Early Childhood Research Quarterly 54 (2021) 136–152

Contents lists available at ScienceDirect

Early Childhood Research Quarterly

Teaching Together: Pilot study of a tiered language and literacy intervention with Head Start teachers and linguistically diverse families

Tricia A. Zucker (Ph.D.)[a,*], Sonia Q. Cabell (Ph.D.)[b], Yaacov Petscher (Ph.D.)[b], Heather Mui (MPH)[a], Susan H. Landry (Ph.D.)[a], Jamie Tock (Ph.D.)[b]

[a] Children's Learning Institute at the University of Texas Health Science Center at Houston, Department of Pediatrics, United States
[b] Florida Center for Reading Research, Florida State University, United States

ARTICLE INFO

Article history:
Received 26 March 2019
Received in revised form 4 August 2020
Accepted 1 September 2020

1. Introduction

This study examines blending of classroom- and family-based supports within a Multi-Tiered System of Support (MTSS). MTSS frameworks typically use primary, secondary, and tertiary interventions to prevent or ameliorate academic difficulties. Preventing reading difficulties should begin in early childhood education (ECE) programs using evidence-based Tier 1 instruction with all children, Tier 2 targeted instruction for children with weak skills, and specialized Tier 3 interventions for children with special needs (Carta & Miller-Young, 2019). The present study focuses on a sample of Head Start children who are both experiencing poverty and scoring below early language and literacy screening benchmarks. MTSS studies in pre-kindergarten (pre-k) classrooms are feasible and promising (e.g., Buysse et al., 2016), yet more information is needed on the intensity of supports children experiencing poverty and children from language minority backgrounds require. These profiles of learners are at elevated risk for longer-term language and literacy difficulties (e.g., Chatterji, 2006; Hoff, 2013; Scheffner Hammer et al., 2014). This study begins to explore which combination of resources are the most parsimonious, yet effective for improving language and literacy skills for such learners.

We examine a model that first establishes Tier 1 and 2 supports in the classroom and then adds two aligned tiers of family engagement supports called *Teaching Together*. Poverty often leads to limited resources and fewer opportunities for rich language and literacy interactions (Heidlage et al., 2019; Hoff, 2013; Wasik & Hindman, 2015). Therefore, we drew on theoretical frameworks that recognized school and home are key microsystems for learning (Bronfenbrenner & Morris, 2006) when responsive adults guide and scaffold learning (Vygotsky, 1978). We posited aligned interventions in both school and home are required to close the achievement gap between preschoolers experiencing poverty and their peers.

The *Teaching Together* intervention primarily addressed early language skills alongside a secondary focus on literacy skills, as these are precursors to later reading comprehension and decoding, respectively (Dickinson, Golinkoff, & Hirsh-Pasek, 2010; van Kleeck, 2008). Focal skills included three related academic language areas: (a) understanding sophisticated vocabulary; (b) using inferential language to articulate ideas beyond the immediate context, such as inferences about emotions, cognition, or reasoning about causality; and (c) narrative abilities for articulating a logical series of events or other forms of extended discourse (Foorman et al., 2016). Focal literacy skills included: (d) print knowledge and understanding how print represents spoken language, and (e) phonological awareness to identify and manipulate sounds in spoken language (National Early Literacy Panel [NELP], 2008; Whitehurst & Lonigan, 1998). Both teacher and parent interventions uses multiple adult learning approaches – sharing information, modeling, prompting, providing feedback, and guided reflection (cf. Haring Biel et al., 2019).

1.1. Effective Tier 1 and Tier 2 approaches in ECE classrooms

Our classroom intervention was built upon extensive research demonstrating the effectiveness of interactive shared reading to support language and literacy skills (Mol, Bus, & de Jong, 2009;

* Corresponding author at: 7000 Fannin St., 2300, Houston, TX 77030, United States.
E-mail addresses: Tricia.Zucker@uth.tmc.edu (T.A. Zucker), scabell@fcrr.org (S.Q. Cabell).

https://doi.org/10.1016/j.ecresq.2020.09.001

NELP, 2008), including for children at risk for reading difficulty (Swanson et al., 2011). Four times per week, all children receive evidence-based Tier 1 shared reading designed as 20-min lessons. Three of these lessons taught academic vocabulary (e.g., Neuman & Dwyer, 2011; Penno, Wilkinson, & Moore, 2002) and supported inferential level conversations (e.g., Blewitt, Rump, Shealy, & Cook, 2009; van Kleeck, Vander Woude, & Hammett, 2006), but one lesson embedded explicit references to print within the text (Justice, Kaderavek, Fan, Sofka, & Hunt, 2009; Justice, Logan, Kaderavek, & Dynia, 2015). Approximately four children per classroom receive four 10-min, Tier 2 small-group lessons per week that review and extend understanding of academic language and vocabulary (e.g., Pullen, Tuckwiller, Konold, Maynard, & Coyne, 2010; Zucker, Solari, Landry, & Swank, 2013) or used more hands-on approaches to increase print knowledge (e.g., Koutsoftas, Harmon, & Gray, 2009; Kruse, Spencer, Olszewski, & Goldstein, 2015). The classroom teacher delivers these Tier 1 and 2 lessons. This extends a tiered language intervention that has shown promise in our past studies by adding literacy instruction (Zucker et al., 2019).

Classroom Tier 1 rationale. Experimental evidence shows Tier 1 shared reading curricular supplements can effectively support broad oral language skills (e.g., Lonigan, Allan, & Lerner, 2011; Wasik & Hindman, 2018) including, academic vocabulary knowledge (Pollard-Durodola et al., 2011). Compared to past studies evaluating a Tier 1 print referencing style of shared reading with two to four sessions per week, this study used a light dose of one session per week (Justice et al., 2009, 2015; Justice, McGinty, Piasta, Kaderavek, & Fan, 2010). Although the present study focuses exclusively on children struggling with language and literacy and eligible for Tier 2 supports, we decided to enhance Tier 1 instruction because the MTSS presumption of high-quality evidence-based Tier 1 practices is often unmet in preschool (Lonigan & Phillips, 2016).

Classroom Tier 2 rationale. Tier 2 and small-group instructional settings provide additional opportunities for extended discourse, allow teachers to better scaffold instruction, and promote review of concepts (Connor, Morrison, & Slominski, 2006; Coyne, Kame'enui, Simmons, & Harn, 2004). An increasing number of early childhood studies show positive effects of Tier 2 literacy instruction focused on alphabet skills and phonological awareness (e.g., Coyne et al., 2004; Koutsoftas et al., 2009; Kruse et al., 2015). A small number of ECE studies show promise by providing Tier 2 interventions that reinforce and extend understanding of vocabulary introduced in Tier 1 (Loftus, Coyne, McCoach, Zipoli, & Pullen, 2010; Pullen et al., 2010; Zucker et al., 2013). Successful pre-k Tier 2 approaches also embed inferential level questions within book reading to elicit causal reasoning and inferences that go beyond what is explicitly stated in the text (van den Broek, Kendeou, Lousberg, & Visser, 2017; van Kleeck et al., 2006; Zucker et al., 2013).

MTSS approaches provide Tier 2 supports, in part, because many high-quality Tier 1 pre-k language and literacy interventions provide less benefit to children with weaker initial skills relative to their peers (e.g., Cabell, Justice, Piasta, et al., 2011; Penno et al., 2002; Reese & Cox, 1999). Thus, screening all children at the beginning of the school year to identify children that need more intensive Tier 2 instruction is recommended practice (Gersten et al., 2008). However, even amongst students who initially qualify for Tier 2, not all profit equally from this support. For example, recent research with elementary students with or at risk for reading difficulties demonstrate heterogeneity of response even within these samples, such that children respond differentially depending on their end of year skill level (e.g., Solari, Denton, Petscher, & Haring, 2018; Wanzek et al., 2016). Therefore, we explored if the *Teaching Together* program showed differential promise for children based on initial skills and/or end of year skills.

1.2. Effective approaches in ECE family interventions

Many ECE models seek to increase parent's roles in their child's education. For example, the federally-funded Head Start program requires centers to provide family education activities to support children's learning at home (US Department of Health and Human Services, 2016). Efforts to increase parents' involvement from preschool to Grade 12 generally show small, positive academic effect sizes (ES; Jeynes, 2012; Wilder, 2014). Yet a 2016 meta-analysis with 46 studies of ECE programs (Grindal et al., 2016) found no benefit of adding family education programs on children's outcomes (ES = .01 for vocabulary, literacy) unless it included modeling and parent skill practice (ES range .08–.22) or parent coaching (ES range .30–.42). These effect sizes align with other meta-analyses of programs that specifically coach parents to read interactively with their child (ES range .51–.65; Mol, Bus, de Jong, & Smeets, 2008; Sénéchal & Young, 2008). More recent studies that combine classroom intervention and parent coaching report diverse child impacts that range from mixed literacy effects (d = –.33 to .14) and null language effects (Landry et al., 2017) to sustained effects into Grade 2 (reading g = .30–.33; Bierman, Heinrichs, Welsh, Nix, & Gest, 2017). These inconsistent effects call into question long-standing beliefs that there is value-added when ECE programs provide parent education.

This study considers how to best invest resources for supporting outcomes of children living in poverty using the *Teaching Together* tiered family engagement approach. Some existing parent intervention models provide tiers of support such that low cost approaches (self-study and group education) are offered universally, whereas individualized parent coaching is reserved for children demonstrating increased risk (McIntyre & Phaneuf, 2008; Sanders, Kirby, Tellegen, & Day, 2014). The present study applied a tiered service model to ECE family services that included both "parent involvement" and "family engagement" activities (White, 2019). Parent involvement encourages parents to support their child's learning and participate in school communications and events. Family engagement is a process in which educators and families develop a relationship focused on supporting children's success via shared perspectives. All *Teaching Together* activities used strength-based and culturally-affirming approaches that can give power to often marginalized parents to serve as their child's first and most important teacher (National Academies of Sciences, Engineering, and Medicine [NASEM], 2017).

Family Tier 1 rationale. As part of *Teaching Together*, all families receive information to promote positive parent–child interactions and extend the classroom learning at the home. First, group-based, after-school workshops that were adapted from a library- and museum-based model (Garibay, 2007) that: (a) uses videotaped models to explain positive parent strategies, (b) models interactive shared book reading, and (c) includes opportunities for strategy practice and feedback at family activity stations. Second, parent communication includes text messages with developmental information and tips for parents to support learning at home (e.g., Cabell, Zucker, DeCoster, Copp, & Landry, 2019; York, Loeb, & Doss, 2018). All family engagement activities are designed to be facilitated by early childhood educators (teacher, social worker) or community partners (librarians, museum educators). However, in this study, parent activities were researcher-led to examine promise of the approach for potential later use in ways that are school-led and facilitate true parent-teacher partnerships.

Family Tier 2 Basic rationale. To dismantle the effects of providing basic learning resources versus a more intensive parent coaching approach, the first randomly assigned Tier 2 condition simply provided additional resources. Many effective ECE approaches feature provision of material resources for extending learning at home in ways that closely aligned with the classroom

curriculum (e.g., Bierman et al., 2017; (Roberts, 2008). For example, preschool (Anthony, Williams, Zhang, Landry, & Dunkelberger, 2014) and kindergarten programs (Jordan, Snow, & Porche, 2000) that provided five group, afterschool family sessions and materials to extend learning at home resulted in small to medium effects on vocabulary outcomes (ES range .15–.64), but smaller effects on literacy (ES range .07–.32).

Family Tier 2 Enhanced rationale. The second randomly assigned Tier 2 condition added parent coaching to other family supports. Parent coaching is a well-established intervention approach for families experiencing poverty and linguistically diverse families (Heidlage et al., 2019; NASEM, 2017). Coaching increases positive parent behaviors and a range of child outcomes, including language and literacy (e.g., Brown & Lee, 2017; Dulay, Cheung, Reyes, & McBride, 2019; Landry et al., 2012). However, parent coaching is resource intensive, requiring highly trained coaches and individualized supports that, per session, cost about $200 per child (Knight et al., 2016) and can impact scalability (Mahoney, McConnell, Larson, Becklenberg, & Stapel-Wax, 2020). The present study modified the Tier 3 type of coaching used by this team (Landry et al., 2017) to a less intensive Tier 2 approach with four remote coaching sessions. Recent parenting interventions, such as Thirty Million Words (Suskind et al., 2016) or Providence Talks (Wong, Boben & Thomas, 2018), include two to eight months of parent coaching visits. Yet, some effective models use less intensive coaching (Bakermans-Kranenburg, Van Ijzendoorn, & Juffer, 2003), such as the Video Interaction Project (Mendelsohn et al., 2011) to reduce costs and improve scalability.

