Published in final edited form as:
NHSA Dialog. 2011 October 1; 14(4): 189–212. doi:10.1080/15240754.2011.613129.

Effects of Web-Mediated Teacher Professional Development on the Language and Literacy Skills of Children Enrolled in Pre-Kindergarten Programs

Jason Downer,
Center for Advanced Study of Teaching and Learning, University of Virginia

Robert Pianta,
Center for Advanced Study of Teaching and Learning, University of Virginia

Xitao Fan,
University of Macau, China

Bridget Hamre,
Center for Advanced Study of Teaching and Learning, University of Virginia

Andrew Mashburn, and
Center for Advanced Study of Teaching and Learning, University of Virginia

Laura Justice
Ohio State University

Abstract

As early education grows in the United States, in-service professional development in key instructional and interaction skills is a core component of capacity-building in early childhood education. In this paper, we describe results from an evaluation of the effects of MyTeachingPartner, a web-based system of professional development, on language and literacy development during pre-kindergarten for 1338 children in 161 teachers' classrooms. High levels of support for teachers' implementation of language/literacy activities showed modest but significant effects for improving early language and literacy for children in classrooms in which English was the dominant language spoken by the students and teachers. The combination of web-based supports, including video-based consultation *and* web-based video teaching exemplars, was more effective at improving children's literacy and language skills than was only making available to teachers a set of instructional materials and detailed lesson guides. These results suggest the importance of targeted, practice-focused supports for teachers in designing professional development systems for effective teaching in early childhood programs.

Keywords

Early childhood education; In-service training; Professional development; Language development; Emergent literacy

Rigorous studies now indicate that teachers' effective implementation of instruction through interactions with children is a mechanism through which the value of enrollment, as well as exposure to well-developed curricula and instructional activities, is transmitted to children in

Correspondence concerning this article should be addressed to Jason Downer, 350 Old Ivy Way, Suite 100, Charlottesville, VA 22903-4897. jdowner@virginia.edu.

preschool programs (Howes et al., 2008; Hamre & Pianta, 2005; National Council on Teacher Quality [NCTQ], 2005; NICHD Early Child Care Research Network [ECCRN], 2000). Standardized observations involving several thousand US early education classrooms clearly demonstrate that, on average, the quality of child-teacher interactions in such programs is not high, particularly for instruction, and effective curriculum implementation is inadequate (NICHD ECCRN, 2002; Peisner-Feinberg & Burchinal, 1997; Pianta et al., 2005). Thus, it is increasingly recognized that the promise of early childhood education depends in large part on the professional development and training of teachers in instructional/interaction skills (Zaslow & Martinez-Beck, 2005), particularly as they apply to supporting children's early literacy and language competencies, key components of early school success (Dickinson & Brady, 2005). In the present study, we report results for effects of MyTeachingPartner (Pianta, Mashburn, Downer, Hamre, & Justice, 2008), a web-based approach to professional development that focuses on exposing teachers to video-based exemplars of other teachers' effective interactions with children and consultation focused on analysis and observation of their own behavior. This approach has proven effective with regard to improvements in the quality of pre-k teachers' interactions with children (Pianta et al., 2008), and in this study we report effects on child language and early literacy outcomes in classrooms with diverse learners.

Professional development of early childhood educators is a key component to providing high quality classroom environments that will help children to succeed (Bogard & Takanishi, 2005; Zaslow & Martinez-Beck, 2005). Effective teaching in early childhood education requires skillful combinations of explicit instruction, sensitive and warm interactions, responsive feedback, and verbal engagement/stimulation intentionally directed to ensure children's learning while embedding these interactions in a classroom environment that is not overly structured or regimented (Burchinal, Roberts, Riggins, Zeisel, Neebe, & Bryant, 2000; Hyson & Biggar, 2005). This approach to early childhood teaching is dually endorsed by those who advocate tougher standards and more instruction, as well as those who argue for child-centered approaches, and has strong parallels in the types of instruction and teacher-child interactions that have been shown to contribute to student achievement growth in K-12 value-added studies (see Hart, Stroot, Yinger, & Smith, 2005; NCTQ, 2005). The challenge is how to produce such teaching in large numbers of highly diverse teachers working in diverse early childhood settings and to ensure that resulting teacher training and preparation can be linked to improvements in child outcomes, particularly language and emergent literacy skills that serve as the building blocks for reading.

One representation of the quality of teacher-child interactions in early childhood classrooms and their effects on child outcomes emanates from observations conducted through the National Center for Early Development and Learning's (NCEDL) Multi-state Study of Pre-kindergarten and the Statewide Early Education Programs study (SWEEP; LoCasale-Crouch et al., 2007; Mashburn et al., 2008; Pianta, Howes, Early, Clifford, Bryant, & Burchinal, 2003). Variation in teachers' use of instructionally and emotionally supportive interactions with children, as assessed by the Classroom Assessment Scoring System (CLASS; Pianta, La Paro, & Hamre, 2008) and Early Childhood Environmental Rating Scale – Revised (Harms, Clifford, & Cryer, 1998), was directly related to growth in children's achievement test scores and social behavior ratings across the pre-k year (Howes et al., 2008; Mashburn et al., 2008) and into kindergarten (Burchinal, Roberts, Riggins, Zeisel, Neebe, & Bryant, 2008). Of particular note was the finding that the quality of teachers' *Instructional Support* (e.g., stimulation of conceptual development, provision of feedback) was most consistently and strongly related to growth in literacy, language, and math skills (Howes et al., 2008; Mashburn et al., 2008). These findings indicate that targeting teachers' interactions with students (emotional, instructional, and management-focused) for professional development

and training may be particularly beneficial because these interactions are the proximal mechanism responsible for effects on children's early academic development.

The present study focuses on MyTeachingPartner (MTP), which is an approach to professional development organized around three principles that promote teachers' capacity to skillfully use interactions with children to promote learning. These principles state that professional development requires extensive opportunities for a) *observation* of high quality instruction, implementation, and language and social interaction through analysis and viewing of multiple video examples, b) *skills training* in identifying in/appropriate instructional, linguistic, and social responses to children's cues, and how teacher responses can contribute to child literacy and language skill growth, and c) *ongoing individualized feedback and support* focused on one's own instruction, implementation, and interactions with children. As part of MTP, teachers learn to observe their interactions and receive feedback related to improving quality and effectiveness based on the validated CLASS Framework (Hamre & Pianta, 2007; Pianta, La Paro, & Hamre, 2008). Because the CLASS reliably and validly measures aspects of teachers' instruction and interaction focused on language and early literacy that predicts gains in these areas during the pre-k years, it serves as one of the central "targets" for teachers' knowledge and skills training in MTP. And, it has been demonstrated in several other controlled trials that coaching targeted on teachers' explicit instructional (Landry, Swank, Smith, Assel, & Gunnewig, 2006; Powell, Diamond, Burchinal, & Koehler, 2010) or social interactions (Raver et al., 2008) has significant impacts not only for improving the nature and quality of interactions, but child outcomes as well. These studies help establish the evidentiary base for the argument that teachers and children in early education settings benefit from professional development resources that target teachers' interactions with children (Pianta, 2005).

Consistent with this emphasis on professional development targeted toward practice, MTP consultation and video exemplars explicitly link dimensions of teacher-child interaction to six high priority skill targets for preschool literacy and language instruction (Lonigan, 2004). Based on meta-analyses (e.g., Hammill, 2004; National Early Literacy Panel [NELP], 2008) and longitudinal studies of early language and literacy predicting later reading and language skills (e.g., Bryant, MacLean, & Bradley, 1990; Catts, Fey, Zhang, & Tomblin, 2001; Chaney, 1998; Christensen, 1997; Gallagher, Frith, & Snowling, 2000; Schatschneider, Fletcher, Francis, Carlson, & Foorman, 2004; Storch & Whitehurst, 2002), we selected six targets for MTP supports: phonological awareness, alphabet knowledge, print awareness, vocabulary/linguistic concepts, narrative, and social communication/ pragmatics (e.g., Bunce, 1995; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999a; Lonigan, Bloomfield, Anthony, Bacon, Phillips, & Samwel, 1999b; Notari-Syverson, O'Connor, & Vadasy, 1998). The first three targets (*phonological awareness, alphabet knowledge, print awareness*) are literacy skills that consistently predict school-age decoding (NELP, 2008), are amenable to change via interventions (e.g., Justice & Ezell, 2002; Ukrainetz, Cooney, Dyer, Kysar, & Harris, 2000; van Kleeck, Gillam, & McFadden, 1998; Whitehurst, Epstein, Angell, Crone, & Fischel, 1994) and are under-developed in at-risk pupils (e.g., Bowey, 1995; Lonigan et al., 1999b; Snowling, Gallagher, & Frith, 2003). The other targets--*vocabulary/linguistic concepts, narrative*, and *social communication/ pragmatics*--are moderately associated with school-age decoding (average $r = .33$; NELP, 2008) and reading comprehension (average $r = .33$; NELP, 2008).

In a recent controlled evaluation, MTP consultation was demonstrated to produce gains in high-quality implementation of instruction and the quality of teacher-child interactions (Pianta et al., 2008). Specifically, teachers exposed to regular, ongoing web-mediated consultation focused on a standardized approach to analysis and observation of their interactions with children produced significant gains in the quality of emotional,

organizational, and instructional supports they provided to children over the course of the intervention year. These gains for the group receiving consultation were greater than those for a group that only received access to a website with video exemplars of effective teacher-child interactions. Interestingly, teachers' review of the web-based video exemplars was related to gains in instructional quality when this was their only intervention resource. In sum, teachers in both the consultation and the video-exemplar groups showed some evidence of benefiting from a focus on effective teacher-child interactions, with the consultation group showing greater improvement. However, the effects of these forms of professional development on gains in children's language and literacy skills during the pre-k year were not addressed in this initial study, and are therefore the explicit focus of the current investigation.

