and connect them to new content (p. 34) are creating "negative zones", where learning is watered-down and development is hindered because teachers do not bring in DLLs' background knowledge and experiences to bolster more challenging academic content (Diaz & Flores, 2001, p. 33). By contrast, "positive zones", where teachers account for and build on students' existing assets, foster academic achievement (August & Hakuta, 1997; Ladson-Billings & Tate, 2006). Our finding that DLL students in classrooms with teachers that stimulated higher-level thinking through making connections to students' real-life experiences had greater gains in vocabulary at the end of kindergarten further supports the important relationship between capitalizing on existing knowledge and academic gains for these learners. Notably, this is not the first study to find this association. Lopez (2012), with 995 late elementary DLLs across 46 classrooms, found a similar association between the Instructional Support domain and students' literacy scores. This study builds and extends Lopez's finding by providing a compelling example of this phenomenon at one of the earliest and most influential moments in students' schooling when, already these practices are, in some cases, influencing the language gap between learners.

Limitations

Several limitations warrant mention. First, this study intended to capture teacher interactions with students during the period of the day dedicated to literacy curriculum. As such, only one CLASS observation cycle was used to assess teacher-child interaction quality (20 minutes). Although prior research has found that CLASS observation scores remain relatively stable across one to four cycles (Curby, Grimm, & Pianta, 2010), increasing the number of cycles or engaging in multiple observations across the course of the program in future studies would provide a more robust picture of the language interactions that occur within the classroom.

Second, the measure used to capture DLL status was collected using school data regarding the languages spoken at home. Without the use of English and Spanish language measures to assess proficiency, it is unclear if different profiles of DLL students in classrooms may have contributed to the findings. The present study included an expressive vocabulary risk indicator as a covariate that was proximal to English proficiency. Future work should use a more nuanced measure to capture language proficiency in students' two languages to explore the interactions between DLL status and teacher-child interactions on vocabulary growth.

Similarly, although the models included a school-level measure of SES, we were unable to control for socio-economic status at the student level given that this variable was not collected as part of the larger study. Prior research indicates that SES influences students' early vocabulary skills (Hammer et al., 2014), and thus, the controlling for this variable would provide a stronger case for the moderating effect of teacher-child interaction quality on the relation between language status and receptive vocabulary growth. Future work should build on this study by including SES at the student level to further isolate the influence of teacher–child interaction quality on DLLs' achievement.

Finally, while we collected data from various sources to capture the experience of students, we did not specifically focus on teachers' understandings of their own instructional moves. Thus, we cannot determine if the less frequent use of certain types of interactions in classrooms with more DLLs was an intentional choice of participating teachers. Subsequent research should include teacher interviews to ascertain teachers' understandings of their pedagogical choices and their interpretations about what kinds of practices best serve DLLs.

Implications

The present study found that teachers in classrooms with higher numbers of DLLs engaged less frequently in student-teacher interactions found in previous research to bolster students' language and literacy development. In addition, the present study found that despite the less frequent use of practices associated with building higher-level thinking skills and providing responsive feedback, DLL students in classrooms with teachers who frequently supported these higher-level thinking skills had greater gains in general vocabulary scores at the end of kindergarten. These results have several implications for practice and research.

First, this finding provides preliminary evidence that higher-order thinking skills or Concept Development may be an important, and in the case of the present study, less common, type of student-teacher interaction in classrooms. Teachers in this study, consistent with the literature more broadly, did engage in practices more commonly associated with language growth, such as Language Modeling. However, this study indicates that teachers should be focusing on developing students' Concept Development as well. Sophisticated ideas reside in complex language, and as was seen in the present investigation, higher-level discussion about content was associated with higher vocabulary scores for this group of learners compared with their monolingual peers who did not demonstrate comparable gains. Pre-service and in-service teacher programs commonly emphasize vocabulary and language modeling as essential for bolstering the performance of these students (Heineke & Neugebauer, 2018; Fillmore, Snow, & Educational Resources Information Center, 2000). Yet, equally important for these teacher preparation programs is an emphasis on how teachers can support DLLs in accessing content and engaging in complex conversations about content. This study did not capture more fine-grained language interactions between teachers and students, so an analysis of the nature of the language used by teachers is not possible (e.g., whether they watered down the language demands of challenging material). We encourage future observational studies to qualitatively capture specific strategies and types of teacher language that teachers with high scores in Concept Development use when working in classrooms with high numbers of DLLs to document robust examples of effective practice. As a preliminary step, subsequent studies should also capture teachers' existing knowledge about these constructs to determine if it is teacher knowledge of these Instructional Support practices that contributes to their use or underuse, or if their pedagogical choices are driven by other classroom or school-level factors.