1.3. Considerations for dual language learners (DLLs)

Early intervention within MTSS has the potential to support certain profiles of children such as DLLs who often face broad risk for later reading disabilities (Catts, Petscher, Schatschneider, Sittner Bridges, & Mendoza, 2009). Young DLLs who are experiencing poverty may not have had the early experiences to optimally prepare for learning in pre-k programs (Ballantyne, Sanderman, & McLaughlin, 2008; Hoff, 2013). More tailored MTSS approaches could better support DLLs' unique pre-k learning needs. Linguistically diverse families of children entering preschool need to understand the value of maintaining their child's home language to maintain cultural and linguistic heritage and provide a strong foundation to learn academic language (NASEM, 2017). The present intervention encouraged parents to maintain use of the home language to reduce unintended first language loss that can occur when children experience ECE in English. Parents received information on how informal learning activities in the home language support school readiness (e.g., Long, 2012; Roberts, 2008). All family materials were available in English and Spanish, whereas classroom instruction focused on English.

The present study included a relatively large proportion of DLLs (about 66%) who unexpectedly spoke diverse home languages beyond English and Spanish. Although early bilingualism is an asset, this linguistic diversity presents challenges for educators to meet the unique needs of DLLs. ECE teachers need PD on classroom practices that support DLLs' language and literacy skills such as visual supports, explicit explanation of vocabulary, and selective use of the home language (Buysse, Peisner-Feinberg, Páez, Hammer, & Knowles, 2014; Fitton, McIlraith, & Wood, 2018). The classroom curriculum in the present study included multiple supports for DLLs such as: (a) accompanying new word learning with corresponding gestures to imagine they were experiencing new words (cf. Adams, Glenberg, & Restrepo, 2018), (b) scaffolding interactive conversations about texts read aloud (Fitton et al., 2018), and (c) improving narrative language skills with story sequencing and retelling activities that included visuals and props (Rogde, Melby-Lervåg, & Lervåg, 2016).

1.4. Study purpose & context

The supplemental *Teaching Together* intervention was iteratively developed using multiple earlier phases to ensure the curriculum was feasible and appealing for users. Thus, this pilot study used an underpowered, randomized controlled trial to evaluate intervention promise for child outcomes. We screened children to focus only on a sample of children eligible for Tier 2 based on state screening measures. Although in authentic MTSS models children who receive Tier 1 versus Tier 2 represent different skill levels, we sought to experimentally contrast the effects of enriched Tier 1 only or added Tier 2 supports for a sample of preschoolers at higher risk due to Tier 2 eligibility. We designed the study to examine the effects of increasingly intensive teacher and parent supports such that children could be randomly assigned to Tier 1 Only or increasingly intensive interventions including Tier 2 Basic or Tier 2 Enhanced family supports. The components of the three school plus home interventions are detailed below; all treatment conditions included an aligned parent component rather than a school only condition.

Our primary research question asked: What are the effects of the increasingly tailored tiers of school plus home interventions (Tier 1 Only, Tier 2 Basic, Tier 2 Enhanced) contrasted with business-as-usual (BAU) practice on children's language skills and literacy skills? Second, we asked for which children the interventions were more or less beneficial depending on children's: (a) DLL status; (b) initial skill levels; and (c) end of year skill level, examining varying points of the conditional distribution of the outcome. We hypothesized that the intervention program would result in improved outcomes relative to the other children receiving business-as-usual instruction. We also expected that each increasing tier and intensity of school and home supports would be associated with modest gains relative to BAU practice. We did not have directional expectations around the second research question but felt such considerations were important in exploring who might benefit from this new program.

2. Method

This study occurred in a large, metropolitan area in the south central region of the United States. It included 33 classrooms from three Head Start agencies (20 school sites). Pre-k classrooms were eligible to participate if: (a) instruction was predominantly in English, and (b) the teacher had not participated in PD projects offered by the university researchers in the past five years. We invited 68 classrooms and screened them for eligibility; 24 teachers declined and 11 teachers were not eligible due to participation in recent research studies. Thus, a total of 33 classrooms were eligible. We requested written teacher and parental consents for research activities but also used a passive consent process, approved by our local Institutional Review Board (IRB), to allow all children to participate in screening tasks.

There were four key steps in the study design that are illustrated in Fig. 1. First, researchers screened all children in participating classrooms to identify a high-risk sample. We screened 426 children and only 45 children (10.56%) were *not* eligible because they scored above the screening criteria. Amongst consented, Tier 2 eligible children, 170 children were allocated in the study design using random assignment at classroom, child, and family levels. Second, the methodologist randomized classroom clusters to either: (a) the *Teaching Together* classroom/family intervention ($n = 17$ classrooms, 98 focal children); or (b) a waitlist control

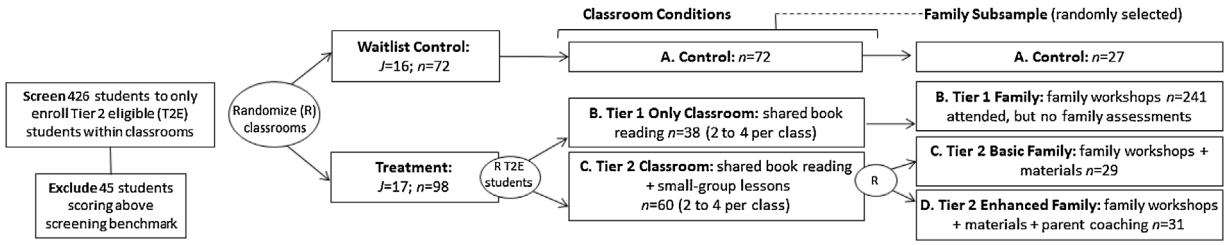

Fig. 1. Study design.

Table 1
Intervention scope & sequence with additive supports by treatment (Tx) conditions.

	Tier 1 Condition	Tier 2 Basic Condition	Tier 2 Enhanced Condition
Classroom supports			
Whole-Group Lessons	Interactive Shared Reading: 90 lessons, 20 minutes each • Day 1: Language comprehension • Day 2: Vocabulary and comprehension • Day 3: Literacy, print referencing • Day 4: Text review and retelling		
Small-Group Lessons	—	+ Review & Extension Activities: 76 lessons, 10 min each• Day 1: Review and build background knowledge • Day 2: Vocabulary review and extension activity • Day 3: Literacy – PA, print and letter knowledge • Day 4: Text review and retelling	
Family Supports			
Workshops	Six Workshops & Museum Passes		
Text Messages	35 Messages• Mondays: Conversation starters linked to shared reading • Thursdays: Activity videos	+34 Messages• Tuesdays: Tips and strategies for supporting language • Wednesdays: Tips and strategies for supporting literacy	
Materials	—	+ Family Materials Kit • Four books with vocabulary and question prompts 11 language activities, four literacy activities	
Coaching	—	—	+ Four Parent Coaching Sessions • Four informational videos and linked activity manual • Parent–child video upload• Coach call to discuss topic of week and reflect on parent–child interactions

that implemented business-as-usual classroom/family programs and received delayed training/materials in the summer ($n = 16$ classrooms, 72 focal children). Consented children scoring below screening benchmarks were rank-ordered within each classroom and the lowest-performing four to eight children per classroom were selected (based on the number of eligible children per classroom). Third, pairs of lowest rank-ordered children were randomly assigned within treatment classrooms, to Tier 1 Only Classroom instruction ($n = 38$), or extra Tier 2 Classroom small-group instruction ($n = 60$). That is, the Child 1 and 2 pair were randomized to Tier 1 or Tier 2 within their classroom, followed by paired Child 3 and 4, and so on. Fourth, researchers further randomized families of children assigned to Tier 2 classroom intervention (see Fig. 1). There were four final family groups: (a) control or business-as-usual (BAU) family engagement experiences; (b) Tier 1 Family workshops and universal text messages; (c) Tier 2 Basic Family, that added learning materials and targeted text messages to the workshops; or (d) Tier 2 Enhanced family that added parent coaching/individualized communication to the above supports. Classroom interventions began before the final fourth step of randomization to Tier 2 family treatment conditions; thus, family conditions (a) and (b) were created default with the earlier randomizations, whereas group (c) and (d) were created with an additional round of randomization. There was no difference between the classroom experience for Tier 2 Basic and Tier 2 Enhanced groups; rather, these conditions test added levels of family support while holding the classroom experience constant, as illustrated in Table 1. As shown in Fig. 1, amongst the 241 families who participated in the universal Tier 1 activities, a subsample was selected that included 87 focal families who participated in an additional battery of home assessments; these families were selected randomly from the control group ($n = 27$) and represented all of the Tier 2 Basic ($n = 29$) and the Tier 2 Enhanced families ($n = 31$).

2.1. Selection of child participants

In October and early November 2017, researchers administered a screener to 426 children using two fluency subtests – rapid

vocabulary and letter naming – of the CIRCLE Assessment and the test-developers' age-based screening benchmarks (Landry et al., 2014). This represented 78.45% of the 543 children enrolled in these classrooms, but 100% of children present on assessment days. Children were eligible for the study if they: (a) demonstrated risk for later reading difficulties based on scores below predetermined benchmarks on a screening measure (rapid vocabulary naming and/or letter naming), (b) were between 3.5 years and 5.0 years on October 1 of the intervention year, and (c) did not have a significant sensory or intellectual disability (although children with mild disabilities such as language impairment were eligible). No children were excluded due to severe disabilities; however 10 enrolled children had an Individualized Education Plan (IEP) for mild disabilities such as speech/language difficulties or epilepsy. About 77.3% of parents provided consent for eligible children to participate.

2.2. Classrooms & teacher participants

Classrooms served an average of 18.09 children and most were staffed by a lead teacher and an assistant (teacher–child ratio of 1:9). Most children were enrolled in full-day programs, but four classrooms were half-day. Head Start policy at the time of the study encouraged the use of children's home languages, allowing for language of instruction to be decided based on staff skills and children's languages (U.S. Department of Health and Human Services, 2010). Despite initial recruitment criteria, there was diversity in language of instruction reported by teachers with 27 classrooms (51.52%) that provided >90% English instruction, three used 80% English, and three used about 50% English. Spanish was the most common secondary language of instruction, but Arabic was used in one classroom. The lead teacher in participating classrooms was invited to take part in the study. Table 2 shows demographic characteristics of teachers. All but two teachers had an Associate's degree or higher and 41.18% held a state certification for teaching pre-k. Most teachers were female. Teachers' reported race/ethnicities were: 19 African American, 10 Hispanic, 4 Asian, and 1 Caucasian ($n = 2$ missing). As detailed in a participant flowchart in Online Supplemental Fig. A1 and Table A1, there was a 26.92% attrition rate for teachers, with a differential attrition rate of 7.7% when comparing treatment and control groups.

2.3. Child & family participants

All participating children ($n = 170$) and their families met federal income or risk criteria to qualify for the Head Start program. Average child age at the start of the study was 4 years, 5 months (range = 3 years, 7 months to 5 years, 1 month). Child gender was about half male. As detailed in Table 3, children were racially and ethnically diverse – mostly Hispanic/Hispanic White or Black/African American. Parents reported diverse home languages spoken to these children: 30.00% spoke English only, 18.82% spoke Spanish only, 34.71% bilingual English and Spanish, 11.18% bilingual in English and another home language besides Spanish, and 1.76% only spoke another language at home (beyond English/Spanish). Children were considered DLLs if parents reported any language other than or in addition to English was spoken at home. Most primary caregivers indicated some college but no degree or high school as their highest level of education. The majority of families who reported their household income (75.88% of sample) earned at or below $11,000. As detailed in Online Supplemental Fig. A1 and Table A1, the overall attrition rate for child participants was 4.09%. There was less child- than classroom-level attrition because most teachers withdrew early and before the pretesting period.