The goal of improving children's language and literacy skills through professional development of pre-k teachers should not be surprising. Many children are lacking in spoken language and literacy competencies at the start of school, particularly those growing up in less socially or economically advantaged households, many of whom attend preschool (Snow, Hemphill, & Barnes, 1991; U.S. Dept. of Ed., 2000; Vernon-Feagans, 1996). In fact, much of the rapid expansion of pre-kindergarten programs in the United States is based on the expectation that preschool experiences will help remediate these gaps in development.

Because early language and literacy skills contribute causally to later reading achievement (e.g., Anthony, Lonigan, Driscoll, Phillips, & Burgess, 2003; Lonigan, Burgess, Anthony, & Barker, 1998; Storch & Whitehurst, 2002), early childhood education can prevent/reduce the prevalence of reading difficulties among at-risk elementary students if programs can effectively implement appropriate curricula (e.g., Torgesen, 1998). Experimental research is available on effective early language and literacy curriculum interventions that can be used in classrooms and integrated into teacher preparation programs (e.g., Byrne & Fielding-Barnsley, 1993, 1995; Girolametto, Pearce, & Weitzman, 1996; Girolametto, Weitzman, & Clements-Baartman, 1998; Justice & Ezell, 2002; Penno, Wilkinson, & Moore, 2002; Wasik & Bond, 2001; Whitehurst et al., 1994). However, even for early childhood educators who have these curricula available to them in their classrooms or who have been exposed to them in courses, observational studies show clearly that these demonstrably effective literacy interventions have *no* effect on child outcomes when the quality of implementation (i.e., instructional interactions) is low (Dickinson & Brady, 2005; Howes et al., 2008). In fact, these teachers appear under-trained in how to *implement* instructional activities in early literacy and engage in interactions and conversations that promote language skills (Justice & Ezell, 1999; Morrison & Connor, 2002; NICHD ECCRN, 2002). Furthermore, they are rarely exposed to field-based examples of objectively-defined high quality practice and have few if any opportunities to receive feedback about the extent to which their classroom interactions and instruction promote these skill domains (Pianta, 2005).

In the present study, we used a randomized controlled trial to evaluate the effects of teachers' participation in MyTeachingPartner on the language and literacy outcomes of children enrolled in these teachers' classrooms. Specifically, we were interested in the extent to which teachers' assignment to three forms of professional development support – 1) *Low Language/Literacy Support* (access to MTP Language/Literacy activities and web-based video exemplars of effective teacher-child interactions during these activities, referred to as Low Support for remainder of this paper), 2) *High Language/Literacy Support* (access to same resources as Low Support group with additional one-on-one web-based consultation, referred to as High Support for remainder of this paper); and 3) *Control* (only had access to the MTP Language/Literacy activities, without any supports) – predicted gains in children's development of language and literacy skills across the pre-kindergarten school year. Although MTP was only designed to support English-language instruction, participating

classrooms served a linguistically diverse population of young learners, and consequently, we also examined whether intervention effects would be moderated by the English-language composition of classrooms. We hypothesized that children's language and literacy skill development would improve the most in the High Support classrooms, particularly in comparison to Control classrooms. And, we anticipated that these effects of High Support would be strongest for children in classrooms where the linguistic composition was predominantly English-speaking.

Method

Participants

The study was conducted in state-funded pre-kindergarten programs across 42 school districts in a single Mid-Atlantic state. Participants were both teachers and students in these programs. This study included 161 pre-kindergarten teachers who participated for two consecutive years in the MyTeachingPartner professional development program. All participating teachers and classrooms were part of a state-funded pre-kindergarten program within a single state that serves children who meet "at-risk" status determined by the following criteria: 1) poverty; 2) homelessness; 3) parents or guardians are school dropouts, have limited education, or are chronically ill; 4) family stress as evidenced by poverty, episodes of violence, crime, underemployment, unemployment, homelessness, incarceration, or family instability; 5) child or developmental problems, or 6) limited English proficiency.

There were 161 teachers who were randomly assigned, clustered by district, into one of three study conditions and who participated in the intervention for two full school years. In addition, approximately four students within each teacher's classroom during each year were randomly selected to participate in the study, for a total of 1385 students across two cohorts, or two years of the study. As is typical of a large-scale study like this, we had the issue of missing data ranging approximately from 5% to 17%, depending on specific variables. To prevent excessive data loss due to missing data, we implemented a data imputation procedure for relevant variables (see "*Missing data treatment*" under "*Data Analysis Plan*" for more details of data imputation), and the final useable child sample size was 1338. The final useable child sample size ($N = 1338$) is less than 4% reduction of the original child sample size under these teachers, which we considered as minimal reduction for a study of this length and scale. The remaining 4% of child sample contained too much missing data on too many relevant variables, making it statistically not viable for data imputation for these cases.

Of the final useable sample, ninety-three percent of the participating teachers were women. The majority of teachers reported their race/ethnicity as Caucasian (68%), 26% reported African American, 3% reported Other, 2% reported Latina/Hispanic, and 1% left race/ethnicity missing. In terms of educational background, 62.1% had a bachelor's degree and 36% had advanced degrees, with 37.3% majoring in early childhood. Teachers reported an average of 15.65 years of classroom experience ($SD = 9.24$), with a range of 0 to 37 years. These teachers were leading classrooms that largely served children from economically disadvantaged (70% with an income-to-needs ratio less than 1.5; range = 0–100%) and linguistically diverse (25% dual language learners; range = 0–100%). Table 1 provides additional information on these teachers and their classrooms.

Of the total of 242 teachers originally selected to participate in the study, 33 dropped out of the study during the first year and 34 dropped out of the study during the second year, resulting in an attrition rate of 14% during year 1 and 16% during year 2. To estimate potential attrition bias, analyses were conducted comparing teacher and classroom characteristics for the 161 teachers who fully participated and the 81 teachers who did not

fully participate. Results indicate that there were no statistically significant differences between the two groups of teachers for the percentage that had an advanced degree ($\chi 2 = 0.64$, $p = 0.42$) or the percentage with certification to teach 4-year-olds ($\chi 2 = 0.72$, $p = 0.40$); however, teachers who fully participated in the study did have more years of experience teaching pre-kindergarten ($t = 4.40$, $p \leq 0.05$). In addition, there were no statistically significant differences between classrooms that did and did not fully participate in relation to the mean pretest scores for all five language/literacy assessments: definitional vocabulary ($t = -0.49$, $p = 0.63$), print knowledge ($t = -0.30$, $p = 0.76$), blending ($t = -1.15$, $p = 0.25$), elision ($t = -0.95$, $p = 0.34$), or the phonological awareness and literacy screening ($t = 0.44$, $p = 0.66$). Also, teachers who fully participated in the study were leading classrooms composed of a higher percentage of children who were dual language learners ($t = 18.46$, $p \leq 0.05$) and from families with an income-to-needs ratio lower than 1.5 ($t = 2.12$, $p \leq 0.05$).

Among the children enrolled in each participating teacher's classroom, four were selected for assessment of literacy and language outcomes. Consent to be selected for outcome assessment was obtained from parents for as many children in each classroom as possible, and all parents who consented for their child to be assessed also completed a demographic survey that included information on the child's disability status and home language. Children with an identified disability or IEP were excluded. Then within each classroom four children were selected randomly from those whose parents consented and who the teacher indicated could follow simple directions in English. Based on surveys completed by parents, annual family income was less than $15,000 for 29% of the families and between $15–25,000 for another 28% of families. Maternal education in these families averaged 12.75 (SD = 2.08) years, and close to 80% of families reported only speaking English at home. Table 1 provides additional information on the children and their families.

Recruitment and Intervention Procedures

Recruitment of teacher participants proceeded through several steps. First, invitation letters went to all district-level coordinators of the state-funded pre-kindergarten program; these letters described the study, the interventions, and that the study team was interested in recruiting teachers in each district's program to enroll in the study. Following the initial mailing, a meeting was held of all interested coordinators to describe the study in more detail and the procedures for recruitment of teachers. Forty-two district coordinators agreed to facilitate recruitment of teachers. These coordinators furnished contact information for pre-k program teachers and co-signed the letter of recruitment, indicating the district's permission for teachers to enroll in the study. Teachers received an individual letter inviting them to participate in the study, informing them that they were consenting to participate in the condition to which all the teachers in that district would be randomly assigned.

In the fall, prior to the start of the school year, teachers in each district attended a training and introductory workshop held at a convenient location. At the workshop, teachers were oriented to the purpose of the study, trained in the intervention to which they were assigned, and informed of a set of data-collection requirements. They also received a laptop computer. Over the course of the year all teachers received a series of MTP newsletters, reminders, and updates.