From a methodological standpoint, this study adds to a limited body of research exploring the CLASS Instructional Support domain with DLLs. We encourage use of the CLASS instrument across diverse student samples with the aim of addressing whether specific pedagogical methods are more or less beneficial for distinctive populations of learners. Such research will provide important insight for best practice for teachers in schools with changing demographics.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, *39*, 1490–1528. doi:10.1177/0149206313478188

Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a supplemental vocabulary program on word knowledge and passage comprehension. *Journal of Research on Educational Effectiveness*, *5*, 160–188. doi: 10.1080/19345747.2012.660240.

August, D., & Hakuta, K. (1997). *Improving schooling for language-minority children: A research agenda*. Washington, DC: National Academies Press.

August, D., & Shanahan, T. (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth*. Mahwah,NJ: Lawrence Erlbaum Associates.

Bailey, A. L. (2007). *The language demands of school: Putting academic English to the test*. New Haven, CT: Yale University Press.

Bailey, A. L. (2010). Implications for instruction and assessment. In M. Shatz & L. Wilkinson (Eds.), *The education of English language learners* (pp. 222–247). New York, NY: Guilford Press.

Ballantyne, K. G., Sanderman, A. R., & Levy, J. (2008). *Educating English language learners: Building teacher capacity*. Washington, DC: National Clearinghouse for English Language Acquisition.

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

Bassok, D., & Galdo, E. (2016). Inequality in preschool quality? Community-level disparities in access to high-quality learning environments. *Early Education and Development*, *27*(1), 128–144. doi:10.1080/10409289.2015.1057463

Beck, I. L., & McKeown, M. G. (2004). Elements of Reading: Vocabulary. Vocabulary supplemental program for kindergarten through grade 5. Austin, TX: Harcourt Achieve (formerly Steck-Vaughn).

Borman, G. D., Slavin, R. E., Cheung, A., Chamberlain, A., Madden, N., & Chambers, B. (2005a). The national randomized field trial of success for all: Second-year outcomes. *American Educational Research Journal*, *42*, 673–696. doi:10.3102/00028312042004673

Borman, G. D., Slavin, R. E., Cheung, A., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2005b). Success for all: First-year results from the national randomized field trial. *Educational Evaluation and Policy Analysis*, *27*, 1–22. doi:10.3102/01623737027001001

Bowers, E. P., & Vasilyeva, M. (2011). The relation between teacher input and lexical growth of preschoolers. *Applied Psycholinguistics*, *32*, 221–241. doi:10.1017/S0142716410000354

Castro, D. C., Espinosa, L., & Paez, M. (2014). Defining and measuring quality in early childhood practices that promote dual language learners' development and learning. In M. Zaslow, I. Martinez-Beck, K. Tout, & T. Halle (Eds.), *Measuring quality in early childhood settings* (pp. 257–280). Baltimore, MD: Paul H. Brookes.

Cazden, C. B. (1992). *Whole language plus: Essays on literacy in the United States and New Zealand*. New York, NY: Teachers College Press.

Cervetti, G. N., Kulikowich, J. M., & Bravo, M. A. (2015). The effects of educative curriculum materials on teachers' Use of instructional strategies for English language learners in science and on student learning. *Contemporary Educational Psychology*, *40*, 86–98. doi:10.1016/j.cedpsych.2014.10.005

Chien, N. C., Howes, C., Burchinal, M., Pianta, R. C., Ritchie, S., Bryant, D. M., … Barbarin, O. A. (2010). Children's classroom engagement and school readiness gains in Prekindergarten. *Child Development*, *81*(5), 1534–1549. doi:10.1111/j.1467-8624.2010.01490.x

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's literacy skill growth. *Journal of Educational Psychology*, *98*, 665–689. doi:10.1037/0022-0663.98.4.665

Costa, A. L. (1985). *Developing minds: A resource book for teaching thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Jr., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth, versus depth. *Elementary School Journal*, *110*(1), 1–48. doi:10.1086/598840

Cummins, J. (1980). The construct of language proficiency in bilingual education. In J. E. Alatis (Ed.), *Georgetown university round table on languages and linguistics 1980* (pp. 81–103). Washington, DC: Georgetown University Press.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. doi:10.1037/0012-1649.33.6.934

Curby, T. W., Grimm, K. J., & Pianta, R. C. (2010). Stability and change in early childhood classroom interactions during the first two hours of a day. *Early Childhood Research Quarterly*, *25*, 373–384. doi:10.1016/j.ecresq.2010.02.004

Cuticelli, M. N. (2016). *Teacher effects on kindergarten students' vocabulary learning* (unpublisheddoctoral dissertation). University of Connecticut, Storrs.