Focal family subsample. The subsample of 87 focal families participated in an additional battery of home assessments that are reported elsewhere (Zucker et al., 2019). Due to limited resources, this subsample (shown in Fig. 1) represents all parent–child dyads in the Tier 2 Basic and Tier 2 Enhanced groups and a random selection of about 37% of the control group. None of the Tier 1 families completed home assessments. Table 3 provides details on background characteristics of this subsample. The focal, primary caregivers who participated were 93.02% mothers, 4.65% fathers, and 2.33% grandparents/great-grandparents. Similar to the full sample, families reported diverse home languages: 27.59% spoke English only, 18.39% spoke Spanish only, 39.08% bilingual English and Spanish, and 14.94% bilingual in English and another home language (i.e., Oromo, Arabic, Urdu, Kachin, Amharic, Ibo, Twi, Creole). Almost half (47.13%) of these caregivers reported a high school as their highest educational award. Most focal families reported annual household income at or below $30,000.

2.4. Data collection

This study occurred in the 2017–18 school year with screening in October through mid-November. In November and early December, children completed pretest language and literacy assessments. Posttest occurred in mid-April to early May 2018. Due to limited pilot study resources, one child measure (narrative) was administered only at post-test. Trained data collectors administered child assessments in a quiet location within the school using a counterbalanced order. Only 40% of assessors were blind to study condition due to limited resources; for example, some assessors were also family coaches, but they were not permitted to assess in schools where they had coached. Assessment staff were required to demonstrate >95% accuracy in administration and scoring procedures before testing at pretest and again at post-test.

2.5. Child proximal measures

Inferential comprehension. To estimate children's ability to understand complex language, children listened to a passage and answered questions that required inferencing. The *Inference Making Task* (Cain & Oakhill, 1999; Oakhill & Cain, 2012) assesses children's ability to make text integration inferences (requires connection between two separate premises in the text) and gap filling inferences (require integration of background knowledge to fill in gaps in the text). Children listened to a practice story, followed by a test story about a surprise party, and then answered eight questions. Responses were scored as incorrect (0), partially correct (1), of fully correct (2). In past studies, the pre-k version demonstrated good test–retest reliability scores ($r = .63$) and concurrent validity scores with other listening comprehension measures ($r = .61–.72$; Language and Reading Research Consortium [LAARC] & Muijselaar, 2018).

Taught vocabulary. Children's understanding of words within intervention texts was assessed with a researcher-developed measure called the *Teaching Together* Taught Vocabulary test (TT Taught). After two practice items with feedback, the computerized task names a word and asks the child to point to the best photo representing the word amongst two foils. For example, the correct picture for the word "conversation" shows two adults talking, whereas the foils show two adults watching television silently or two adults working in a garden but not talking. The screening/pretest version included 10 items to reduce testing time amongst the full sample of 426 children, but the post-test version added 18 additional taught words items. Internal consistency of the measure was poor for this sample at pretest (Cronbach's alpha = .36) but the reliability score improved somewhat (alpha = .56) when we added the posttest only items. Words assessed with the TT Taught measure included sophisticated, academic vocabulary, such as *shel-*

Table 2
Teacher demographics ($N = 34$).

	Control (BAU)	Intervention (TT)	Total	
			n	%
Race/ethnicity				
Black/African American	9	10	19	55.88
Asian	2	2	4	11.76
Hispanic	3	2	5	14.71
Hispanic White	2	3	5	14.71
White/Caucasian	0	1	1	2.94
Highest level of education				
High school diploma or equivalent	0	1	1	2.94
Some college but no degree	0	1	1	2.94
AA/AS2 year degree	2	6	8	23.53
Bachelor's degree	8	7	15	44.12
At least one year of course work beyond Bachelor's degree	2	1	3	8.82
Master's Degree	4	2	6	17.65
Doctoral Degree	0	0	0	0.00
Years of experience with pre-k				
0 years	0	1	1	2.94
1 year	0	0	0	0.00
2 years	1	1	2	5.88
3–4 years	4	3	7	20.59
5–10 years	5	6	11	32.35
11+years	5	7	12	35.29
No response	1	0	1	2.94
State certification for teaching pre-k				
Yes	5	9	14	41.18
No	11	9	20	58.82

Note. BAU = Business as Usual control condition. TT = Teaching Together intervention condition.

Table 3
Child demographics.

	Full sample ($N = 170$)				Family subsample ($N = 87$)				
	BAU	TT	*n*	%	BAU	TT Basic	TT Enh[a]	*n*	%
Gender									
Female	39	45	84	49.41	15	13	11	39	44.83
Male	33	53	86	50.59	12	16	20	48	55.17
Race/Ethnicity									
Hispanic/Hispanic White	31	57	88	51.76	9	15	19	43	49.43
Black/African American	27	27	54	31.76	13	8	9	30	34.48
Mixed	5	5	10	5.88	2	2	2	6	6.90
Asian	2	4	6	3.53	1	3	0	4	4.60
White/Caucasian	1	3	4	2.35	1	1	1	3	3.45
No Response	6	2	8	4.71	1	0	0	1	1.15
Home Languages Spoken[b]									
English only	21	30	51	30.00	9	7	8	24	27.59
Spanish only	15	17	32	18.82	5	5	6	16	18.39
Bilingual English/Spanish	21	38	59	34.71	7	13	14	34	39.08
Bilingual English/Other	11	8	19	11.18	6	4	3	13	14.94
Other only	1	2	3	1.76	0	0	0	0	0.00
No Response	3	3	6	3.53	0	0	0	0	0.00
Caregiver's Education[c]									
Eighth grade or less	14	11	25	14.71	7	4	4	15	17.24
Some high school[d]	4	8	12	7.06	2	1	3	6	6.90
High school diploma/GED	10	22	32	18.82	3	5	8	16	18.39
Diploma/GED & technical training certificate	5	7	12	7.06	0	2	2	4	4.60
Some college, no degree	15	21	36	21.18	9	5	7	21	24.14
Associate's, Bachelor's, or Master's Degree	8	6	14	8.24	1	3	2	6	6.90
No Response	16	23	39	22.94	5	9	5	19	21.84
Yearly Family Income									
$11,000 or less	24	31	55	32.35	8	11	9	28	32.18
$11,001–$20,000	18	21	39	22.94	3	2	9	14	16.09
$20,001–$30,000	6	14	20	11.76	5	4	5	14	16.09
$30,001 or more	6	9	15	8.82	3	1	5	9	10.34
No response	18	23	41	24.12	8	11	3	22	25.29

Note. BAU = Business as usual condition. TT = Teaching Together conditions.
[a] Teaching Teaching Enhanced.
[b] Languages Spoken at Home.
[c] Primary Caregiver's Highest Level of Education.
[d] Some high school, no diploma.

ter, disappointed, or *gather* (all words are shown in Online Appendix Table A2).

Alphabet knowledge. The Rapid Letter Naming (RLN) subtest of the CIRCLE Assessment (Landry et al., 2014) was both a screener and outcome measure. This computer-administered task presents a series of upper- and lower-case letters for children to identify as many as possible in 60 s. If a child does not respond in 3 s, a new letter appears.

2.6. Child distal measures

Oral language. Children's broad, receptive vocabulary was measured with the *Peabody Picture Vocabulary Test – Fourth Edition* (PPVT-IV; Dunn & Dunn, 2007) Form A. Children identify the picture of a target word given four possible choices. The PPVT-IV has internal consistency of .94 and test retest reliability of .92. Oral language skills were measured using three subtests from the *Clinical Evaluation of Language Fundamentals Preschool-2* (CELF-P2; Wiig, Secord, & Semel, 2006). First, the Sentence Structure (CELF-SS) subtest measures children's ability to understand spoken sentences of increasing complexity, including prepositional phrases, verb tense, and relative clauses. The Word Structure (CELF-WS) subtest measures children's ability to apply morphological rules to mark inflections and derivations and to use subjective, objective and reflexive pronouns. The Expressive Vocabulary (CELF-EV) subtest measures children's ability to label pictures of objects and actions. These subtests internal consistency coefficient alphas are: SS = .78, WS = .83, and EV = .82.

Narrative language. At posttest only, children's ability to retell simple narratives was assessed with two passages from the *Narrative Language Measures: Preschool* (NLM; Petersen & Spencer, 2012). Children listened to two short stories from the NLM Spring Benchmark Protocol – the first about a character who drops his ice cream and the second about a character who fell while roller skating. Each of children's two retellings were scored in real-time for: (a) story grammar elements present in the child's retelling (14 points); (b) episodes retold (5 points); (c) language complexity including use of sequential and causal terms (i.e., *then, because, when, after*; 9 points), and (d) presence of less common vocabulary in the retelling (e.g., *avoided, toppled*; 6 points). Before administering this task, examiners demonstrated >85% agreement in scoring a set of three recorded retellings. All retellings were videotaped and a subsample of 20% were scored later by a second examiner; average inter-rater reliability was 86.31% in this sample.

Print knowledge. Children's literacy skills were assessed with the Print Knowledge subtest of the *Test of Preschool Early Literacy* (TOPEL; Lonigan, Wagner, Torgesen, & Rashotte, 2007). This 36 item subtest measures letter discrimination, word discrimination, letter-name identification, and letter sound identification. The internal consistency for the Print Knowledge subtest of the TOPEL is .95, test retest reliability is .89 and inter-rater agreement is .96.

Classroom intervention. The classroom-based supplemental curriculum included 23 one-week units used from November to May of the school year. Teachers attended three trainings (7 h each). The first training in early November explained how to use Tier 1 lessons. The second training in late November explained how to use the Tier 2 lessons with four children randomly selected by the researchers. In January, teachers attended a booster training that addressed scaffolding children' linguistic productions, supports for DLLs, and components that fall fidelity observations indicated were inconsistently implemented as designed (see Fidelity section below). Of the 17 randomized *Teaching Together* teachers, 11 teachers (64.71%) attended all three trainings. Lessons were organized into one-week thematic units each consisting of four Tier 1 and four Tier 2 lessons. Units addressed social-emotional themes from 17 narrative texts and science themes from six informational texts. The program embedded most instructional targets on stickers placed at the point of use within the text itself or on the back of instructional materials (e.g., picture cards). These softly scripted, lessons included three meaning-related lessons per week and one code-related lesson per week. Table 1 shows the routine sequence of activities and supports; note how this reflects an additive conditions that increase supports in the classroom and home across tiers. Classroom materials cost an estimated $1100 per classroom.

Tier 1 classroom shared reading. Whole-class interactive book reading lessons occurred four times per week and were designed to take 20 min (90 lessons, total estimate: 1840 min). The four-day Tier 1 lesson sequence comprised: (1) activating background knowledge before reading and asking mostly literal comprehension questions during book reading; (2) teaching academic vocabulary words before reading, asking inferential questions during reading, and challenging children after reading to fill their class "word bank" by using vocabulary in other school/home contexts; (3) referencing print during reading, followed by shared writing after reading; and (4) eliciting children's oral retellings or responses to the text as a culminating activity. After the first lesson, a short 1–2 min phonological awareness (PA) activity used movement or songs to teach sounds in language; this ring of PA activities could be used repeatedly as a transition to small-group time. The Tier 1 curriculum targets are detailed in Online Appendix Fig. A2 and in total included: (a) 66 "big questions" or inferential questions previewed before reading and answered after reading; (b) 181 in-text comprehension questions moving from literal to inferential levels across readings; (c) 117 taught vocabulary words; (d) 111 print references addressing four domains – print meaning, book and organization, letters, and words; (d) 22 review or narrative retelling activities; and (f) 14 PA transition activities.

Tier 2 classroom small-group. The small-group Tier 2 lessons were provided to the subsample of four selected children per classroom. These 10-min lessons occurred four times per week with the lead teacher while other children worked in independent activities or with an assistant teacher. Tier 2 lessons were designed to take 10 min per lesson and were delivered later that day based on each classroom's schedule. Most teachers conducted Tier 1 in the morning and Tier 2 in late morning or after lunch. Tier 2 lessons began in Unit 5 for a total of 19 weeks (76 lessons, total estimate: 760 min; see Online Fig. A3 details). The 4-day lesson sequence included: (1) reviewing the background knowledge activity and the text; (2) reviewing vocabulary cards and introducing a new vocabulary extension activity focused on a high-mileage word; (3) singing songs about letters or PA; and (4) responding to the text with drawing and dictation/writing or playful retellings. The Tier 2 instruction reviewed many Tier 1 activities presented earlier in that day's Tier 1 lesson with adjustments to simplify and scaffold learning of that content. For example, 22 vocabulary extensions sought to breakdown previously taught words using graphic organizers, visual aids, and more explicit instruction. For literacy, 22 print and letter knowledge activities were unique to small- group lessons and included more hands-on manipulatives and explicit instruction than the Tier 1 approaches.