This cluster-randomized controlled trial was conducted over two school years (2004–2005 and 2005–2006), and all teachers participated in either the Control, Low Language/Literacy Support, or High Language/Literacy Support groups for all or part of both years. Random assignment was conducted at the district level for several reasons. First, in initial meetings with district coordinators it was clear they preferred all teachers in their program to receive the same professional development opportunities. Nearly all of these pre-k program

classrooms (and teachers) were organized by district-level program, not by school building, in terms of their training, professional development, curriculum, and program procedures. Thus "district" was a distinct program-relevant feature. On a related note, there was concern about contamination of intervention effects across conditions if teachers in the same district (often in the same building) were enrolled in different conditions. For example, teachers participating in the High Support intervention would, by definition, be exposed to more detailed descriptions of effective practices and receive direct facilitation of their use of the video exemplars on the website. Thus, they could easily "tutor" Low Support teachers in their district in the use of the site's resources, thereby potentially reducing the key distinction between these conditions. We sought to eliminate this possibility through random assignment at the district level and recognized that causal inferences cannot be made at the classroom level. Assignment by district was also stratified by district size (in terms of number of classrooms in the pre-k program) before teachers were recruited individually. Districts were classified into large, medium, and small, according to numbers of pre-k classrooms, then assigned randomly by size to condition.

The MyTeachingPartner (MTP) intervention was a web-based professional development program for pre-kindergarten teachers consisting of three different conditions – Control, Low Language/Literacy Support, and High Language/Literacy Support, as follows:

1. *Control*. Teachers received lesson plans for MyTeachingPartner-Language & Literacy activities (MTP-LL; Justice, Pullen, Hall, & Pianta, 2003) in print and on-line form. The literacy activities focused on developing children's alphabet knowledge, print concepts and phonological awareness. The language activities focused on developing children's narrative skills, understanding and use of new vocabulary, understanding of linguistic concepts, and pragmatic and social communication skills. These activities were not meant to serve as a stand-alone curriculum, but instead to offer a language/literacy supplement to any existing curricula.
2. *Low Language/Literacy Support*. Teachers received the same MTP-LL activities listed above. In addition, they received access to the MTP website, including access to descriptions and demonstrations of the MTP-LL activities, video exemplars of high-quality interactions during MTP-LL activities, and teaching challenges that encouraged teachers to examine video clips of classroom interactions, consider how they would behave in the situation, and then to compare their ideas with an expert response to the situation. These teachers also received print copies of the Preschool PATHS-Promoting Alternative Thinking Strategies curriculum (Domitrovich, Greenberg, Kusche, & Cortes, 2004) intended to promote social and emotional development.
3. *High Language/Literacy Support*. Teachers received the same MTP-LL activities and website access listed above, as well as the PATHS curriculum. In addition, they participated in a one-on-one consultation process aimed at improving the implementation of language/literacy activities through effect student-teacher interactions. This consultation process addressed teachers' individual needs by having teachers videotape their teaching practices for 30 minutes every other week. The teacher then sent the video to his or her consultant who selected a short segment from the video and posted it on the secure MTP website along with a prompt or question to which the teacher responded. The prompts and questions were designed to provide targeted feedback to pre-k teachers through a standardized protocol that focused on specific dimensions of teachers' emotional, organizational, and instructional interactions with students, using the Classroom Assessment Scoring System (CLASS; Pianta et al., 2007). After the teacher

responded, the teacher and consultant met via videoconference to discuss teaching practices and set future goals.

Measures

Data were collected on a variety of teacher, classroom, and child characteristics. Pre-kindergarten teachers completed questionnaires regarding demographic information (e.g., degree status, years of teaching experience), attitudes and beliefs about children, and sense of self-efficacy. Information regarding individual pre-kindergarten classrooms was also collected, including the percent of children who were dual language learners, and the average family income. During each school year, the four selected children within each participating classroom were assessed at the beginning and end of the pre-kindergarten school year using a direct assessment of pre-literacy skills.

For children's language and literacy skills, we used the individually-administered direct assessment Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP; Lonigan, Wagner, Torgesen, & Rashotte, 2002), which consists of four subscales: Vocabulary, Blending, Elision, and Print. The Pre-CTOPPP was administered at the beginning of the fall semester (baseline measures) and near the end of the spring semester (outcome measures). It was designed for use with children from 3 to 5 years of age and is a precursor to the slightly revised, recently published Test of Preschool Early Literacy (Lonigan, Wagner, & Torgesen, 2007). The Pre-CTOPPP provides scores from four subtests: blending, elision, print awareness, and receptive vocabulary. Based on consultation with the assessment's authors, raw scores are reported and used in analyses.

The Blending subtest includes items that measure whether children can blend initial phonemes onto one-syllable words, initial syllables onto two-syllable words, and ending phonemes onto one-syllable words. The Elision subtest measures whether children can break apart initial and ending phonemes, as well as initial syllables, from one- and two-syllable words. Print Awareness items measure whether children recognize individual letters and letter-sound correspondences, and whether they differentiate words in print from pictures and other symbols. And, Receptive Vocabulary items measure children's word knowledge. The Pre-CTOPPP subtests have shown adequate internal consistency, test-retest reliability, and concurrent validity in past research by the test developers and in several large, federally funded studies, including the Head Start Impact Study, IES Even Start Classroom Literacy Interventions and Outcomes Study, IES Preschool Curriculum Evaluation Research Study, and IES Early Reading First National Evaluation (Lonigan, McDowell, & Phillips, 2004; National Center for Education Evaluation, 2007).

Teachers in this study administered the Pre-CTOPPP after receiving training by the investigators. Small-scale pilot tests have demonstrated the reliability and validity of data collected by assessments administered by teachers in Head Start classrooms (Vogel, Nogales, Aikens, & Tarullo, 2008). In addition, these teachers had administered a similar, statewide literacy/language assessment to children in their classrooms in the past and were therefore very familiar with standardized testing procedures. At the beginning of the project, all teachers completed training focused on administration of the language and literacy battery, and fidelity of administration was randomly checked via videotape for 20% of teachers in the Fall of year 1. Teachers accurately administered standardized items over 90% of the time and reported that for 96% of the assessments children's performances were "most typical" or "very typical" of their usual classroom functioning.

Data Analysis Plan

We were interested in knowing if the intervention conditions (High Support, Low Support, and Control) had any effect on children's language and literacy skills, as measured by the Pre-CTOPPP. For our analysis, we carefully considered and weighed the advantages and disadvantages of a multivariate analytical framework (structural equation modeling; SEM) and a univariate analytical framework (hierarchical linear modeling; HLM), as detailed below.

Hierarchical linear modeling (HLM) framework—The most obvious statistical advantage of a HLM analytical framework is to take care of the issue of clustering sampling design and its resultant design effect in the analysis. Using HLM for cluster sample data would provide a correct standard error of a sample statistical estimator (Raudenbush & Bryk, 2002). Because many educational datasets have nested data structure as a result of cluster sampling design, HLM has become a popular analytical approach. HLM, however, is a univariate analytical framework; as a result, it has its own disadvantages. First, HLM is essentially a univariate regression analysis strategy; as such, it has no mechanism for accommodating multiple outcome variables in the same analysis, because it only analyzes one outcome variable at a time, and therefore ignores any relationships among multiple outcome variables. The second disadvantage of HLM is its inability to take measurement error into consideration while estimating model coefficients of research interest. As is well known in the research and measurement literature (e.g., Fan, 2003; Gulliksen, 1987), measurement error may considerably attenuate the relationships among variables/constructs, thus making it much more difficult to detect true relationships in statistical analysis. For our research involving young children, measurement error is a serious concern, because it is usually more difficult to obtain reliable measurements for these children.

Structural equation modeling—Structural equation modeling (SEM) joins methods from econometrics, psychometrics, sociometrics, and multivariate statistics (Bentler, 1994). Many statistical techniques can be considered as special cases of SEM, including regression analysis, canonical correlation analysis, confirmatory factor analysis, and path analysis (Bentler, 1994; Fan, 1996; Jöreskog & Sörbom, 2001). SEM offers some unique advantages, in comparison to HLM. First, it is a multivariate analytical framework, and multiple outcome variables with their relationships can be modeled in the same analysis. Second, SEM analysis not only models the relationships among multiple outcome variables/constructs, but it also models the measurement errors of the observed indicators. In other words, SEM can statistically correct for the attenuation of measurement error on the relationships among the variables/constructs, and provides better estimates of the true relationships among the variables of research interest (Fan, 2003; Jöreskog & Sörbom, 2001). As Thompson (1994) discussed, "The failure to consider score reliability in substantive research may exact a toll on the interpretations within research studies. For example, we may conduct studies that could not possibly yield noteworthy effect sizes given that score reliability inherently attenuates effect sizes. Or we may not accurately interpret the effect sizes in our studies if we do not consider the reliability of the scores we are actually analyzing" (p. 840). After considering the advantages and disadvantages of two analytical frameworks (HLM and SEM), we decided to use SEM as the major analytical tool for our analysis. At the same time, we statistically corrected for the clustering sampling effect while estimating the model coefficients of our research interest, as detailed below, in order to avoid the problem of inflated Type I error rate.

Cluster sampling and design effect—We had a cluster-sampling design in our study, with classroom (teacher) being the higher sampling unit, and children being nested under classrooms (on average, about four children with direct assessment measures in each

classroom). Cluster sampling presents some analytical challenges for data analyses. Standard statistical procedures assume that data are collected by simple random sampling (SRS), and violation of this assumption creates problems for data analyses (e.g., Cochran, 1977; Kish, 1965; Kott, 1991; Lee, Forthofer, & Lorimor, 1989; NCES, 2002). The effect of cluster sampling design is usually quantified by "design effect" (Kish, 1965; NCES, 2002), which is the ratio of the correct variance of an estimator under the cluster sampling design to the variance of the estimator obtained while ignoring the cluster sampling design (i.e., assuming a simple random sample). Typically, design effect is greater than one, indicating larger variation of a statistic from a cluster sample than from a simple random sample. In data analysis, ignoring clustering sampling design (and its design effect) would result in inflated Type I error in statistical significance testing. It is important to note that design effect affects the standard error of a statistic only, but not the statistic itself. In other words, design effect does not cause biased estimates, but only causes higher Type I error rate in statistical inferential testing.