Delpit, L. (1988). *Other people's children: Cultural conflict in the classroom*. New York, NY: W. W. Norton.

Diaz, E., & Flores, B. (2001). Teacher as sociocultural, sociohistorical mediator. In M. de la Luz Reyes & J. J. Halcon (Eds.), *The best for our children: Critical perspectives on literacy for Latino students* (pp. 29–47). New York, NY: Teachers College Press.

Dickinson, D. K., Flushman, T. R., & Freiberg, J. B. (2009). Language, reading, and classroom supports: Where we are and where we need to be going. In B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton, & J. Trefers-Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 23–38). Hampshire, UK: Palgrave-MacMillan.

Dickinson, D. K., Darrow, C., & Tinubu, T. A. (2008). Patterns of teacher–Child conversations in head start classrooms: Implications for an empirically grounded approach to professional development. *Early Education & Development*, *19*, 396–429. doi:10.1080/10409280802065403

Donovan, M. S., & Cross, C. T. (Eds.). (2002). *Minority students in special and gifted education*. Washington, DC: National Academy Press, National Research Council Committee on Minority Representation in Special Education.

Downer, J. T., Lopez, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher-child interactions in classrooms serving Latinos and dual language learners: Applicability of the classroom assessment scoring system in diverse settings. *Early Childhood Research Quarterly*, *27*, 21–32. doi:10.1016/j.ecresq.2011.07.005

Dunn, L. M., & Dunn, D. M. (2007). *The peabody picture vocabulary test* (4th ed.). Bloomington, MN: NCS Pearson, Inc.

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 Childhood Research Quarterly*, *25*(2), 207–217. doi:10.1016/j.ecresq.2009.10.002

Echevarria, J., Vogt, M., & Short, D. J. (2008). *Making content comprehensible for English language learners: The SIOP model*. Boston, MA: Allen Bacon.

Ellis, P. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge, MA: Cambridge University Press.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. *Structural Equation Modeling*, *8*, 430–457. doi:10.1207/S15328007SEM0803_5

Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevelmodels: A new look at an old issue. *Psychological Methods*, *12*, 121–138. doi:10.1037/1082-989X.12.2.121

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234–248. doi:10.1111/desc.12019

Fillmore, L. W., Snow, C. E., & Educational Resources Information Center (U.S.). (2000). What teachers need to know about language. Washington, DC: U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Finch, W. H., & Bolin, J. E. (2017). *Multilevel modeling using Mplus*. Boca Raton, FL: Taylor & Francis Group LLC.

Francis, D., Rivera, M., Lesaux, N., Kieffer, M., & Rivera, H. (2006). *Practical guidelines for the education of English language learners: Research-based recommendations for serving adolescent newcomers* (Under cooperative agreement grant S283B050034 for U.S. Department of Education). Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Gamez, P., Neugebauer, S. R., Coyne, M., McCoach, D. B., & Ware, S. (2017). Linguistic and social cues for vocabulary learning in dual language learners and their english-only peers. *Early Childhood Research Quarterly*, *40*, 25–37.

Garcia, E., & Jensen, B. (2009). Early educational opportunities for children of Hispanic origin. *SRCD Social Policy Report*, *23*, 3–19.

Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2006). *Educating English language learners*. New York, NY: Cambridge University Press.

Gibbons, P. (2010). Classroom talk and the learning of new registers in a second language. *Language and Education*, *12*, 99–118. doi:10.1080/09500789808666742

Gibbons, P. (2012). Mediating language learning: Teacher interactions with ESL students in a content-based classroom. *TESOL Quarterly*, *37*, 247–273. doi:10.2307/3588504

Goldenberg, C. (2008). Teaching English language learners: What the research does—And does not—Say. *American Educator*, *8–23*, 42–44.