Classroom coaching supports. Coaches co-facilitated PD trainings with the lead author. During implementation, coaches completed classroom coaching visits and emailed 11 modeling videos to teachers on new instructional approaches as they occurred in the curriculum. Teacher coaching began in November and ended in March with a minimum of four coaching sessions per teacher and up to six sessions for teachers who needed additional support ($M = 5.35$ sessions per teacher, $SD = 1.88$). There were three female coaches who coached four to six teachers each. All coaches were curriculum experts who contributed to writing the intervention in earlier development phases. In terms of coaches' backgrounds, two coaches held master's degrees and one held a

bachelor's degree in education and related fields. Coaches used a technical coaching approach (Denton & Hasbrouck, 2009); the goal of technical coaching is to help teachers use new instructional practices and implement curriculum with fidelity (cf. Landry et al., 2017).

Coaches followed a set protocol and logged activities during classroom visits; about 75% of coaches' time was spent observing lessons, providing feedback, and answering teacher questions. During initial coaching visits (and longer for teachers needing support), coaches also modeled lessons or prompted teachers to use each lesson step. At the end of coaching sessions, a brief reflection conversation occurred and a written report was given to teachers summarizing the lesson steps implemented (or skipped). This also recorded areas identified as improvement goals.

2.7. Family Intervention

Family sessions were scheduled according to school and parent availability between November and April of the school year. To reduce costs, Tier 2 Enhanced families were split randomly into two repeated implementation waves of coaching. Wave 1 occurred December–January; wave 2 occurred in February–March. Tier 1 family workshop materials cost an estimated $1500 per facilitator, including consumable materials for about five schools. The Tier 2 family print materials, books and home activities cost an estimated $35–$45 per family. Text message costs follow a tiered pricing model based on number of schools using the service; the estimated costs were about $100 per school ($1750 total for all 17 intervention schools).

Tier 1 parent–child workshops. All families with children enrolled in treatment classrooms were invited to approximately monthly after-school workshops on these topics: (1) Building Trust and Cooperation, (2) Talking is Teaching, (3) Making Books Come Alive, (4) Playing with Letters and Sounds, (5) Building Your Child's Vocabulary, and (6) Writing Together. The workshops were developed in partnership with the local children's museum and advertised to families by teachers and research staff through the use of flyers sent home in children' backpacks or given directly to parents. Each 60-min workshop included three components: (a) a 2-min video and discussion of two strategies to support early language and literacy skills, (b) an interactive shared reading session led by the workshop facilitator, and (c) four or five hands-on activity stations where parents and children applied the focal strategies with facilitator support. Parents received an informational handout on the strategies and a family museum pass to encourage informal learning with activities and exhibits designed to promote language and literacy. Bilingual workshop facilitators were trained in a 7-h session by the lead author and a museum parenting expert. After this training, facilitators completed role-play with feedback from the museum expert. The four female facilitators all had a bachelor's degree or higher and were fluent in English and Spanish; they had one to eight years of past experience providing interventions to parents. Each facilitator served an average of four schools. Most workshops occurred in English with Spanish subtitles in videos and bilingual handouts; however, in schools with high numbers of Spanish-speaking families a bilingual approach was used.

Tier 1 text messages. For 15 weeks of the school program, all parents in intervention classrooms were invited to receive two SMS text messages per week with information about how to support their child's school learning. Parents could sign up for text messages via a form sent by the child's teacher or at the workshops. The first text message was sent on Mondays and explained what book teachers were reading and provided conversation starters to extend learning at home. The second text message was sent on Thursdays and featured short videos (English/Spanish) designed to show families how to find learning moments in day-to-day activities. These modeling videos were developed by a commercial partner that provides mobile technology to support family engagement. Sample texts are in Online Fig. A3.

Tier 2 Basic parent resources. Families assigned to the Tier 2 Basic group received a phone call from research staff explaining they were eligible to receive books and materials to support their child's language and literacy learning at home. Parents were offered materials in English or bilingual English/Spanish, according to their preference. Through "backpack mail" from the child's teacher, parents were given a set of materials including: (a) four books that included questions and vocabulary to discuss that mirrored classroom instruction with the same titles; (b) an alphabet mat with plastic letters and three letter activity instructions; (c) a ring of conversation starter prompts for use at mealtime or during wait times; (d) an activity to support emotion understanding; and (e) finger puppet materials to retell a provided book.

Tier 2 Enhanced parent coaching. Families assigned to the Tier 2 Enhanced group received a phone call from research staff explaining they were eligible for four, one-on-one coaching sessions along with the set of books and materials described above. Coaching sessions were provided by the same four workshop facilitators; thus, some families had already made personal connections with these coaches. Each coach served 4–14 families. Coaching sessions occurred in a four-week period (range 3–5 weeks). The first coaching session occurred in the family's home and the remaining sessions were facilitated remotely via phone call.

At the first session, the participating parent/caregiver received (a) a loaned iPad that included a series of four instructional videos (10 min each) that explained strategies to support their child's language and literacy skills including modeling videos that exemplified each strategy; (b) a bilingual family manual that further detailed caregiver strategies and included a list of activities that could support learning; (c) a schedule for viewing the video content and videotaping the parent–child dyad applying the strategies during a short home activity; and (d) instructions for securely sharing the video with their coach and a reflection sheet the parent to complete before their coaching call. Each coaching cycle include: (1) a training video that shares information and models strategies parents use to support their child's learning; (2) practicing strategies while video recording a short parent–child interaction as well as during provided activities; (3) discussing the video with their coach as a basis for self-reflection with oral support from their coach. The coaching sessions addressed similar topics as the workshops but with greater depth of content: (1) Building Trust and Cooperation, (2) Talking is Teaching, (3) Making Books Come Alive, and (4) Playing with Letters and Sounds. The first in-home coaching session required an average of 1 hour, 48 min. The remaining phone sessions required an average of 24 min each (Range = 15–30 min) to discuss that week's topic, check-in on specific interactions the parent and child had during the week to support language and literacy, and to review ways the parent supported his/her child's learning in the videos uploaded that week. Before the call, coaches noted specific moments from the uploaded video; on the call, they provided praise or positive feedback and asked parents questions to encourage reflection.

Tier 2 text messages. For up to 15 weeks, parents of children selected for Tier 2 basic and enhanced groups were invited to receive two additional SMS text messages per week about how to support their child's school learning; Tier 2 text messages were identical across these two conditions. Parents could sign up for text messages during their first coaching session or via a form sent by the child's teacher. Tier 2 families received additional texts on Tuesdays and Wednesdays. These included specific things for parents to say and do with their child such as, "As you read a book ask your child, 'What does this make you think about?' It's better to focus on what your child is interested in than finishing a whole book in

one sitting." or "Have fun hunting for letters your child knows! At the grocery store, say 'Look for letters you know!'" Text messages were adapted from our prior work (Cabell et al., 2019).

2.8. Description of business-as-usual condition

Across conditions, all teachers reported using a commercial curriculum as their core curriculum with 75.76% naming Frog Street Pre-K as their core curriculum and 15.15% using the Creative Curriculum Preschool. At pretest, we surveyed teachers on their classroom instructional activities and curricular focus using a set of 17 items adapted from the Early Childhood Longitudinal Study-Kindergarten Spring Teacher Questionnaire (1999). Questions used 5-point rating (*0*-never to *4*-three or four times per week) and asked how often do children in your class do: (a) code-related activities (six items), such as work on learning the names of letters, or (b) 11 meaning-related activities, such as retelling stories read aloud. In general, teachers reported providing greater exposure to code-related activities ($M = 3.42$) than meaning-related activities ($M = 2.69$) at baseline, $t(32) = 7.83$, $p < .001$. Differences between conditions were not significant for exposure to code-related activities [$t(30) = .73$, $p = .47$] or meaning-related activities [$t(30) = .73$, $p = .47$]. Given that Head Start prioritizes family engagement, in February, we surveyed all teachers and families on their school's provision of learning and volunteer opportunities for families. We adapted a Head Start Family Engagement questionnaire survey (Zill et al., 2000). This included 12 items assessing the quantity of family opportunities offered this school year (0 times, 1 time, . . ., 4 or more times) and 10 items assessing the quality of these offerings on a 3-point rating scale (*1*-does not do it all, *2*-does it okay, *3*-does it very well). These data showed a large number of parent education opportunities at all schools with an average of 31.57 ($SD = 8.81$) parent engagement opportunities in BAU schools and 26.00 ($SD = 8.18$) in schools in the *Teaching Together* group. These group differences were not significant, $t(26) = -1.73$, $p = .095$. Most teachers rated these family engagement opportunities as moderate to high quality [BAU $M = 2.80$, $SD = .16$; Intervention groups $M = 2.45$, $SD = .48$]. Teachers rated the quality significantly higher in the BAU condition, $t(15.93) = -2.59$, $p = .021$.

2.9. Implementation fidelity

From a design perspective within this pilot study, we set a minimum threshold of 75% implementation, representing moderate high implementation (cf. Hill & Erickson, 2019). The Online Supplementary Tables A3–A4 (pp. 12–13) detail fidelity monitoring methods and results.

Classroom implementation. We collected dosage and adherence data to understand classroom intervention delivery. Key findings were that teachers reported delivering an average dosage of 79.05% of the 91 whole-group lessons and 66.14% of the 76 small-group lessons monitored. Teachers' adherence to implementing lesson steps in small group (82.81%) was slightly higher than whole group (77.12%).

Family implementation. To monitor fidelity of delivering the family intervention, we collected: (a) attendance data for family workshops, (b) parents' receipt of text messages, and (c) engagement with the individualized coaching sessions. Average attendance at the six workshops did not noticeably differ by treatment groups, but was well below our minimum threshold: Tier 1 $M = 1.68$ ($SD = 1.56$), Tier 2 Basic $M = 1.41$ ($SD = 1.43$), Tier 2 Enhanced $M = 1.57$ ($SD = 1.22$). For the parent text messages, 79.38–84.65% of messages of messages were delivered, but only 24.96% of modeling video links were clicked. Tier 2 families reported using the eight home activities/books an average of only 2.50–2.75 times ($SD = 1.26$).

For the Tier 2 Enhanced families offered four coaching sessions, >80% of families completed some coaching, but five families declined. Parents' adherence to the coaching components was 84.65% and their engagement during coaching calls was high.

Researcher implementation. We also monitored the research staff's delivery of the intervention. This was high for all components, ranging from 87.30% to 94.39% adherence.

Fidelity interpretation. In sum, participants' average implementation of the intervention met our design criteria (>75%) for the classroom curriculum implementation. Yet, we observed moderate-low implementation of literacy components signaled the need for classroom literacy revisions before future implementation. Over 80% of parents participated in some coaching and with good adherence and engagement levels, signaling feasibility. Yet other family components required revision because parents only attended about a quarter of the intended workshops and opened 25% of text messages with modeling videos.

2.10. Data analytic plan

Primary analyses used linear mixed methods models (i.e., multilevel models; LMM) to examine main effects of the intervention on proximal and distal outcome variables, followed by exploratory moderator models to examine potential interactive effects of the intervention moderated by baseline performance or DLL status. Finally, exploratory analyses via linear quantile mixed methods models (i.e., multilevel quantile models; LQMM) tested the extent to which the impact on the selected outcomes varied across the conditional distribution of the outcome variables themselves. LMM and LQMM were both used to account for the nested nature of the data such that variance in the selected outcomes could be decomposed and attributed to group-level (i.e., classroom) and individual-level (i.e., children) clustering. Subsequently, the standard errors for the individual predictors of the model are appropriately estimated account for the clustering of children in classrooms.

Linear mixed models (LMM). Proximal outcomes on which we tested for effects included the Inferential comprehension, TT Taught, and RLN assessments, and the distal outcomes included the PPVT, CELF-EV, CELF-SS, CELF-WS, NLM, and TOPEL. For each of these outcomes, the respective baseline pretest measure, race, DLL status, age, and sex were included as covariates in order to increase power. Following the test of primary impacts, moderator analyses tested whether the level of intervention impact was separately dependent on either DLL status or initial pretest performance. The nature of the study design was an underpowered efficacy trial, thus, to account for the increased number of participants required to detect significant interaction effects, a p-value of .10 was used to determine significance whether any moderated effects should be probed for simple slope differences. The lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015) was used to estimate all main effect and moderator LMMs, which helped us address research questions 1a and 1b, as well as 2a and 2b.