For these reasons, we estimated the design effect, and took into consideration the design effect while testing for the model coefficients in SEM modeling analysis. As is known in the statistical literature, design effect (DEFF) of a cluster sampling design is defined as (NCES, 2002):

$$DEFF=\frac{(Design\ SE)^2}{(SRS\ SE)^2}$$

where "Design SE" is the standard error from cluster sampling design, and "SRS SE" is the standard error assuming simple random sample design. In research applications, the square root of design effect (also known as root design effect, or DEFT):

$$DEFT=\frac{Design\ SE}{SRS\ SE}$$

is more useful (NCES, 2002), because DEFT can be used as a correction factor in statistical testing in order to avoid the problem of inflated Type I error caused by cluster sampling design.

In research practice, root design effect (DEFT) can be estimated as (e.g., Kerry & Bland, 1998):

$$\text{DEFT}=\sqrt{1+(n-1)\rho}$$

where n is the cluster sample size and ρ is the intraclass correlation coefficient (ρ) of the nested data. In our later SEM analysis, we used DEFT as a correction factor for testing relevant model parameter estimates such that the potential of inflated Type I error due to cluster sampling design was avoided.

Missing data treatment—As is typical in longitudinal studies, this study has missing data; specifically, the proportions of missing values on various variables (e.g., demographic variables, and both pretest and post-test direct assessment measures of language and literacy skills) ranged from 5% to about 17%. Missing data can be treated in different ways. The simplest approach is to do a listwise deletion (i.e., eliminating any observations with missing values on the variables used in an analysis), and conduct analysis by using the cases with

complete data. However, as discussed in the literature, listwise deletion is not a statistically optimal approach, because of the loss of statistical efficiency and potential statistical bias that may be introduced by such a simplistic treatment of missing data. A different approach is to attempt to estimate the missing values through data imputation. For this purpose, there are more traditional single imputation approaches (e.g., mean substitution, regression-based substitution), and more recent multiple imputation approaches. The current "gold standard" for missing data treatment is through multiple imputation (e.g., Little & Rubin, 2002; Rubin, 1996), as single substitution does not reflect the uncertainty of the predictions of the missing values. As Rubin (1996) discussed, as few as three to five imputations should be adequate in multiple imputation. In our analysis, we used multiple imputation (five imputations) to address the missing data issue in this study, and used imputed data for our modeling analysis.

Covariates and moderators—We considered several potential covariates to be included in the modeling analysis (e.g., children's age, ethnicity and family income). In preliminary analyses once the pre-test measures of the Pre-CTOPPP (administered at the beginning of the fall) were included as covariates in the model, all other covariates were not statistically related to the outcome measures (i.e., post-test direct assessment measures of language and literacy skills). For the sake of model parsimony, we made the decision to not include these demographic variables as covariates in presentation of the final model results.

As was noted earlier, a constraint of the present study is that the results focus on children assessed in classrooms in which the teacher reported that the primary language of instruction was English and the children, if they spoke a language other than English, were reported by their teacher to be capable of following simple verbal directions in English. However, despite this initial screening mechanism there were a variety of languages spoken by the students for whom we obtained outcome assessments. Classrooms varied in the extent to which English was the only language spoken and that Spanish was used as the language of informal interaction. In the demographic and classroom survey completed by teachers, there were two items that asked teachers to report (a) if Spanish was spoken or not by the children in class; and (b) if other languages were spoken or not by the children in class. Because the professional development resources provided to teachers (lesson/activity guides and instructional materials in language and literacy; web-based video exemplars of effective interactions; and web-mediated consultation) were developed in English and assumed English as the language of instruction, there was reason to believe that children speaking Spanish or other non-English languages in the classroom could moderate the effects of these professional development resources on children's performance in English on assessments of language and literacy. For this reason, we examined this dichotomous variable (i.e., children spoke Spanish or other languages in class, versus those where the children only spoke English in class) as a moderator variable for the hypothesized intervention effect.

Modeling intervention effects—Within the SEM modeling framework, we hypothesized that intervention conditions (High Support, Low Support, and Control) would differentially affect children's performance on language and literacy skills at the end of academic year (post-test measures taken near the end of spring), while controlling for their performance at the beginning of the year (pre-test measures taken at the beginning of the prior fall). Because the Pre-CTOPPP has four scales (Vocabulary, Blending, Elision, and Print), we modeled the potential intervention effect on the latent construct of "language/literacy" with four measured indicators. This latent variable modeling approach not only allowed us to conduct the analysis within a multivariate framework, but also allowed us to take measurement error into consideration in the multivariate analysis. The basic SEM model was conceptually a multivariate ANCOVA model in which the potential effects of

intervention conditions on post-test language/literacy were modeled while controlling for prior achievement and taking measurement error into account in the modeling process.

Results

Average Change Pre- to Post- for Language/Literacy Skills in Pre-K

Table 2 presents the descriptive statistical information for the language/literacy measures (pre- and post) for the intervention groups under the two conditions of the moderator variable discussed previously. There was considerable growth on the language/literacy measures from pre-test to post-test. For example, for the High Support group under the first moderator condition ("Only English spoken in class by students"), growth on the measure of Print is approximately 1.62 standard deviations: (29.15 − 16.07) / [(7.30 + 8.81)/2] = 1.62. Other measures of language/literacy for different groups also showed substantial magnitudes of growth (typically close to or above one standard deviation) from pre-test measures to post-test measures. In sum, on average, children in all conditions showed considerable gains in early literacy and language skills during the pre-k year.

At the same time, a closer look at Table 2 data also reveal some growth differences from pre-test to post-test measures among the intervention groups. For example, as described above, the High Support group under the first moderator condition showed approximately 1.62 standard deviation growth from pre- to post-test, while for the same moderator condition, the Control group showed less change, 1.24 standard deviation: (26.89 − 16.15) / [(8.00 + 9.47)/2] = 1.24. Similarly, under the second moderator condition ("Spanish and other languages spoken in class by students"), for the measure of Print, the Low Support group showed about 1.86 standard deviation growth: (29.24 − 15.62) / [(6.66 + 7.96)/2] = 1.86, while the Control group showed 1.42 standard deviation growth: (27.23 − 15.29) / [(7.75 + 9.02)/2] = 1.42. The primary point here is that there appears to be some variation across groups in the extent of growth in language and literacy scores, and the next set of analyses is designed to test the extent of these observed descriptive differences.

Analytic Modeling Results

We primarily were interested in knowing if the High Support and Low Support intervention conditions would lead to children's better performance on language/literacy outcomes, relative to the Control group. It should be noted that we did not have a true "control group" in this study, because even the Control group did receive some form of intervention that could have contributed to children's gains in language and literacy skills. This lack of a true "control group" could have attenuated the estimates of the High Support and Low Support intervention effects.

As discussed previously, we used a structural equation modeling (SEM) approach as our primary analysis tool, while correcting for the effect of clustered sampling on the standard error of model parameter estimates. The SEM model was conceptually a multivariate ANCOVA model in which the potential effects of intervention conditions on post-language/literacy were modeled while controlling for prior achievement and taking measurement error into account in the modeling process.

Correcting for clustering effect—For the four univariate outcome measures (post-language/literacy measures that we used in the modeling analysis), the intraclass correlations ranged from 0.12 to 0.19, very typical of intraclass correlation magnitudes observed in some large scale educational datasets (e.g., *High School and Beyond*). The cluster sample size (i.e., children under each participating classroom teacher) is approximately eight. To correct for the cluster sampling in statistical testing, we chose to use the average of the intraclass

correlations (0.17) to derive the root design effect (DEFT) for correcting for standard error of parameter estimate from SEM analysis:

$$\text{DEFT}=\sqrt{1+(n-1)\rho}=\sqrt{1+(8-1)0.17}=1.48$$

The DEFT of 1.48 indicates that the standard error from the cluster sampling is 1.48 times larger that the standard error obtained from SEM modeling analysis, which assumed simple random sampling. Again, this correction factor is very typical in educational data (e.g., NCES, 2002). We applied this correction factor while conducting statistical significance testing for SEM model parameter estimates described in the following.

The model and modeling results are presented in Figure 1. As discussed previously, we considered the moderator variable describing whether or not "Spanish/other languages were spoken by children in class" as reported in the teacher survey, and we conducted modeling analyses for the two conditions under this moderator variable, as shown in (a) and (b) of Figure 1. Table 3 presents the model fit assessment for models of both moderator conditions, including the major model fit indices.

It is noted that the model for both moderator conditions showed very good model fit, with all major model fit indices suggesting that the model fit the data very well. For example, *RMSEA*, a widely used model fit index, is 0.05, a value typically considered as representing good model fit (e.g., Browne & Cudeck, 1993; Steiger, 1989). The probabilities of testing for close fit are also statistically non-significant, leading to the conclusion that one should *not* reject the model. All other fit indices would obviously lead to similar conclusion about model fit.