Gutiérrez, K. D., Zepeda, M., & Castro, D. C. (2010). Advancing early literacy learning for all children: Implications of the NELP report for dual-language learners. *Educational Researcher*, *39*, 334–339. doi:10.3102/0013189X10369831

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C, & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 715–733. doi: 10.1016/j.ecresq.2014.05.008

Heineke, A., & Neugebauer, S. (2018). The complexity of language and learning: Deconstructing teachers' conceptions of academic language. *Issues in Teacher Education*, *27*(3), 73–89.
Heineke, A., & Neugebauer, S. R. (2018). The complexity of language and learning: Deconstructing teachers' conceptions of academic language. *Issues in Teacher Education*, *27*, 73–89.
Hemphill, L., & Tivnan, T. (2008). The importance of early vocabulary for literacy achievement in high-poverty schools. *Journal of Education for Students Placed at Risk (JESPAR)*, *13*(4), 426–451. doi:10.1080/10824660802427710
Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*, 337–374. doi:10.1016/S0010-0285(02)00500-5
Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51–68. doi:10.1016/j.ecresq.2007.09.004
Klingner, J., & Soltero-González, L. (2009). Culturally and linguistically responsive literacy instruction for English language learners with learning disabilities. *Multiple Voices for Ethnically Diverse Exceptional Learners*, *12*(1), 4–20.
Krashen, S. (1981). *Second language acquisition and second language learning*. New York, NY: Pergamon.
Ladson-Billings, G., & Tate, W. F. (2006). *Education research in the public interest: Social justice, action, and policy*. New York, NY: Teachers College Press.
Limbos, M. M., & Geva, E. (2001). Accuracy of teacher assessments of second-language students at risk for reading disability. *Journal of Learning Disabilities*, *34*, 135–151. doi:10.1177/002221940103400204
Lopez, F. A. (2012). Moderators of language acquisition models and reading achievement for English language learners: The role of emotional warmth and instructional support. *Teachers College Record*, *114*(8), 1–30.
Lucas, T., Villegas, A. M., & Freedson-Gonzalez, M. (2008). Linguistically responsive teacher education: Preparing classroom teachers to teach English language learners. *Journal of Teacher Education*, *59*(4), 361–373. doi:10.1177/0022487108322110
Mancilla-Martinez, J., & Lesaux, N. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development*, *82*(5), 1544–1560. doi:10.1111/j.1467-8624.2011.01633.x
Mashburn, A., Pianta, R., Hamre, B., Downer, J., Barbarin, O., Bryant, D., … Early, D. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79*(3), 732–789. doi:10.1111/j.1467-8624.2008.01154.x
Michaels, S., O'Connor, C., & Resnick, L. B. (2008). Deliberative discourse idealized and realized: Accountable talk in the classroom and in civic life. *Studies in Philosophy and Education*, *27*(4), 283–297. doi:10.1007/s11217-007-9071-1
Moll, L. C., Amanti, C., Neff, D., & González, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and classrooms. *Theory into Practice*, *31*(2), 132–141. doi:10.1080/00405849209543534
Moll, L. C., & González, N. (1994). Lessons from research with language-minority children. *Journal of Reading Behavior*, *26*(4), 439–456. doi: 10.1080/10862969409547862
Mulligan, G. M., McCarroll, J. C., Flanagan, K. D., & Potter, D. (2016). *Findings from the third-grade round of the early childhood longitudinal study, kindergarten class of 2010–11 (ECLS-K:2011)* (NCES 2016-094). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthén & Muthén.
Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. doi:10.1002/RRQ.011
National Center for Education Statistics. (2018). *The condition of education: Concentration of public school students eligible for free or reduced-price lunch*. Retrieved from https://nces.ed.gov/programs/coe/indicator_clb.asp
National Center for Education Statistics (NCES). (2004). *The condition of education 2004* (NCES 2004-077). Washington, DC: Institute of Education Sciences, U.S. Department of Education. Retrieved from https://nces.ed.gov/pubs2004/2004077.pdf
National Center for Education Statistics (NCES). (2017). *Fast facts: Back to school statistics*. Washington, DC: Institute of Education Sciences, U.S. Department of Education.
National Center on Quality Teaching and Learning. (2013). *Improving teacher–Child interactions: Using the CLASS in head start preschool programs* (Grant # 90HC0002). U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start. Retrieved from eclkc.ohs.acf.hhs.gov/hslc/tta-system/teaching/docs/using-the-class.pdf
Neugebauer, S. R., Coyne, M., McCoach, B., & Ware, S. (2017). Teaching beyond the intervention: The contribution of teacher language extensions to vocabulary learning in urban kindergarten classrooms. *Reading and Writing*, *30*, 543–567.
Newton, D. (2000). *Teaching for understanding: What it is and how to do it*. Florence, KY: Routledge Falmer.
Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554–566. doi:10.1037/0022-0663.98.3.554
Ouellette, G. P., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing: an Interdisciplinary Journal*, *23*(2), 189–208. doi:10.1007/s11145-008-9159-1

Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J.-E. (2010). A validation of the classroom assessment scoring system in finnish kindergartens. *Early Education and Development*, *21*(1), 95–124. doi:10.1080/10409280902858764

Park, M., O'Toole, A., & Katsiaficas, C. (2017, October). *Dual language learners: A national demographic and policy profile*. Washington, DC: Migration Policy Institute.