Linear quantile mixed models (LQMM). The models estimated with LMM estimate the mean effect of intervention on the outcome; however, it is plausible that the *average* treatment effect may mask the fact that the impact was more localized at a point of the conditional distribution of the outcome that was not the mean (Koenker & Bassett, 1978; Petscher, 2016). Like other studies with heterogeneous effects for children with or at risk for reading difficulties (Solari et al., 2018; Wanzek et al., 2016), it is possible that the effect of the different intensities of intervention may be larger at the lower end of the conditional distribution of the dependent variable compared to the higher end, or vice versa. Quantile regression examines the effects at different points of the conditional distribution of the outcome variable other than just the conditional mean as in LMMs. Quantile regression does not split the sample into multiple groups

Table 4
Correlations, means, and standard deviations of included variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. InfPost																								
2. TTPost	**.57**																							
3. RLNPost	**.20**	**.15**																						
4. PPVTPost	**.74**	**.60**	**.31**																					
5. CELF-EV Post	**.70**	**.52**	**.23**	**.76**																				
6. CELF-SSPost	**.65**	**.43**	**.22**	**.69**	**.58**																			
7. CELF-WSPost	**.74**	**.56**	**.24**	**.72**	**.75**	**.61**																		
8. NarrativePost	**.47**	**.30**	**.23**	**.40**	**.40**	**.42**	**.43**																	
9. TOPELPost	**.38**	**.31**	**.77**	**.49**	**.34**	**.40**	**.39**	**.35**																
10. InfPre	**.71**	**.44**	.15	**.65**	**.68**	**.61**	**.68**	**.47**	**.31**															
11. TTPre	**.34**	**.32**	**.23**	**.41**	**.42**	**.26**	**.38**	**.18**	**.35**	**.30**														
12. RLNPre	**.20**	**.24**	**.66**	**.30**	**.27**	**.26**	**.29**	.13	**.57**	**.20**	**.24**													
13. PPVTPre	**.72**	**.54**	**.26**	**.84**	**.81**	**.65**	**.75**	**.44**	**.37**	**.69**	**.36**	**.37**												
14. CELF-EVPre	**.67**	**.47**	.12	**.70**	**.86**	**.54**	**.73**	**.45**	**.26**	**.69**	**.37**	**.20**	**.77**											
15. CELF-SSPre	**.56**	**.40**	**.23**	**.59**	**.56**	**.59**	**.61**	**.35**	**.36**	**.54**	**.33**	**.28**	**.60**	**.54**										
16. CELF-WSPre	**.67**	**.59**	**.24**	**.72**	**.70**	**.55**	**.80**	**.43**	**.42**	**.61**	**.42**	**.31**	**.75**	**.70**	**.59**									
17. TOPELPre	**.32**	**.32**	**.66**	**.44**	**.35**	**.41**	**.41**	**.23**	**.69**	**.36**	**.20**	**.81**	**.46**	**.31**	**.36**	**.40**								
18. Race	**.24**	**.18**	**.21**	**.29**	**.37**	**.21**	**.35**	**.24**	**.21**	**.29**	.07	**.22**	**.35**	**.41**	**.21**	**.28**	**.29**							
19. DLL	**-.17**	-.04	**.20**	-.13	**-.25**	-.04	**-.22**	-.09	.15	-.14	-.03	**.26**	**-.16**	**-.27**	-.02	**-.27**	**.20**	-.12						
20. Age	**.32**	**.22**	**.35**	**.37**	**.31**	**.29**	**.27**	**.27**	**.35**	**.34**	**.30**	**.31**	**.32**	**.25**	**.27**	**.33**	**.32**	.06	.13					
21. Sex	**-.16**	-.03	-.02	.00	.05	-.13	-.11	-.14	-.05	-.01	.01	.04	.00	-.01	.02	-.06	.06	-.07	**.17**	.00				
22. Tier1	.00	.07	-.09	.10	.12	.07	-.02	-.03	-.07	-.05	.07	-.02	.03	.11	-.09	.05	-.06	-.03	-.13	.05	-.04			
23. Tier2B	-.13	-.07	-.09	-.11	-.11	-.12	-.11	-.05	-.09	-.01	-.11	-.08	**-.16**	-.06	-.03	**-.19**	-.15	.01	.05	-.03	.04	**-.24**		
24. Tier2E	**.17**	.15	-.11	.05	.01	-.01	.08	.08	-.08	.10	.04	-.05	.10	.05	.14	.13	.05	-.09	.01	.12	.13	**-.25**	**-.21**	
Mean	2.14	14.16	16.53	60.47	11.86	10.22	7.12	7.65	20.12	1.32	9.37	6.26	49.45	8.30	7.48	5.18	12.7	2.88	.82	53.35	1.51	.22	.17	.18
SD	1.75	3.88	12.42	20.01	7.38	4.11	5.02	4.24	10.57	1.38	2.77	8.42	19.00	6.18	4.13	4.3	9.59	2.18	.65	4.49	.50	.42	.38	.39
SKEW	.71	-0.20	.64	.00	.43	-.20	.56	.06	-.18	1.17	-.04	1.95	.06	.51	.30	.90	.73	.52	.18	-.09	-.05	1.33	1.75	1.65
KURT	2.72	3.20	2.78	2.64	2.34	2.34	2.32	2.63	1.69	3.77	2.7	6.11	2.7	2.33	2.24	3.15	2.40	1.76	2.34	2.14	1.00	2.76	4.07	3.71
% Missing	.05	.03	.03	.05	.05	.05	.05	.03	.05	.01	.01	.00	.04	.00	.00	.01	.00	.04	.04	.00	.00	.00	.00	.00

Note. InfPost = Inferential posttest; TTPost = Teaching Together taught posttest; RLNPost = Rapid letter naming posttest; PPVTPost = Peabody picture vocabulary test posttest; CELF-EV Post = Clinical Evaluation of Language Fundamentals-Expressive vocabulary posttest; CELF-SSPost = Clinical Evaluation of Language Fundamentals-Sentence structure posttest; CELF-WSPost = Clinical Evaluation of Language Fundamentals-Word structure posttest; NarrativePost = Narrative Language Measures posttest; TOPELPost = Test of preschool early literacy-Print Knowledge posttest; Inferential pretest; TTPre = Teaching Together taught pretest; RLNPre = Rapid letter naming pretest; PPVTPre = Peabody picture vocabulary test pretest; CELF-EV Pre = Clinical Evaluation of Language Fundamentals-Expressive vocabulary pretest; CELF-SSPre = Clinical Evaluation of Language Fundamentals-Sentence structure pretest; CELF-WSPre = Clinical Evaluation of Language Fundamentals-Word structure pretest; TOPELPre = Test of Preschool Early Literacy-Print Knowledge pretest; DLL = Dual language learner status; Sex = Female(1), Male(2); Tier1 = Tier 1 intervention; Tier2B = Tier 2B intervention; Tier2E = Tier 2E intervention; SKEW = Skewness; KURT = Kurtosis.
All statistically significant correlations are bolded. Relations between interventions and posttest in dashed box.
$r \geq .16$ or $\leq -.16$ (sig. at $p < .05$). $r \geq .20$ or $\leq -.20$ (sig. at $p < .01$).

(e.g., quartiles, quintiles, deciles) but rather uses the full sample. Just as conditional mean models – like multiple regression, multilevel models, and structural equation models – use the full sample of data but results are estimated based on the conditional mean of the dependent variable given the independent variables (i.e., all data points are used, not just those at the conditional mean), so too does quantile regression use the full sample of data but provides estimates based on other conditional points of the dependent variable given the independent variable. In this way we were able to test for effects specified in research question 2c. The LQMM package developed in R (Geraci, 2014) leverages the advantages of quantile regression while also partitioning the variance into group and individual level variance similar to LMM.

3. Results

3.1. Descriptive statistics

Values for descriptive statistics, correlation coefficients, normality, and data missingness are listed in Table 4. For the inferential listening comprehension measure with a maximum score 16 points, means are very low at both at pretest ($M = 1.32$) and posttest ($M = 2.14$). On the proximal *Teaching Together* vocabulary task, children only identified about half of the 28 presented words ($M = 14.16$) at posttest. The pretest literacy means ($M = 6.26$) show children scored below expected screening benchmarks (raw scores below 10) on RLN, but made gains by posttest ($M = 16.53$). None of the values indicated departures from normality for skewness ($\leq \pm 2$) or kurtosis ($\leq \pm 7$) suggested by Curran, West, and Finch (1996). Missing data ranged from 0% to 5%; no variable had substantial missing values. Little's test of data missing completely at random (MCAR) resulted in a fail-to-reject, $\chi^2(111) = 103.03$, $p = .692$ indicating that full information maximum likelihood, as used in this study, was appropriate in our estimation. Of the relations between levels of the intervention (i.e., Tier 1, Tier 2 Basic, Tier 2 Enhanced) and the outcome measures (indicated by the dashed box in Table 4, only the Inferential comprehension post-test was significantly related to the Tier 2 Enhanced ($r = .17$, $p < .05$). Finally, of the nine outcome variables, the covariates child race and age were significantly related to all nine, DLL status was related to four (three negatively r's = −.17 to −.25), and sex was related to one variable.

3.2. Baseline differences

Baseline mean and standard deviations for treatment groups are in Table 5. Independent sample t-tests were performed to determine if any group differences existed prior to receiving treatment. For the Tier 2 Basic group, results indicated a statistically significant difference between the PPVT treatment ($M = 42.71$, $SD = 17.23$) and the BAU group ($M = 49.87$, $SD = 16.63$), $t(162) = -2.08$, $p < .05$, Hedge's $g = -.42$), and the CELF-WS treatment ($M = 3.41$, $SD = 3.65$) and BAU group ($M = 5.18$, $SD = 3.92$, $t(167) = -2.47$, $p < .05$, Hedge's $g = -.46$). The remaining measures indicated no significant differences at baseline; however, some other non-significant, small to moderate baseline differences were observed (e.g., $g = -.49$ Tier 2 Basic lower vs. BAU on TOPEL PK; $g = .33$ Tier 2 Enhanced higher vs. BAU on CELF-SS).

We recognize that the lack of baseline equivalence (i.e., $g > .25$) introduces levels of bias in the estimation of treatment effects and would not let this study meet a group design standard (What Works Clearinghouse, 2019). Because this study is a preliminary design and pilot testing study, we opted to use a residualized gain score as our outcome to understand group differences. 3.65) and BAU group ($M = 5.18$, $SD = 3.92$, $t(167) = -2.47$, $p < .05$, Hedge's $g = -.46$). The remaining measures indicated no significant differences at baseline.

3.3. Research Question 1: What are the Effects of Increasing Tiers of School & Home Interventions Relative to the Control Condition?

Linear mixed model (LMM). LMMs were performed for three models, with the results of the random effect variance due to classroom level factors listed in Supplementary Table A5. The portion of

Table 5
Baseline mean and SD for individual intervention groups for each baseline measure.

	Tier 1					Tier 2 Basic					Tier 2 Enhanced				
	M_{TX}	SD_{TX}	M_{BAU}	SD_{BAU}	*g*	M_{TX}	SD_{TX}	M_{BAU}	SD_{BAU}	*g*	M_{TX}	SD_{TX}	M_{BAU}	SD_{BAU}	*g*
Baseline															
Inferential	1.20	1.28	1.26	1.31	0.05	1.29	1.58	1.26	1.31	−.02	1.61	1.48	1.26	1.31	.25
TT Taught	9.73	2.65	9.38	2.70	.13	8.69	3.20	9.38	2.70	−.23	9.58	2.67	9.38	2.70	.07
RLN	5.92	8.96	7.39	9.05	−.16	4.79	6.63	7.39	9.05	−.32	5.42	7.70	7.39	9.05	−.23
PPVT	50.50	23.62	49.87	16.63	.03	42.71*	17.23	49.87	16.63	−.42	53.5	19.19	49.87	16.63	.20
CELF-EV	9.55	6.60	7.72	5.75	.29	7.45	7.28	7.72	5.75	−.04	8.90	5.47	7.72	5.75	.21
CELF-SS	6.82	4.35	7.43	4.18	−.14	7.17	4.00	7.43	4.18	−.06	8.68	3.80	7.43	4.18	.31
CELF-WS[!]	5.59	4.72	5.18	3.92	.09	3.41*	3.65	5.18	3.92	−.46	6.35	4.85	5.18	3.92	.26
TOPEL-PK	11.70	9.41	14.03	9.60	−.25	9.55	8.58	14.03	9.60	−.49	13.81	10.30	14.03	9.60	.02

Note. M_{TX} = Mean treatment group; SD_{TX} = Standard deviation treatment group; M_{BAU} = Mean business as usual group; SD_{BAU} = Standard deviation business as usual group; g = Hedge's *g* effect size for difference between treatment and non-treatment group; Tier1 = Tier 1 intervention; Tier2B = Tier 2 Basic intervention; Tier2E = Tier 2 Enhanced intervention; TT Taught = Teaching Together taught; RLN = CIRCLE Rapid Letter Naming subtest; PPVT = Peabody Picture Vocabulary Test; CELF-EV = Clinical Evaluation of Language Fundamentals-Expressive vocabulary; CELF-SS = Clinical Evaluation of Language Fundamentals-Sentence structure; CELF-WS = Clinical Evaluation of Language Fundamentals-Word structure; [!] = CELF-WS was used as baseline measure for NLM outcome measure; TOPEL-PK = Test of Preschool Early Literacy-Print Knowledge subtest.
* $p < .05$.