The factor loadings for the observed language/literacy measures (Vocabulary, Blending, Elision, Print) suggest that these measures are statistically reasonable indicators for the language/literacy latent construct, but each measure has considerable amount of residual (i.e., variance not shared with other measures). It is likely that part of the residual for each observed measure (e.g., vocabulary) represents its unique content, and such unique content would materialize as correlated residuals across time. For this reason, we hypothesized correlated residual variances for the same measure across two times (pre- and post-tests), as shown in the model. Not surprisingly, the standardized coefficients from prior achievement (Pre-Language/Literacy) to later performance (Post-Language/Literacy) are strong (0.85 and 0.84, respectively for the two moderator conditions). With R^2 = 0.73 and 0.72 for the outcome construct (Post-Language/Literacy) in the two conditions, the model accounts for the post-language/literacy outcome measures well.

It is observed that model fit and model estimates for the two conditions of the moderator variable are remarkably similar, except the hypothesized intervention effects: the dummy coded comparison between High Support and Control, and that between Low Support and Control. For the children in the classrooms where "Only English was spoken in class by students", relative to the Control condition, consultation showed a positive effect on children's literacy outcome measures, after controlling for the prior achievement (fall measures). The effect (0.11 in the form of standardized regression coefficient) is statistically significant after correcting for the cluster sampling effect, and this could be characterized as a "small" effect. On the other hand, the effect (0.06) of Low Support condition in contrast to the Control condition is statistically non-significant.

For the children in the classrooms where Spanish/other languages *were* spoken in class, we hypothesized muted effects of any of the conditions (because of selection procedures that

stipulated a focus of instruction on English). Indeed, after controlling for the prior achievement (fall measures), neither the comparison of the High Support condition versus the Control condition (standardized coefficient of −.05), nor the comparison of the Low Support condition versus the Control condition (standardized coefficient of .04), were statistically significant.

Discussion

In this paper, we demonstrate the positive effects of MyTeachingPartner, a web-based system of professional development resources, on children's language and literacy development, as a function of their teachers' exposure to certain forms of support. High levels of support for teachers' implementation of language/literacy activities showed modest but significant effects for improving early language and literacy for children in classrooms where English was the dominant language spoken by the students and teachers. This combination of web-based supports, including video-based consultation *and* web-based video teaching exemplars, was more effective in improving literacy and language skills than was only providing teachers with a set of instructional materials and detailed lesson guides. These results have implications for designing systems of support for effective teaching in early childhood programs.

Before discussing the results and their implications in greater detail, it is important to consider that this study was a rather conservative test of professional development effects for two reasons. First, these were intent-to-treat analyses focused on *child outcomes*, when in fact the primary target of intervention was the quality of teachers' interactions with children (Pianta et al., 2008). Although the intervention was certainly designed to have an effect on children's language and literacy skills, the actual nature of the intervention was focused on teacher behavior. The second reason this study may be a conservative test of intervention efficacy is that we lacked a true control group in which "business as usual" was the condition and there was an absence of investigator-provided supports for teachers. Indeed, the "control group" in this case received a set of instructional materials specially designed to promote language and literacy development that included lessons with specific guides for how to implement an activity well. It is possible that these resources could have contributed to gains in children's skills that were greater than those for a "business as usual" control and therefore narrowed the difference between the High Support and Low Support intervention conditions and the Control group.

The pattern of evidence emerging from empirical work on consultation and coaching in early childhood education indicates quite clearly the benefits of focused consultation for improving teacher behavior (Burchinal, 2008). In prior work, we have established the benefit of teachers' engagement in ongoing, one-on-one consultation and feedback loops focused on teacher-child interactions (Pianta et al., 2008). This positive effect of focused coaching or consultation for improving teachers' interactions with children has been found in other recent studies of early childhood classrooms (Domitrovich, Gest, Gill, Bierman, Welsh, & Jones, 2009; Neuman & Wright, 2010; Raver, Jones, Li-Grining, Metzger, Smallwood, & Sardin, 2008) and in the k-12 grades (e.g., Cantrell & Hughes, 2008; Sailors & Shanklin, 2010; Walpole, McKenna, Uribe-Zarain, & Lamitina, 2010).

Such results have promise for addressing a number of concerns related to both the capacity of the early childhood workforce to foster desired changes in developmental skills of children and the capacity of professional development systems to adequately and effectively support this effort (Zaslow & Martinez-Beck, 2005). However, to date, research and development efforts focused on identifying effective and replicable approaches to teacher training that produce positive, and even accelerated, gains in children's academic

performance have come up short, particularly empirically-supported approaches that address the needs of a large number of teachers (Brandon & Martinez-Beck, 2005; Ramey & Ramey, 2005). The present study extends findings regarding the value of consultation for changing teachers' behavior by demonstrating benefits for children's learning outcomes within a large, diverse sample. Continued rigorous research and development in this area is clearly needed, in part to establish the extent to which the gains in children's learning outcomes associated with teacher consultation are mediated by changes in teacher behavior, as would be suggested by non-experimental work (NICHD ECCRN, 2002), and to further examine questions related to dosage, engagement, costs, and focus of consultation supports.

The MyTeachingPartner approach was developed in response to the need for professional development and training of teachers (Zaslow & Martinez-Beck, 2005) that focuses on embedding support for children's early literacy and language competence within interactions throughout the day (Dickinson & Brady, 2005). This approach targets the improvement of teacher-child interactions, and subsequently children's language and literacy development, through consultation and video-based teaching exemplars that are tied to a standardized and validated observational measure of interactions, and in the present study there were differences with regard to the effects of this support on children's outcomes conditioned upon the linguistic diversity of children in these classrooms. In particular, teachers' engagement in ongoing, web-mediated consultation showed modest but significant effects for improving early language and literacy for children, but only in classrooms where English was the dominant language spoken by children.

As noted earlier, we established with previous work that consultation produces improvements in the quality of teachers' interactions with children, so this downstream effect on students was to be expected. At first, the fact that the benefits of consultation did not extend to students in classrooms in which Spanish or other languages were spoken was somewhat surprising. After all, two contemporary research reviews recently indicated that high quality instructional practices, such as emotionally, organizationally, and instructionally supportive teacher-child interactions, seem to be just as beneficial to dual language learners as non-DLL children (August & Shanahan, 2006; Genesee, Lindholm-Leary, Saunders, & Christian, 2006). And recent findings indicate that the teacher-child interactions targeted by MTP are positively linked to early social and academic competence, regardless of a child's primary home language (Downer, Lopez, Hamagami, Howes, & Pianta, 2009). Then again, the magnitude of the impacts of high quality, global instructional practices are typically more modest for DLL children than for primarily English-speaking children (August & Shanahan, 2006). And perhaps most important, there is growing consensus that DLLs learn to read best when first taught in their primary home language, or at the very least when instruction is in some combination of English and their primary home language (Espinosa, 2007; Garcia & Jensen, 2009; Vaughn, Linan-Thompson, Pollard-Durodola, Mathes, & Cardenas Hagan, 2006). In this study, teachers in all classrooms reported that the language of instruction was English, so the professional development supports were specifically developed to address English only instruction. It may be that the utility of focused consultation around teacher-child interaction, for improving children's language development and literacy skills, was limited because it was not well-matched to the set of classroom processes necessary to ensure such gains for all children. As the number of dual language learners continues to grow at a rapid pace in the United States (Hernandez, Denton, & Macartney, 2007), there is a pressing need to better understand what combination of high quality global instructional practices and DLL-specific strategies best support children's learning, as well as ensuring that early childhood teachers receive training and support to do this well.

There were a number of limitations in the present study that constrain inferences and implications made on the basis of the reported results. First, although this was a cluster-randomized study, children were not assigned randomly to teachers, and so there could be some selection bias that was not estimated and limits inferences about causality of the interventions. Second, we did not formally test mediation of the detected intervention effects on children's gains as a function of changes in teacher behavior. This is the focus of a set of analyses currently underway. In addition, there are a number of other putative moderator variables that could be examined in subsequent studies, including the skill level of the children or other features of the classroom setting. Further work will attempt to address these and other factors that could help understand the nature and limits of the efficacy of these resources. Finally, it is amply evident that in the present study we both constrained the sampling plan in ways that reduced the generalizability of the findings to state-funded pre-k contexts in which there are very large numbers of children who are dual language learners and instruction may be conducted in languages other than English, and limited the inferences that could be drawn about effects on DLL children.

Finally, we call attention to these results for developing new models of professional development for early childhood educators. We argue that the professional development needs of teachers, particularly early childhood educators, should be conceptualized not in terms of the hypothesized knowledge that accrues as a function of credentials or degrees, but rather direct supports that are relevant and linked to teachers' actual demonstrated skill in the classroom. If early education programs are going to achieve high quality at scale (Pew Charitable Trusts, 2005), then new and effective mechanisms of training teachers must be developed and tested both in preservice teacher training and in alternate certification and re-training routes used by large school districts or alternative suppliers (Birman et al., 2000; Borko, 2004; Clifford & Maxwell, 2002; Cochran-Smith & Zeichner, 2005; Pianta, 2005). Results of the present study suggest the promise of web-mediated interactions with professional development resources that are video-based, individualized, and tied to valid conceptual and assessment frameworks for identifying effective practices and interactions, in which the target is teachers' delivery of instruction and provision of social and emotional supports. Such approaches differ in several ways from traditional courses or even web-based delivery of course content in that they are scalable as well as closely linked to practice. Clearly, as state and Federal policy-makers and educational leaders work to build an effective system of early childhood education in the United States, this study and other relevant work (Bierman et al., 2008; Raver et al., 2008) can inform efforts to supply effective professional development supports to an ever-growing number of early educators and identify appropriate incentive and credentialing systems that maximize the value of such supports.