Pearson, P. D., & Gallagher, M. C. (1983). The instruction of reading comprehension. *Contemporary Educational Psychology*, *8*, 317–344. doi:10.1016/0361-476X(83)90019-X

Penno, J. F., Wilkinson, I. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *92*, 22–33.

Pianta, R. C., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., & Barbarin, O. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, *9*(3), 144–159. doi:10.1207/s1532480xads0903_2

Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *The classroom assessment scoring system manual, K-3*. Baltimore, MD: Brookes Publishing Co.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interaction effects in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, *31*, 437–448. doi:10.3102/10769986031004437

Rueda, R., Monzo, L. D., & Higareda, I. (2004). Appropriating the sociocultural resources of Latino paraeducators for effective instruction with Latino students: Promises and problems. *Urban Education*, *39*(1), 52–90. doi:10.1177/0042085903259213

Salvia, J., & Ysseldyke, J. E. (1998). *Assessment* (*7th* ed.). Boston, MA: Houghton Mifflin.

Sandilos, L. E., Cycyk, L. M., Scheffner Hammer, C., Sawyer, B. E., López, L., & Blair, C. (2015). Depression, control, and climate: An examination of factors impacting teaching quality in preschool classrooms. *Early Education and Development*, *26*, 1111–1127. doi:10.1080/10409289.2015.1027624

Satorra, A., & Muthen, B. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, *25*, 267–316. doi:10.2307/271070

Sawyer, B. E., Scheffner Hammer, C., Cycyk, L. M., López, L., Blair, C., Sandilos, L., & Komaroff, E. (2016). Preschool teachers' language and literacy practices with dual language learners. *Bilingual Research Journal*, *39*(1), 35–49. doi:10.1080/15235882.2016.1138904

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

Sprachman, S., Caspe, M., & Atkins-Burnett, S. (2009). *Language interaction snapshot (LISn) field procedures and coding guide*. Princeton, NJ: Mathematica Policy Research.

Stipek, D. (2004). Teaching practices in kindergarten and first grade: Different strokes for different folks. *Early Childhood Research Quarterly*, *19*(4), 548–568. doi:10.1016/j.ecresq.2004.10.010

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp. 125–144). Oxford, UK: Oxford University Press.

Valencia, S. W., Wixson, K. K., Ackerman, T., & Sanders, E. (2017). *Identifying text-task-reader interactions related to item and block difficulty in the national assessment for educational progress reading assessment*. San Mateo, CA: National Center for Education Statistics.

Valentino, R. (2018). Will public Pre-K really close achievement gaps? Gaps in prekindergarten quality between students and across states. *American Educational Research Journal*, *55*(1), 79–116. doi:10.3102/0002831217732000

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walqui, A. (2006). Scaffolding instruction for English language learners: A conceptual framework. *International Journal of Bilingual Education and Bilingualism*, *9*(2), 159–180. doi:10.1080/13670050608668639

White, L. (1987). Against comprehensible input: The input hypothesis and the development of second language competence. *Applied Linguistics*, *8*, 95–110. doi:10.1093/applin/8.2.95

Williams, K. T. (2007). *Expressive vocabulary test-2*. Minneapolis, MN: NCS Pearson.

Wolf, M. K., Kai, J., Griffin, N., Herman, J. L., Bachman, P. L., & Chang, S. M., Tim Farnsworth CRESST/University of California. (2008). *Issues in Assessing English Language Learners: English Language Proficiency Measures and Accomodation uses. Practice Review (Part 2 of 3)*. CRESST report 732. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Wong Fillmore, L., & Valadez, C. (1986). Teaching bilingual learners. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 648-685). New York: Macmillan.

Zaslow, M., Martinez-Beck, I., Tout, K., & Halle, T. (2011). *Quality measurement in early childhood settings*. Baltimore, MD: Brookes.

Appendix

Descriptive Statistics for Additional CLASS Dimensions Not Included in Primary Analyses

Dimensions	*M*	*SD*	Range
Positive Climate	5.23	1.21	2.00–7.00
Negative Climate	1.27	0.53	1.00–3.00
Teacher Sensitivity	4.77	1.24	2.00–6.00
Regard for Student Perspectives	4.77	1.11	2.00–6.00
Behavior Management	5.85	1.08	3.00–7.00
Productivity	5.92	0.98	3.00–7.00
Instructional Learning Formats	4.92	1.09	2.00–6.00