Table 6
Main effects models for individual interventions on proximal outcome measures ($N = 170$).

	Inferential				TT Taught				RLN			
	β	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*
Model 1 – Proximal Subgroup												
Intercept	1.34	1.29	.31		11.95	4.00	**		−1.04	9.50	.91	
Race	.02	.05	.68		.25	.15	.10		.18	.37	.64	
DLL	−.10	.16	.52		.10	.50	.85		.55	1.24	.66	
Age	.03	.02	.26		.03	.07	.70		.38	.17	*	
Sex	−.46	.20	*		−.41	.60	.49		−1.36	1.28	.29	
Tier1	.03	.26	.91	.02	.73	.86	.40	.19	−3.37	2.53	.19	−.27
Tier2B	−.44	.27	.11	−.25	.15	.89	.87	.04	−2.69	2.57	.30	−.22
Tier2E	.44	.27	.11	.25	1.78	.87	*	.46	−3.71	2.56	.16	−.30
Pretest	.84[!]	.08[!]	***[!]		.40[$]	.11[$]	***[$]		.94[^]	.09[^]	***[^]	

Note. TT Taught = Teaching together taught words; RLN = CIRCLE Rapid Letter Naming subtest; DLL = Dual language learner status; Sex = Female(1), Male(2); Pretest = Pretest baseline control for post-test measure ([!] = Inferential pretest; [$] = TT taught pretest; [^] = RLN pretest); Tier1 = Tier 1 intervention; Tier2B = Tier 2B intervention; Tier2E = Tier 2E intervention. All predictor variables were centered (0/1 for dichotomous variables and grand-mean centered for continuous variables) such that Beta coefficients reflect fitted deflections from the intercept where the intercept represents the referent BAU.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

variance attributed to classroom level factors was substantial for RLN (ICC = 32%) in Model 1 and Model 3, but did not exceed 10% for any of the other eight outcome variables in any of the models. The fixed effects of the individual intervention factors for the proximal and distal outcome measures are summarized in Tables 6 and 7, respectively. A positive statistically significant effect was indicated for the relation between the Tier 2 Enhanced intervention relative to the BAU and the proximal outcome TT taught, $t = 2.05$, $p < .05$, $g = .46$. No other significant proximal impacts were identified relative to the BAU; however, some effect sizes may be meaningful (e.g., $g = .25$ Tier 2 Enhanced effect on inferential comprehension; range of negative effects on RLN $g = -.22$ to $-.30$). For distal outcomes, a positive effect size, though not statistically significant, was observed for the relation between Tier 1 Only and the distal outcome PPVT relative to the BAU, $t = 1.95$, $p = .058$, $g = .39$. Surprisingly, a negative impact was indicated for the relation between Tier 2 Enhanced and the distal outcome relative to the BAU on the TOPEL, $t = -2.25$, $p < .05$, $g = -.39$, and a negative marginal effect was observed between Tier 2 Basic, relative to BAU, on the distal outcome CELF-EV, $t = -1.72$, $p = .09$, $g = -.20$.

3.4. Research Question 2: Which Children Benefit More or Less from the Interventions?

Our second research question used two exploratory models to: (1) understand main effects at different points of the conditional distribution of the outcomes (i.e., LQMM Main Effects), and (2) understand the extent to which the impact of the intervention on the post-test was moderated by key child baseline characteristics (i.e., LMM Moderation).

LQMM Main Effects. The results of the LQMMs clarified the primary impact findings (Supplementary Table A6) and indicated patterns of effects that were not identifiable in LMM.

TT Taught. The result from the main effect LMM previously reported for the statistically significant medium effect of the Tier 2 Enhanced intervention on TT Taught relative to the BAU ($g = .46$) indicated an increase across the distribution from the lowest quantile (.10 quantile $g = .30$) to the highest quantile (.90 quantile $g = .79$) in the LQMM, with the largest effects concentrated in the .70 through .90 quantile.

PPVT. The small but not statistically significant effect of the Tier 1 intervention on PPVT relative to the BAU found in the LMM ($g = .39$) showed a pattern of results in the LQMM where the effects were concentrated at the lowest (.10 quantile = .27; .20 quantile = .25) and highest quantiles (.80 quantile = .61; .90 quantile = .57), with the effects being more than two times as large at the higher end of the distribution.

CELF-EV. Negative effects of Tier 2 Basic on CELF-EV had an effect size at the .50 quantile in the LQMM ($g = -.17$) similar to that of the LMM ($g = -.20$), but decreased toward the .10 quantile ($g = .00$) and increased toward the .90 quantile (−.32).

Table 7
Main effects models for individual interventions on distal outcome measures ($N = 170$).

	PPVT				CELF-EV				CELF-SS				CELF-WS				Narrative				TOPEL			
	β	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*	*β*	S.E.	*p*	*g*
Model 2 – Distal Subgroup																								
Intercept	37.02	12.00	**		−.66	3.97	.87		5.25	3.41	.14		5.76	3.31	†		1.51	4.07	.72		2.36	8.35	.78	
Race	−.06	.45	.90		.10	.15	.55		.15	.13	.26		.28	.12	*		.25	.15	.11		.00	.32	1.00	
DLL	.14	1.46	.93		−.37	.51	.48		.08	.42	.85		.08	.42	.85		.32	.53	.55		.01	1.05	.99	
Age	.39	.22	†		.21	.07	**		.12	.06	†		.03	.06	.60		.12	.08	.11		.38	.15	*	
Sex	.27	1.84	.88		1.12	.61	†		−1.03	.54	†		−.82	.50	.11		−.89	.60	.14		−1.00	1.25	.43	
Tier1	4.85	2.49	†	.39	−.25	.84	.77	−.03	.52	.72	.48	.13	−.53	.70	.45	−.11	−.90	.90	.32	−.21	−2.70	1.78	.14	−.25
Tier2B	1.86	2.61	.48	.15	−1.50	.87	†	−.20	−1.21	.75	.11	−.29	.28	.73	.71	.06	.15	.93	.87	.04	−1.01	1.86	.59	−.09
Tier2E	1.16	2.55	.65	.09	−1.25	.86	.15	−.17	−.91	.74	.23	−.22	−.04	.72	.95	−.01	.35	.92	.71	.08	−4.09	1.82	*	−.39
Pretest	.86[!]	.05[!]	***!		.96	.06	***		.55	.07	***		.88	.07	***		.36	.08	***		.71	.07	***	

Note. PPVT = Peabody picture vocabulary test; CELF-EV = Clinical Evaluation of Language Fundamentals-Expressive vocabulary test; CELF-SS = Clinical Evaluation of Language Fundamentals-Sentence structure test; CELF-WS = Clinical Evaluation of Language Fundamentals-Word structure test; Narrative = Narrative Language Measures; TOPEL = Test of Preschool Early Literacy-Print Knowledge; DLL = Dual language learner status; Sex = Female(1), Male(2); Pretest = Pretest baseline control for post-test measure ([!] = PPVT pretest; [^] = CELF-EV pretest; [$] = CELF-SS pretest; [#] = CELF-WS pretest; [&] = TOPEL pretest). Tier1 = Tier 1 intervention; Tier2B = Tier 2B intervention; Tier2E = Tier 2E intervention. All predictor variables were centered (0/1 for dichotomous variables and grand-mean centered for continuous variables) such that Beta coefficients reflect fitted deflections from the intercept where the intercept represents the referent BAU.

* $p < .05$.
** $p < .01$.
*** $p < .001$.
† $p < .10$.

CELF-SS. Other notable effects revealing a discrepancy between the results of the LMM and LQMM included the relation between the Tier 1 treatment and CELF-SS, where the effect size was .13 in the LMM but indicated a practically meaningful, positive effect of .58 and .27 at the .10 and .20 quantiles respectively relative to the BAU, but negative coefficients for the .80 quantile ($g = -.22$) and .90 quantiles ($g = -.14$) relative to the BAU.

CELF-WS. On CELF-WS, Tier 2 Enhanced had a negligible effect size of −.01 in the LMM but departed from that effect in the negative direction for the .20 through .60 quantiles ($g = -.14, -.60$) and the positive direction for the .80 ($g = .16$) and .90 quantiles ($g = .15$), all relative to the BAU.

NLM. The negligible effect of Tier 2 Enhanced on NLM may have been deceiving in the LMM ($g = .08$) given the large difference between the largest negative effect (.10 quantile = −.20) and largest positive effect (.90 quantile = .99) at the poles of the distribution.

TOPEL. The significant negative effect of Tier 2 Enhanced on TOPEL relative to the BAU in the LMM ($g = -.39$) was fairly consistent across the distribution in the LQMM.

LMM Moderation. Results from the LMM moderator analyses show that DLL status significantly moderated the relation between Tier 2 Basic and RLN, $t = 2.07$, $p < .05$, such that individuals with DLL status in Tier 2 Basic performed lower on RLN than DLL status in the control (see Table 8).

4. Discussion

This underpowered, randomized controlled trial was the initial evaluation of a tiered school plus home curricular supplement, *Teaching Together*, which primarily targets academic language, but also supports early literacy skills. We examined the effects of increasingly intensive classroom instruction blended with family supports. Our sample included Head Start children demonstrating risk for later reading difficulties; about 66% of these children were DLLs. All focal children scored below predetermined language and literacy screening benchmarks and were Tier 2 eligible. Classrooms were randomly assigned to either the control or treatment condition. Within treatment classrooms, these high-risk children were further randomized to receive Tier 1 Only Classroom or additional Tier 2 Classroom lessons. After classroom intervention began, parents and children were further randomized to one of three treatment groups – Tier 1 Family, Tier 2 Basic Family, or Tier 2 Enhanced Family.

The additive study design sought to understand the benefits of aligned interventions delivered by teachers and parents. The Tier 1 Only condition included large-group shared book reading activities at school and parent education via workshops and text message communication; this condition showed non-significant, albeit promising, effects on one vocabulary measure. The Tier 2 Basic treatment added small-group instruction at school and provided parents with home learning materials and additional text messages. Interestingly, this Tier 2 Basic condition showed the least evidence of promise. The Tier 2 Enhanced condition added four individualized parent coaching sessions to the above supports and showed the most evidence of promise on language outcomes. Unexpectedly, all treatment groups were negatively related to literacy outcomes. We explored differential effects by skill level and DLL status because, even within an at risk sample, we expected heterogeneity in response (Cabell, Justice, Konold, & McGinty, 2011; NASEM, 2017; Wanzek et al., 2016).

4.1. Impacts on child language outcomes

The *Teaching Together* program trained teachers and parents to explain sophisticated words and facilitate inferential conver-

Table 8
Moderator analyses for individual interventions on selected outcome measures ($N = 170$).