Acknowledgments

The development of this paper was supported by a grant awarded to Dr. Robert Pianta by the Department of Health and Human Services (DHHS) -- National Institute of Child Health and Human Development (NICHD), Administration for Children and Families (ACF) and Office of the Assistant Secretary for Planning and Evaluation (ASPE); and the Department of Education (ED) -- Office of Special Education and Rehabilitative Services (Grant #5R01HD046061-03). The opinions expressed herein are those of the investigators and do not reflect the views of the funding agencies. We extend our gratitude to the set of teachers who provided us with the opportunity to experiment with new ways of supporting them.

References

Anthony JL, Lonigan CJ, Driscoll K, Phillips BM, Burgess SR. Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. Reading Research Quarterly. 2003; 38:470–487.

August, D.; Shanahan, T., editors. Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth. Mahwah, NJ: Lawrence Earlbaum; 2006.

Bentler, PM. Forward. In: Byrne, BM., editor. Structural equation modeling with EQS and EQS/Windows (vii–ix). Thousand Oaks, CA: Sage Publications; 1994.

Bierman KL, Domitrovich CE, Nix R, Gest S, Welsh JA, Greenberg MT, Blair C, et al. Promoting academic and social-emotional school readiness: The Head Start REDI program. Child Development. 2008; 79:1802–1817. [PubMed: 19037951]

Birman BF, Desimone L, Porter AC, Garet MS. Designing professional development that works. Educational Leadership. 2000; 57:28–33.

Bogard K, Takanishi R. PK-3: An aligned and coordinated approach to education for children 3 to 8 years old. SRCD Social Policy Report. 2005; 19:1–23.

Borko H. Professional development and teacher learning: Mapping the terrain. Educational Researcher. 2004; 33:3–15.

Bowey J. Socioeconomic status differences in preschool phonological sensitivity and first-grade reading achievement. Journal of Educational Psychology. 1995; 87:476–487.

Brandon, R.; Martinez-Beck, I. Estimating the size and characteristics of the United States early care and education workforce. In: Zaslow, M.; Martinez-Beck, I., editors. Critical issues in early childhood professional development. Baltimore, MD: Paul H. Brookes; 2005. p. 49-76.

Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. Testing structural equation models. Thousand Oaks, CA: SAGE; 1993. p. 136-162.

Bryant P, Maclean M, Bradley L. Rhyme, language, and children's reading. Applied Psycholinguistics. 1990; 11:237–252.

Bunce, BH. Building a language-focused curriculum for the preschool classroom. Baltimore, MD: Paul H. Brookes; 1995.

Burchinal, MR. Competencies and credentials for early childhood educators: What do we know and what do we need to know?. Invited keynote presented at Head Start's 9th National Research Conference; Washington, DC. 2008.

Burchinal MR, Roberts JE, Riggins R, Zeisel SA, Neebe E, Bryant D. Relating quality of center child care to early cognitive and language development longitudinally. Child Development. 2000; 71:339–357. [PubMed: 10834469]

Burchinal M, Howes C, Pianta RC, Bryant D, Early D, Clifford R, Barbarin O. Predicting child outcomes at the end of kindergarten from the quality of pre-kindergarten teacher-child interactions and instruction. Applied Developmental Science. 2008; 12:140–153.

Byrne B, Fielding-Barnsley R. Evaluation of a program to teach phonemic awareness to young children: A one-year follow-up. Journal of Educational Psychology. 1993; 85:104–111.

Byrne B, Fielding-Barnsley R. Evaluation of a program to teach phonemic awareness to young children: A 2- and 3- year follow-up and a new preschool trial. Journal of Educational Psychology. 1995; 87:488–503.

Cantrell SC, Hughes HK. Teacher efficacy and content literacy implementation: An exploration of the effects of extended professional development with coaching. Journal of Literacy Research. 2008; 40:95–127.

Catts HW, Fey ME, Zhang X, Tomblin JB. Language basis of reading and reading disabilities: Evidence from a longitudinal investigation. Scientific Studies of Reading. 2001; 3:331–361.

Chaney C. Preschool language and metalinguistic skills are links to reading success. Applied Psycholinguistics. 1998; 19:433–446.

Christensen CA. Onset, rhymes, and phonemes in learning to read. Scientific Studies of Reading. 1997; 1:341–358.

Clifford, RM.; Maxwell, K. The need for highly qualified prekindergarten teachers. Presented at the Preparing Highly Qualified Prekindergarten Teachers Symposium at Frank Porter Graham Child Development Institute; University of North Carolina; 2002 Apr.

Cochran, WG. Sampling techniques. New York, NY: Wiley; 1977.

Cochran-Smith, M.; Zeichner, K. Studying teacher education: The report of the AERA panel on research and teacher education. Mahwah, NJ: Erlbaum; 2005.

Cunningham AE, Perry KE, Stanovich KE, Stanovich PI. Disciplinary knowledge of K-3 teachers and their knowledge calibration in the domain of early literacy. Annals of Dyslexia. 2004; 54:139–167. [PubMed: 15765007]

Dickinson, DK.; Brady, J. Toward effective support for language and literacy through professional development: A decade of experiences and data. In: Zaslow, M.; Martinez-Beck, I., editors. Critical issues in early childhood professional development. Baltimore, MD: Paul H. Brookes; 2005. p. 141-170.

Domitrovich CE, Gest SD, Gill S, Bierman KL, Welsh J, Jones D. Fostering high quality teaching with an enriched curriculum and professional development support: The Head Start REDI program. American Educational Research Journal. 2009; 46:567–597.

Domitrovich, CE.; Greenberg, MT.; Kusche, C.; Cortes, R. The preschool PATHS curriculum. State College, PA: Pennsylvania State University; 2004. Unpublished curriculum.

Downer, J.; Lopez, M.; Hamagami, A.; Howes, C.; Pianta, R. The Classroom Assessment Scoring System in Dual Language Learner classrooms: Factor structure and predictive validity. Poster session presented at the 4th Annual IES Research Conference; Washington, DC.

Espinosa, LM. English-Language Learners as they enter school. In: Pianta, RC.; Cox, MJ.; Snow, KL., editors. School readiness & the transition to kindergarten in the era of accountability. Baltimore, MD: Paul H. Brookes; 2007. p. 175-196.

Fan X. Structural equation modeling and canonical correlation analysis: What do they have in common? Structural Equation Modeling. 1996; 4:64–78.

Fan X. Attenuation of correlation by measurement unreliability and two approaches for correcting the attenuation. Educational and Psychological Measurement. 2003; 63:915–930.

Gallagher A, Frith U, Snowling MJ. Precursors of literacy-delay among children at genetic risk of dyslexia. Journal of Child Psychology & Psychiatry. 2000; 41:203–213. [PubMed: 10750546]

Garcia E, Jensen B. Early educational opportunities for children of Hispanic origins. SRCD Social Policy Report. 2009; 23:3–19.

Genesee, F.; Lindholm-Leary, K.; Saunders, W.; Christian, D. Educating English language learners: A synthesis of research evidence. New York, NY: Cambridge University Press; 2006.

Girolametto L, Pearce PS, Weitzman E. Interactive focused stimulation for toddlers with expressive vocabulary delays. Journal of Speech and Hearing Research. 1996; 39:1274–1283. [PubMed: 8959612]

Girolametto L, Weitzman E, Clements-Baartman J. Vocabulary intervention for children with Down syndrome: Parent training using focused stimulation. Infant Toddler Intervention: A Transdisciplinary Journal. 1998; 8:109–126.

Gulliksen, H. Theory of mental tests. Hillsdale, NJ: Lawrence Erlbaum; 1987.

Hammill DD. What we know about correlates of reading. Exceptional Children. 2004; 70:453–468.

Hamre, BK.; Pianta, RC. Learning opportunities in preschool and early elementary classrooms. In: Pianta, R.; Cox, M.; Snow, K., editors. School readiness & the transition to kindergarten in the era of accountability. Baltimore, MD: Paul H. Brookes; 2007. p. 49-84.

Hamre BK, Pianta RC. Can instructional and emotional support in the first grade classroom make a difference for children at risk of school failure? Child Development. 2005; 76:949–967. [PubMed: 16149994]

Harms, T.; Clifford, RM.; Cryer, D. The Early Childhood Environment Rating Scale: Revised edition. New York, NY: Teachers College Press; 1998.

Hart, P.; Stroot, S.; Yinger, R.; Smith, S. Meeting the teacher education accountability challenge: A focus on novice and experienced teacher studies. Mount Vernon, OH: Teacher Quality Partnership; 2005.

Hernandez, DJ.; Denton, NA.; Macartney, SE. Demographic trends and the transition years. In: Pianta, RC.; Cox, MJ.; Snow, KL., editors. School readiness & the transition to kindergarten in the era of accountability. Baltimore, MD: Paul H. Brookes; 2007. p. 217-282.

Howes C, Burchinal M, Pianta R, Bryant D, Early D, Clifford R, et al. Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. Early Childhood Research Quarterly. 2008; 23:27–50.

Hyson, M.; Biggar, H. NAEYC's standards for early childhood professional preparation: Getting from here to there. In: Zaslow, M.; Martinez-Beck, I., editors. Critical Issues in Early Childhood Professional Development. Baltimore, MD: Paul H. Brookes; 2005. p. 283-308.

Jöreskog, KG.; Sörbom, D. LISREL 8: User's reference guide. Lincolnwood, IL: SSI Scientific Software International; 2001.

Justice, L.; Pullen, P.; Hall, S.; Pianta, R. MTP Activities for Early Literacy and Oral Language Support. Charlottesville: University of Virginia; 2003. Unpublished manual.

Justice LM, Ezell HK. Enhancing children's print and word awareness through home-based parent intervention. American Journal of Speech-Language Pathology. 2000; 9:257–269.

Justice LM, Ezell HK. Use of storybook reading to increase print awareness in at-risk children. American Journal of Speech-Language Pathology. 2002; 11:17–29.

Justice LM, Ezell HK. Knowledge of syntactic structures: A comparison of speech-language pathology graduate students to those in related disciplines. Contemporary Issues in Communication Science and Disorders. 1999; 26:119–127.

Justice LM, Ezell HK. Written language awareness in preschool children from low-income households: A descriptive analysis. Communication Disorders Quarterly. 2001; 22:123–134.

Justice LM, Ezell HK. Use of storybook reading to increase print awareness in at-risk children. American Journal of Speech-Language Pathology. 2002; 85:388–396.

Kerry SM, Bland JM. The intracluster correlation coefficient in cluster randomization. British Medical Journal. 1998; 316:1455. [PubMed: 9572764]

Kish, L. Survey sampling. New York, NY: Wiley; 1965.

Kott L. A model-based look at linear regression with survey data. The American Statistician. 1991; 45:107–112.

Landry SH, Swank PR, Smith KE, Assel MA, Gunnewig SB. Enhancing early literacy skills for preschool children: Bringing a professional development model to scale. Journal of Learning Disabilities. 2006; 39:306–324. [PubMed: 16895156]

Lee, ES.; Forthofer, RN.; Lorimor, RJ. Analyzing complex survey data. Newbury Park, CA: Sage Publications; 1989.

Little, RJA.; Rubin, DB. Statistical Analysis with Missing Data. 2nd edition. New York, NY: John Wiley & Sons, Inc.; 2002.

LoCasale-Crouch J, Konold T, Pianta R, Howes C, Burchinal M, Bryant D, Clifford R, Early D, Barbarin O. Profiles of observed classroom quality in state-funded pre-kindergarten programs and associations with teacher, program, and classroom characteristics. Early Childhood Research Quarterly. 2007; 22:3–17.

Lonigan, CJ. Emergent literacy skills and family literacy. In: Wasik, B., editor. Handbook of family literacy. Mahwah, NJ: Erlbaum; 2004. p. 57-82.

Lonigan CJ, Anthony JL, Bloomfield BG, Dyer SM, Samwel CS. Effects of two shared-reading interventions on emergent literacy skills of at-risk preschoolers. Journal of Early Intervention. 1999; 22:306–322.

Lonigan CJ, Burgess SR, Anthony JL, Barker TA. Development of phonological sensitivity in 2- to 5-year-old children. Journal of Educational Psychology. 1998; 90:294–311.

Lonigan, CJ.; McDowell, K.; Phillips, B. Assessment of children's pre-literacy skills. In: Wasik, B., editor. Handbook on family literacy: Research and services. Mahwah, NJ: Lawrence Erlbaum; 2004. p. 525-550.

Lonigan, CJ.; Wagner, RK.; Torgesen, JK. Test of Preschool Early Literacy (TOPEL). Austin, TX: PRO-ED; 2007.

Lonigan, CJ.; Wagner, RK.; Torgesen, JK.; Rashotte, C. Preschool Comprehensive Test of Phonological and Print Processing. Department of Psychology, Florida State University; 2002.

Lonigan C, Bloomfield B, Anthony J, Bacon K, Phillips B, Samwel C. Relations among emergent literacy skills, behavior problems, and social competence in preschool children from low- and middle-income backgrounds. Topics in Early Childhood Special Education. 1999; 19:40–53.

Mashburn AJ, Pianta RC, Hamre BK, Downer JT, Barbarin O, Bryant D, Burchinal M, et al. Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. Child Development. 2008; 79:732–749. [PubMed: 18489424]

Moats L. The missing foundation of teacher education: Knowledge of the structure of spoken and written language. Annals of Dyslexia. 1994; 44:81–102.

Morrison FJ, Connor CM. Understanding schooling effects on early literacy: A working research strategy. Journal of School Psychology. 2002; 40:493–500.

National Center of Education Statistics (NCES). National Education Longitudinal Study of 1988: User's manual. Washington, DC: U.S. Department of Education; 2002.

National Center for Education Evaluation (NCEE). National evaluation of Early Reading First: Final report to Congress. Washington, DC: U.S. Department of Education; 2007.

National Council on Teacher Quality. Increasing the odds: How good policies can yield better teachers. Washington, DC: Author; 2005.

National Early Literacy Panel (NELP. Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Center for Family Literacy; 2008.

National Institute of Child Health and Human Development, Early Child Care Research Network. The relation of child care to cognitive and language development. Child Development. 2000; 71:960–980. [PubMed: 11016559]

National Institute of Child Health and Human Development, Early Child Care Research Network. Child-care structure → Process → Outcome: Direct and indirect effects of child-care quality on young children's development. Psychological Science. 2002; 13:199–206. [PubMed: 12009038]

Neuman SB, Wright TS. Promoting language and literacy development for early childhood educators: A mixed-methods study of coursework and coaching. The Elementary School Journal. 2010; 111:63–86.

Notari-syverson, A.; O'Connor, RE.; Vadasy, PF. Ladders to literacy: A preschool activity book. Baltimore, MD: Paul H. Brookes; 1998.

Peisner-Feinberg ES, Burchinal MR. Relations between preschool children's child-care experiences and concurrent development: The Cost, Quality, and Outcomes Study. Merrill-Palmer Quarterly. 1997; 43:451–477.

Penno JF, Wilkinson AG, Moore DW. Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? Journal of Educational Psychology. 2002; 94:22–33.

Pew Charitable Trusts. National Early Childhood Accountability Task Force Briefing. Philadelphia, PA: Pew; 2005.

Pianta, RC. Standardized observation and professional development: A focus on individualized implementation and practices. In: Zaslow, M.; Martinez-Beck, I., editors. Critical issues in early childhood professional development. Baltimore, MD: Paul H. Brookes; 2005. p. 231-254.

Pianta, RC.; Howes, C.; Early, D.; Clifford, R.; Bryant, D.; Burchinal, M. Observations of quality and practices in pre-k classrooms: Associations with child outcomes and teacher attributes. Paper presented at the biennial Meeting of the Society for Research in Child Development; Tampa, FL. 2003 Apr.

Pianta, RC.; La Paro, KM.; Hamre, BK. Classroom assessment scoring system [CLASS]. Baltimore, MD: Paul H. Brookes; 2008.

Pianta R, Howes C, Burchinal M, Bryant D, Clifford R, Early D, Barbarin O. Features of pre-kindergarten programs, classrooms, and teaches: Do they predict observed classroom quality and child-teacher interactions? Applied Developmental Science. 2005; 9:144–159.

Pianta R, Mashburn A, Downer J, Hamre B, Justice L. Effects of web-mediated professional development resources on teacher-child interactions in pre-kindergarten classrooms. Early Childhood Research Quarterly. 2008; 23:431–451.

Powell DR, Diamond KE, Burchinal MR, Koehler MJ. Effects of an early literacy professional development intervention on Head Start teachers and children. Journal of Educational Psychology. 2010; 102:299–312.

Ramey, S.; Ramey, C. Creating and sustaining a high-quality workforce in child care, early intervention, and school readiness programs. In: Zaslow, M.; Martinez-Beck, I., editors. Critical issues in early childhood professional development. Baltimore, MD: Paul H. Brookes; 2005. p. 355-368.

Raudenbush, S.; Bryk, A. Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage Publications; 2002.

Raver CC, Jones AS, Li-Grining CP, Metzger M, Smallwood K, Sardin L. Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in head Start settings. Early Childhood Research Quarterly. 2008; 23:10–26.

Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association. 1996; 91:473–489.

Sailors M, Shanklin NL. Introduction: Growing evidence to support coaching in literacy and mathematics. The Elementary School Journal. 2010; 111:1–6.

Schatschneider C, Fletcher JM, Francis DJ, Carlson C, Foorman BR. Kindergarten prediction of reading skills: A longitudinal comparative analysis. Journal of Educational Psychology. 2004; 96:265–282.

Snow, CE.; Hamphill, L.; Barnes, WS., editors. Unfulfilled expectations: Home and school influences on literacy. Cambridge, MA: Harvard University Press; 1991.

Snowling MJ, Gallagher A, Frith U. Family risk of dyslexia is continuous: individual differences in the precursors of reading skill. Child Development. 2003; 74:358–373. [PubMed: 12705560]

Steiger, JH. EzPATH: Causal modeling. Evanston, IL: SYSTAT Inc.; 1989.

Storch, SA.; Whitehurst, GJ. The role of family and home in the literacy development of children from low-income backgrounds. In: Britto, PR.; Brooks-Gunn, J., editors. The role of family literacy environments in promoting young children's emerging literacy skills: New directions for child and adolescent development. New York, NY: Jossey-Bass; 2001. p. 53071

Storch S, Whitehurst G. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. Developmental Psychology. 2002; 38:934–947. [PubMed: 12428705]

Thompson B. Guidelines for authors. Educational and Psychological Measurement. 1994; 54:837–847.

Torgesen JK. Catch them before they fall: Identification and assessment to prevent reading failure in young children. American Educator, Spring/Summer. 1998:1–8.

U.S. Department of Education. America's Kindergarteners: Findings from the Early Childhood Longitudinal Study, Kindergarten Class of 1998–99. Washington, DC: National Center for Education Statistics; 2000.