	Inferential				TT Taught				RLN				TOPEL			
	β	S.E.	p	g	β	S.E.	p	g	β	S.E.	p	g	β	S.E.	p	g
Model 3 – Moderator Primary Subgroup																
Intercept	1.38	1.30	.30		12.43	4.11	**		−3.48	9.72	.73		.86	8.47	.92	
Race	.02	.05	.72		.25	.15	.10		.26	.38	.51		.01	.32	.98	
DLL	−.08	.23	.75		−.06	.72	.94		−.79	1.77	.66		−.46	1.54	.77	
Age	.03	.02	.30		.02	.08	.82		.44	.18	*		.42	.16	*	
Sex	−.43	.20	*		−.28	.61	.65		−1.31	1.30	.32		−1.12	1.27	.39	
Tier1	.34	.40	.40	.19	1.52	1.33	.26	.39	−4.29	3.33	.20	−.35	−1.81	2.64	.50	−.17
Tier2B	−.95	.51	†	−.55	−1.10	1.57	.49	−.28	−9.06	3.90	*	−.75	−5.31	3.23	.11	−.51
Tier2E	.55	.46	.23	.31	.62	1.45	.67	.16	−3.84	3.67	.30	−.31	−4.66	2.99	.13	−.38
Pretest	.74!	.12!	***!		.35$	.17$	*$		.94^	.12^	***^		.67&	.11&	***&	
Tier1*Pretest	.27!	.20!	.18!		.14$	.31$	.65$		−.06^	.21^	.79^		.23&	.19&	.22&	
Tier2B*Pretest	.03!	.19!	.86!		−.03$	.27$	.90$		−.26^	.27^	.33^		−.05&	.20&	.82&	
Tier2E*Pretest	.23!	.19!	.24!		.18$	.30$	.57$		.20^	.22^	.38^		.04&	.18&	.82&	
DLL*Tier1	−.45	.41	.29		−1.28	1.29	.33		1.03	2.96	.73		−1.37	2.65	.61	
DLL*Tier2B	.54	.47	.26		1.26	1.39	.37		6.66	3.22	*		4.34	2.89	.14	
DLL*Tier2E	−.18	.43	.68		1.35	1.35	.33		.33	3.17	.92		.57	2.87	.85	

Note. TT Taught = Teaching together taught words; RLN = CIRCLE Rapid Letter Naming subtest; TOPEL = Test of Preschool Early Literacy-Print Knowledge subtest; DLL = Dual language learner status; Sex = Female(1), Male(2); Pretest = Pretest baseline control for post-test measure (! = Inferential pretest; $ = TT taught pretest; ^ = RLN pretest, & = TOPEL pretest); Tier1 = Tier 1 intervention; Tier2B = Tier 2B intervention; Tier2E = Tier 2E intervention. All predictor variables were centered (0/1 for dichotomous variables and grand-mean centered for continuous variables) such that Beta coefficients reflect fitted deflections from the intercept where the intercept represents the referent BAU.

* $p < .05$.
** $p < .01$.
*** $p < .001$.
† $p < .10$.

sations. We found promising results for four of seven language outcomes (taught vocabulary, receptive vocabulary, inferential comprehension, narrative) that were strongest for the Tier 2 Enhanced group and somewhat promising for the Tier 1 Only group. No main effects resulted for the Tier 2 Basic condition. We analyze possible explanations for these findings beginning with the most promising treatments.

Enhanced Tier 2 with parent coaching. The most intensive, Tier 2 Enhanced group significantly improved children's proximal vocabulary for taught words ($g = .46$). Children in this group also made practically meaningful, albeit non-significant, gains in inferential comprehension ($g = .25$). Despite non-significant main effects, quantile regression showed the highest .90 percentile of this treatment resulted in large gains ($g = .99$) on a narrative recall task. We similarly observed a pattern of larger, localized effects for Tier 2 Enhanced at the higher end of the distribution for the proximal vocabulary and the CELF word structure measures. Perhaps this pattern of findings indicates that children with stronger language skills profited more from the academic language supports because skill begets skill. Other research has demonstrated that children with higher initial vocabulary skill benefit more from language intervention (e.g., Cabell, Justice, Piasta, et al., 2011). This trend of greater benefits for children at the higher end of the distribution suggests additional support may be needed at school and/or home to ensure more children respond. Future studies should also consider if Tier 2 non-responders need Tier 3.

Nonetheless, the most intensive Tier 2 Enhanced aligned teacher-parent intervention approach appeared most promising overall. Such findings suggest that Tier 2 classroom instruction was more effective when paired with family coaching on strategies to responsively facilitate conversations, explain words, and share books at home. It is possible that Enhanced family engagement approach, that included modeling videos and feedback, helped parents improve naturalistic language interactions. Although further observational data is needed to confirm, this aligns with decades of research that show the effectiveness of increasing adults' use of language input that is contingent with children's talk/interest (see Greenwood, Schnitz, Carta, Wallisch, & Irvine, 2019).

In terms of implementation, over 80% of families elected to participate in Tier 2 coaching and adhered to the coaching cycle. This suggests the feasibility of remote coaching for use at large scale. The feasibility and tailored communications within coaching may be mechanisms that motivated parents. Although these findings with four parent coaching sessions align with meta-analytic findings (Bakermans-Kranenburg et al., 2003) that less than five parent coaching sessions can be as effective ($d = .42$) as five to 16 sessions ($d = .38$), further information is needed on the optimal number of sessions to see large, sustained effects. Moreover, it is unclear if schools have the capacity for practitioners to deliver parent coaching, rather than researchers.

Tier 1 effects with parent workshops/texts. Participation in the Tier 1 Only group showed promise for children's language outcomes, but only on one distal vocabulary measure (PPVT) and with a non-significant but meaningful effect size ($g = .39$). We observed interesting localized effects on the PPVT that were slightly larger at the lowest performing quantiles (.10, .20) and much stronger at highest performing quantiles (.80, .90). Quantile regression analyses also showed stronger Tier 1 Only effects at the low end of the distribution for the CELF sentence structure task. In terms of implementation, teachers implemented most aspects of Tier 1 classroom instruction at adequate levels, but across all conditions families participated in few of the Tier 1 offerings, opening only 25% of the modeling videos in text messages and attending only a quarter of the offered workshops. More information is needed on the conditions in which these lower-cost, frequently used ECE parent involvement approaches are motivating or effective.

It is also worth noting that it is difficult to explain the inconsistent effects on the PPVT across treatment groups (i.e., Tier 1 $g = .39$; Tier 2 Basic $g = .15$, Tier 2 Enhanced $g = .09$). Given that all children received Tier 1 support, we would expect more consistent effect sizes across all conditions if these universal intervention activities were changing this broad vocabulary skill. These inconsistencies may be, in part, because Tier 2 Basic children's PPVT scores were lower at pretest than the control group. Past research with tiered language interventions tends to show small effects on norm-referenced language measures like the PPVT with larger effects on proximal measures (Kelley, Goldstein, Spencer, & Sherman, 2015;

Loftus et al., 2010; Pullen et al., 2010; Zucker et al., 2013). Therefore, the presence of any promising findings on this distal vocabulary measure warrants further study in future more diverse and fully-powered samples.

Limited effects for Basic Tier 2. Tier 2 Basic supports – comprised of extra home activities, books, and text messages – did not increase outcomes in this Head Start sample. The Tier 2 Basic group even showed some non-significant but potentially meaningful negative trends for CELF expressive vocabulary and the inferential comprehension task ($g = -.20$ to $-.25$). Given that all Tier 2 children received the same classroom instruction and were randomized into two levels of family supports, these differences may be due to factors outside of the classroom, such as the way parents supported learning at home. Although low-income, linguistically diverse families report access to learning resources is a barrier (e.g., Marrapodi, 2016; Neuman & Knapczyk, 2018), these results and others (Anthony et al., 2014; Neuman, 2017) suggest there are negligible to small child benefits for providing books and home learning materials without more intensive parent education supports. Instead, parent education with opportunities for practice and feedback seem to be key mechanisms for supporting parents in ways that change child outcomes (Haring Biel et al., 2019). For example, simply receiving a book during a pediatric well child visit had little effect on family reading practices, whereas significant differences emerged when one book is given with short guidance from a pediatrician on the importance of reading with young children (Canfield et al., 2020). Likewise, Tier 2 Basic without tailored parent guidance simply was not enough to produce meaningful child impacts.

Overall, for language outcomes, the theory of change was generally supported, which asserted that school-based interventions with simple parent supports are not of sufficient intensity to narrow large vocabulary and achievement gaps for children eligible for Tier 2 and at risk for later reading difficulties (Chatterji, 2006; Wasik & Hindman, 2015). In contrast, the more tailored coaching support and feedback based was promising.

4.2. Impacts on child literacy outcomes

Unexpectedly, there was a negative pattern across all *Teaching Together* groups on literacy outcomes, one of which was statistically significant for the Tier 2 Enhanced group on the print knowledge measure (TOPEL, $p < .05$, $g = -.39$). The overall disappointing literacy findings may not be surprising when considering that literacy was a secondary intervention focus, comprising about 25% of activities at school and home. This imbalanced curriculum design was based on earlier iterative testing that showed only minor literacy instructional enhancements were needed for most classrooms. Yet the result was a relatively low dose of literacy support.

Baseline instruction in most of these Head Start classrooms featured a curriculum with a large amount of literacy instruction that included whole-group and small-group lessons that follow research-based principles (Schiller, 2013). When we asked teachers to implement a new curriculum supplement, this shift of instructional focus likely came at a cost given the finite amount of instructional time and teacher capacity. Several teachers reported replacing some of their core literacy curriculum with *Teaching Together*, which was not designed as a replacement of the core curriculum. Researchers should have given more guidance on how to maintain the core literacy curriculum while layering in a supplement. Teachers may also experience cognitive overload when asked to implement new curricula like this with two tiers of instructional support.

A critical analysis suggests another problem was likely the design of the Tier 1 classroom literacy components. That is, the approach to print referencing during shared reading only included one session per week, resulting in a quarter to half of the recommended dose of explicit print references used in past, efficacious pre-k research (Justice et al., 2009, 2015). These Tier 1 print references were also more highly scripted than the past studies by Justice and colleagues, likely resulting in more abbreviated discussion of print targets. In addition, parents in the Tier 2 Enhanced group only received one coaching session on literacy; this may not have been sufficient to fine-tune their scaffolding of literacy activities in responsive, developmentally appropriate ways (cf. DeBaryshe, Bueller, & Binder, 1996).

Not only were the literacy components a low dose, but amongst all lesson types, teachers implemented literacy components with the lowest average adherence to the planned lesson steps (e.g., 65% for phonological awareness, 72% for print knowledge). These literacy implementation levels are considered moderately low; inadequate implementation is often linked to null findings (Hill & Erickson, 2019), but these negative findings are troubling. Thus, we also think it is possible that print and letter instruction within Tier 2 lessons (e.g., magnetic letter matching to alphabet arcs, initial sound picture sorting) or home activities (e.g., alphabet arc placements, phonological awareness word play) was too advanced or somehow confusing to children who already demonstrated weak literacy skills. Future research with *Teaching Together* should consider higher intensity literacy supports or focus only on the language components.

4.3. Effects for DLLs

Understanding how to support pre-k DLLs is essential because this diverse population faces risk for academic difficulties, which can increase upon the transition to kindergarten without early intervention (Ansari & Crosnoe, 2018). Parents are a natural partner for enhancing DLLs' broad language skills (NASEM, 2017). Providing bilingual family engagement programs is increasingly required to meaningfully involve parents (Jeynes, 2012). The *Teaching Together* intervention was iteratively designed to serve English and Spanish speaking families because earlier phases of our work, as well as national data (NASEM, 2017), suggested the vast majority of DLLs speak Spanish at home. Yet in some parts of the U.S., preschool classrooms increasingly represent "superdiverse" settings where the majority of DLLs speak a variety of home languages (Baker & Páez, 2018). Sixty six percent of the full sample within this study were DLLs; about 54% of all children spoke Spanish at home and 13% spoke other home languages, such as Arabic, Urdu, and Creole. This presented challenges to serve a more linguistically diverse sample than planned; yet we did not have predetermined exclusionary criteria in our protocol that allowed us to target families who spoke only English or Spanish.

We found no evidence of differential effects of the intervention for DLL children's language outcomes. We interpret this as a promising outcome that suggests, like their peers, DLLs can learn academic language with aligned support from teachers and parents. However, we found one significant negative interaction in which DLLs in Tier 2 Basic group performed worse on the English letter naming fluency task relative to DLL peers in the control group. Perhaps the home literacy activities alone (without coaches) were less beneficial for these DLLs. Home letter activities were provided in English and Spanish; however, the program did not explain how these two languages share some graphophonemic properties, but are not equivalent. Thus, if parents of Spanish-speaking DLLs used the literacy materials without sufficient guidance to support biliteracy, this may have temporarily increased letter name/sound confusion for children assessed only in English. Future iterations of *Teaching Together* should consider local contexts and how to best support home literacy in superdiverse linguistic settings.