Ukrainetz TA, Cooney MH, Dyer SK, Kysar AJ, Harris TJ. An investigation into teaching phonemic awareness through shared reading and writing. Early Childhood Research Quarterly. 2000; 15:331–355.

van Kleeck A, Gillam RB, McFadden TU. A study of classroom-based phonological awareness training for preschoolers with speech and/or language disorders. American Journal of Speech-Language Pathology. 1998; 7:65–76.

Vaughn, S.; Linan-Thompson, S.; Pollard-Durodola, SD.; Mathes, PG.; Cardena Hagan, E. Effective interventions for English Language Learners (Spanish-English) at risk for reading difficulties. In: Dickinson, DK.; Neuman, SB., editors. Handbook of early literacy research. New York, NY: The Guilford Press; 2006. p. 185-197.

Vernon-Feagans, L. Children's talk in communities and classrooms. Cambridge, MA: Blackwell Publishers; 1996.

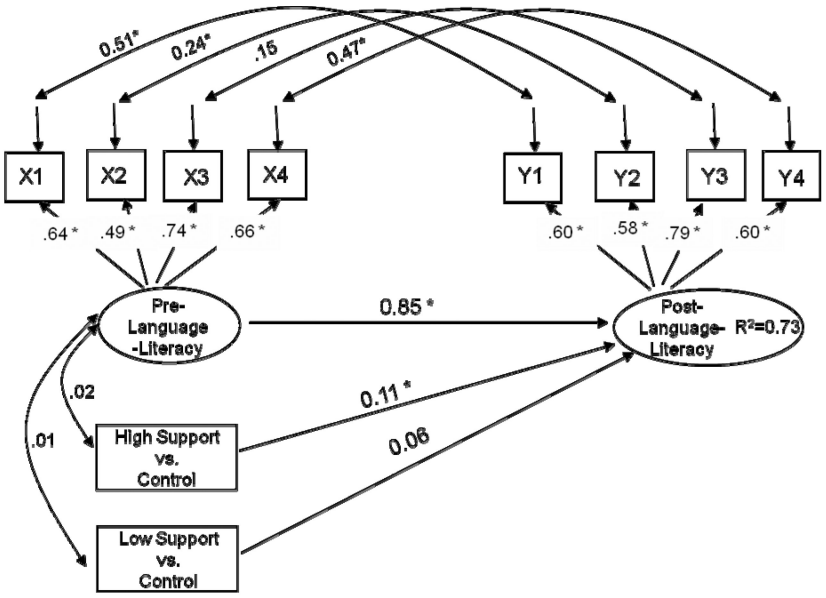
Vogel, CA.; Nogales, R.; Aikens, N.; Tarullo, L. Implementation of the Head Start National Reporting System: Spring 2006. Princeton, NJ: Mathematica Policy Research, Inc.; 2008.

Walpole S, McKenna MC, Uribe-Zarain X, Lamitina D. The relationships between coaching and instruction in the primary grades: Evidence from high-poverty schools. The Elementary School Journal. 2010; 111:115–140.

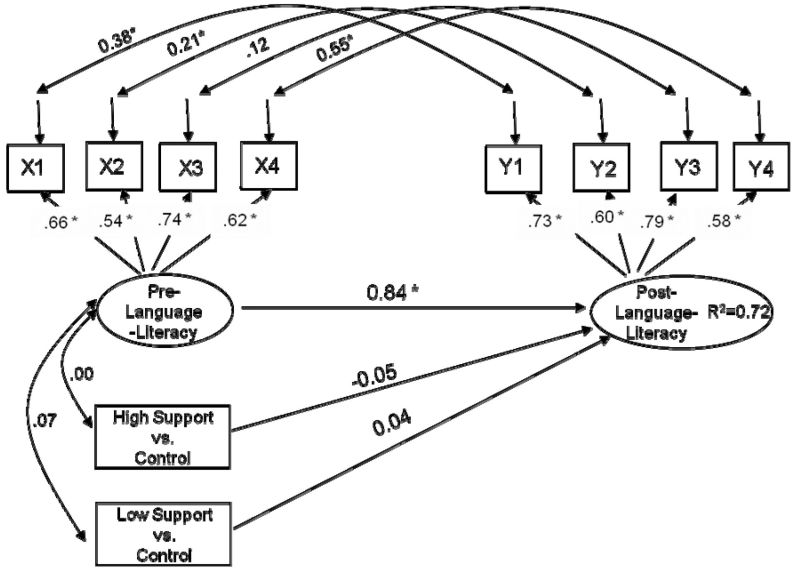
Wasik BA, Bond MA. Beyond the pages of a book: Interactive reading and language development in preschool classrooms. Journal of Educational Psychology. 2001; 93:243–250.

Whitehurst GJ, Epstein AL, Angell AC, Crone DA, Fischel JE. Outcomes of an emergent literacy intervention in Head Start. Journal of Educational Psychology. 1994; 86:542–555.

Zaslow, M.; Martinez-Beck, I., editors. Critical issues in early childhood professional development. Baltimore, MD: John H. Brookes; 2005.



(a) Only English Spoken by Children in Class (N = 702)



(b) Spanish or Other Languages Spoken by Children in Class (N = 636)

Figure 1. Modeling intervention effect for two moderator conditions

*Statistically significant at α=.05 after correcting the clustering sampling design effect. X1 – X4: Pre-Language/Literacy Measures: Vocabulary, Blending, Elision, Print. Y1 – Y4: Post-Language/Literacy Measures: Vocabulary, Blending, Elision, Print

Table 1

Classroom, Teacher, Child, and Family Characteristics by Moderator Variable Groups (n=1338 children/161 classrooms [a])

	English [b]		Other		English		Other	
	n	%	*n*	%	*M*	*SD*	*M*	*SD*
Classroom Characteristics								
Proportion Poor					0.63	0.27	0.77	0.26
Mean Maternal Education (years)					12.77	0.81	12.48	1.14
Teacher Characteristics								
Level of Education								
Bachelor's Degree	68	69	30	52				
Advanced Degree	29	30	28	48				
Field of Study								
Early Childhood Education	39	40	20	35				
Other	59	60	37	65				
Years Teaching					15.32	9.33	15.79	9.01
Child and Family Characteristics								
Gender								
Boy	354	51	291	47				
Girl	337	49	322	53				
Language(s) spoken by child at home [c]								
English	692	99	475	77				
Spanish	11	2	165	17				
Other languages	6	1	100	16				
Maternal education (years)					12.77	1.78	12.72	2.35
Family income [d]								
Poverty	406	65	402	74				
Not Poverty	221	35	141	26				

[a] Variables have different amount of missing values; as a result, the numbers may not add up to these totals.

[b] English: Classrooms where only English was spoken by students (*N = 84*); Other: Classrooms where Spanish/other language were spoken by students (*N = 75*).

[c]Because some children spoke another language at home in addition to English, the numbers under this category add up to be more than the total *n* of 1338.

[d]This variable has the largest amount of missing values (12.5%).

Table 2

Descriptive Statistics of Outcome Measures for Intervention Groups under Moderator Conditions

Moderator	Intervention Groups		Pre-Language/Literacy				Post-Language/Literacy			
			Vocabulary	Blending	Elision	Print	Vocabulary	Blending	Elision	Print
Only English Spoken in Class by Students	High Support	*M*	14.99	6.85	4.81	16.07	18.67	8.66	7.52	29.15
	n = 215	*SD*	3.78	2.95	3.14	8.81	2.96	2.68	3.06	7.30
	Low Support	*M*	14.89	6.25	4.59	17.40	18.16	8.23	7.48	29.22
	n = 348	*SD*	3.68	2.98	3.16	8.82	3.10	2.57	3.31	6.90
	Control	*M*	15.34	5.98	4.23	16.15	18.46	7.84	7.00	26.89
	n = 139	*SD*	4.28	3.21	3.20	9.47	3.08	2.79	3.40	8.00
Spanish, Other Language Spoken in Class by Students	High Support	*M*	13.31	6.05	4.00	16.62	17.22	7.87	6.82	27.32
	n = 299	*SD*	4.09	3.06	3.07	9.02	3.55	3.00	3.33	8.11
	Low Support	*M*	14.50	6.29	4.38	15.62	18.44	8.16	7.37	29.24
	n = 162	*SD*	3.96	2.66	3.09	7.96	3.26	2.74	3.34	6.66
	Control	*M*	13.21	5.90	3.87	15.29	17.71	7.72	6.72	27.23
	n = 175	*SD*	3.99	3.36	3.08	9.02	3.34	3.11	3.44	7.75

Note. The descriptive information in this table is based on the average of five imputations for missing data.

Table 3

Model Fit Assessment for Two Moderator Conditions

	Two Moderator Conditions	
Model Fit Indices	**Only *English* Spoken by Children in Class (*N* = 702)**	***Spanish or Other Languages* Spoken by Children in Class (*N* = 636)**
$\chi^2_{(df=27)}$	69 ($p < 0.01$)	65 ($p < 0.01$)
χ^2/df ratio	*2.55*	*2.41*
GFI	*0.98*	*0.98*
AGFI	*0.96*	*0.96*
RMSEA	0.05 CI_{95} = (0.03, 0.06) *p* (for close fit) = .60	0.05 CI_{95} = (0.03, 0.06) *p* (for close fit) = .58
CFI	*0.98*	*0.98*
TLI	*0.97*	*0.97*
NFI	*0.97*	*0.97*