4.4. Limitations & future directions

Several limitations of this initial pilot study should be considered. First, this is an underpowered pilot of a new program that shows some promise for improving pre-k children's language skills, but not literacy skills within this sample of children, all of whom qualified for Head Start *and* demonstrated risk at screening. Additional study of the *Teaching Together* intervention is needed in more diverse samples with more heterogeneous child skills so that Tier 1 can be delivered to children meeting screening benchmarks. Future iterations of the program should remove or improve the literacy components as well as family components that did not reach desired implementation levels. For example, increased parent attendance at workshops and text messages usage are areas for improvements. We are separately examining relations between teacher and parent fidelity/quality of implementation and child outcomes. Another important limitation was that the psychometric properties of the proximal taught vocabulary measure were weak. This may indicate we did not measure a unidimensional construct or that there was too much variation in the vocabulary; these words were representative of a wide range of difficulty in past research (Dale & O'Rourke, 1981). A better proximal measure in future studies could more fully capture intervention effects on taught academic vocabulary. Also, despite past success using the inferential task in other pre-k samples (e.g., LARRC & Muijselaar, 2018), there were overall floor effects on the Inferential Comprehension measure for this sample.

Two important analytic limitations are worth discussing. First, two baseline measures presented with moderate effect size differences between the TT Basic and BAU conditions (PPVT $g = -.42$; CELF-WS $g = -.46$). However, we think our overall conclusions about the limited benefits of TT Basic are valid because negative effects were found for this group on other measures where baseline equivalence was established (CELF-EV, Inference Making) and because of low parent reported uptake of the Tier 2 Basic activities. Second, our multilevel models accounted for the nesting of children within classrooms but did not include classrooms within schools. We opted to use the two-level model to account for the primary nesting and random assignment mechanisms, both at the classroom-level, and to provide analytic symmetry between the LMM and LQMM approaches (i.e., LQMM only allows for two-level modeling). Future work might explore ways to improve LQMM in similar nested models (e.g., inclusion of extended random effects or inclusion of sandwich estimators).

This study design could not isolate the effects and costs of each component of the classroom and family programs. Future studies of the *Teaching Together* program could examine incremental effects of individual components in the classroom alone or classroom plus school for preschoolers eligible for Tier 2. If combined with cost effectiveness analyses, this could provide educators with important information to make decisions about what types of ECE-sponsored family engagement activities are the most valuable investments. Likewise, future studies should consider additional intensity of Tier 2 and/or Tier 3 for non-responders and more authentic teacher-parent partnerships and family engagement approaches.

5. Conclusion

These data suggest that *Teaching Together* holds promise as a school- and home-based intervention for increasing language skills amongst pre-k children at risk for later reading difficulties and eligible for Tier 2 services. Academic language is an essential skill for later reading and success in school. An aligned approach to explaining sophisticated vocabulary and engaging in extended, inferential conversation across the classroom and home could help close the vocabulary gap observed for DLLs and children from low-income backgrounds. Yet we found heterogeneity in response with greater vocabulary benefits mostly for children at the higher end of the distribution, suggesting future efforts should continue to examine for whom and under what conditions such academic language interventions are most beneficial.

Authors' contribution

Tricia A. Zucker: conceptualization; methodology; supervision; writing – original draft; funding acquisition. Sonia Q. Cabell: conceptualization; validation; writing – original draft; funding acquisition. Yaacov Petscher: conceptualization; methodology; software; writing – original draft. Heather Mui: resources; writing – review & editing; supervision. Susan H. Landry: project administration; writing – review & editing. Jamie Tock: software; writing – review & editing.

Acknowledgements

The research reported here was supported by a grant from the Institute of Education Sciences, U.S. (R305A150319). Opinions expressed do not represent the views of the U.S. Department of Education.We thank the children, teachers, and research staff that made this work possible. Special thanks to Maria Carlo, Hilary Kalin, Erin Jacobs, Azahara de la Vega, Susie Morales, Laura Fox, Stephanie Copp, Carrie Gomez, Nancy VanderLinden, Mayra Monreal, HyeJin Hwang, Cheryl McCallum, Gisela Trevino, and Emily Roden for their thoughtful work on this project. Research reported in this publication was supported by the Institute of Education Sciences (IES) under award number R305A150319.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at https://doi.org/10.1016/j.ecresq.2020.09.001.

References

Adams, A. M., Glenberg, A. M., & Restrepo, M. A. (2018). Moved by reading in a Spanish-speaking, dual language learner population. *Language, Speech, and Hearing Services in Schools*, *49*, 582–594.

Ansari, A., & Crosnoe, R. (2018). The transition into kindergarten for English language learners. In *Kindergarten transition and readiness*. pp. 185–204. Cham: Springer.

Anthony, J. L., Williams, J. M., Zhang, Z., Landry, S. H., & Dunkelberger, M. J. (2014). Experimental evaluation of the value added by Raising a Reader and supplemental parent training in shared reading. *Early Education and Development*, *25*(4), 493–514.

Baker, M., & Páez, M. (2018). *The language of the classroom: Dual language learners in Head Start, public pre-k, and private preschool programs*. Washington, DC: Migration Policy Institute.

Bakermans-Kranenburg, M. J., Van Ijzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, *129*(2), 195–215.

Ballantyne, K. G., Sanderman, A. R., & McLaughlin, N. (2008). *Dual language learners in the early years: Getting ready to succeed in school*. Washington, DC: National Clearinghouse for English Language Acquisition & Language Instruction Educational Programs. Available at http://www.ncela.gwu.edu/resabout/ecell/earlyyears.pdf

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Bierman, K. L., Heinrichs, B. S., Welsh, J. A., Nix, R. L., & Gest, S. D. (2017). Enriching preschool classrooms and home visits with evidence-based programming: Sustained benefits for low-income children. *Journal of Child Psychology and Psychiatry*, *58*(2), 129–137.

Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*(2), 294–304.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner, & W. Damon (Eds.), *Handbook of child psychology: Theoretical models of human development* (pp. 793–828). Hoboken, NJ, USA: John Wiley & Sons Inc.

Brown, A. L., & Lee, J. (2017). Evaluating the efficacy of children participating in Home Instruction for Parents of Preschool Youngsters and Head Start. *Journal of Early Childhood Research*, *15*(1), 61–72. http://dx.doi.org/10.1177/1476718X15577006

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

Buysse, V., Peisner-Feinberg, E., Soukakou, E., Fettig, A., Schaaf, J., & Burchinal, M. (2016). Using Recognition & Response (R&R) to improve children's language and literacy skills: Findings from two studies. *Early Childhood Research Quarterly*, *36*, 11–20.

Cabell, S. Q., Justice, L. M., Konold, T. R., & McGinty, A. S. (2011). Profiles of emergent literacy skills among preschool children who are at risk for academic difficulties. *Early Childhood Research Quarterly*, *26*(1), 1–14.

Cabell, S. Q., Justice, L. M., Piasta, S. B., Curenton, S. M., Wiggins, A., Turnbull, K. P., & Petscher, Y. (2011). The impact of teacher responsivity education on preschoolers' language and literacy skills. *American Journal of Speech-Language Pathology*, *27*(3), 387–400.

Cabell, S. Q., Zucker, T. A., DeCoster, J., Copp, S. B., & Landry, S. (2019). Impact of a parent text messaging program on pre-kindergarteners' literacy development. *AERA Open*, *5*(1), 2332858419833339.

Cain, K., & Oakhill, J. V. (1999). Inference making and its relation to comprehension failure. *Reading and Writing: An Interdisciplinary Journal*, *11*, 489–503.

Canfield, C. F., Seery, A., Weisleder, A., Workman, C., Cates, C. B., Roby, E., . . . & Mendelsohn, A. (2020). Encouraging parent–child book sharing: Potential additive benefits of literacy promotion in health care and the community. *Early Childhood Research Quarterly*, *50*, 221–229. http://dx.doi.org/10.1016/j.ecresq.2018.11.002

Carta, J. J., & Miller-Young, R. (Eds.). (2019). *Multi-tiered systems of support for young children: Driving change in early education*. Paul H. Brookes Publishing Company.

Catts, H. W., Petscher, Y., Schatschneider, C., Sittner Bridges, M., & Mendoza, K. (2009). Floor effects associated with universal screening and their impact on the early identification of reading disabilities. *Journal of Learning Disabilities*, *42*(2), 163–176.

Chatterji, M. (2006). Reading achievement gaps, correlates, and moderators of early reading achievement: Evidence from the Early Childhood Longitudinal Study (ECLS) kindergarten to first grade sample. *Journal of Educational Psychology*, *98*, 488–507. http://dx.doi.org/10.1037/0022-0663.98.3.489

Connor, C. M., Morrison, F. J., & Slominski, L. (2006). Preschool instruction and children's emergent literacy growth. *Journal of Educational Psychology*, *98*(4), 665–689. http://dx.doi.org/10.1037/0022-0663.98.4.665

Coyne, M. D., Kame'enui, E. J., Simmons, D. C., & Harn, B. A. (2004). Beginning reading intervention as inoculation or insulin: First-grade reading performance of strong responders to kindergarten intervention. *Journal of Learning Disabilities*, *37*(2), 90–104.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, *1*, 16–29. http://dx.doi.org/10.1037/1082-989X.1.1.16

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary: A national vocabulary inventory*. World Book-Childcraft International.

DeBaryshe, B. D., Bueller, M. J., & Binder, J. C. (1996). What a parent brings to the table: Young children writing with and without parental assistance. *Journal of Literacy Research*, *28*, 71–90.

Denton, C., & Hasbrouck, J. (2009). A description of instructional coaching and its relationship to consultation. *Journal of Educational and Psychological Consultation*, *19*(2), 150–175. http://dx.doi.org/10.1080/10474410802463296

Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language why language is central to reading development. *Educational Researcher*, *39*(4), 305–310.

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

Dulay, K. M., Cheung, S. K., Reyes, P., & McBride, C. (2019). Effects of parent coaching on Filipino children's numeracy, language, and literacy skills. *Journal of Educational Psychology*, *111*(4), 641–662. http://dx.doi.org/10.1037/edu0000315

Fitton, L., McIlraith, A. L., & Wood, C. L. (2018). Shared book reading interventions with English learners: A meta-analysis. *Review of Educational Research*, *88*(5), 712–751.

Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., . . . & Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016-4008)*. Washington, DC: National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. Retrieved from the NCEE website: http://whatworks.ed.gov.

Garibay, C. (2007). *Para Los Niños: Phase two evaluation*. Illinois: Garibay Group.

Gersten, R., Compton, D., Connor, C. M., Dimino, J., Santoro, L., Linan-Thompson, S., & Tilly, W. D. (2008). *Assisting students struggling with reading: Response to Intervention and multi-tier intervention for reading in the primary grades. A practice guide (NCEE 2009-4045)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, US Department of Education. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides.

Geraci, M. (2014). Linear quantile mixed models: The lqmm package for Laplace quantile regression. *Journal of Statistical Software*, *57*, 1–29. http://dx.doi.org/10.18637/jss.v057.i13

Greenwood, C. R., Schnitz, A. G., Carta, J. J., Wallisch, A., & Irvine, D. W. (2019). A systematic review of language intervention research with low-income families: A word gap prevention perspective. *Early Childhood Research Quarterly*, http://dx.doi.org/10.1016/j.ecresq.2019.04.001

Grindal, T., Bowne, J. B., Yoshikawa, H., Schindler, H. S., Duncan, G. J., Magnuson, K., & Shonkoff, J. P. (2016). The added impact of parenting education in early childhood education programs: A meta-analysis. *Children and Youth Services Review*, *70*, 238–249.

Haring Biel, C. H., Buzhardt, J., Brown, J. A., Romano, M. K., Lorio, C. M., Windsor, K. S., . . . & Goldstein, H. (2019). Language interventions taught to caregivers in homes and classrooms: A review of intervention and implementation fidelity. *Early Childhood Research Quarterly*.

Heidlage, J. K., Cunningham, J. E., Kaiser, A. P., Trivette, C. M., Barton, E. E., Frey, J. R., & Roberts, M. Y. (2019). The effects of parent-implemented language interventions on child linguistic outcomes: A meta-analysis. *Early Childhood Research Quarterly*, http://dx.doi.org/10.1016/j.ecresq.2018.12.006 (in press)

Hill, H. C., & Erickson, A. (2019). Using implementation fidelity to aid in interpreting program impacts: A brief review. *Educational Researcher*, *48*(9), 590–598. http://dx.doi.org/10.3102/0013189X19891436

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Developmental Psychology*, *49*(1), 4.

Jeynes, W. (2012). A meta-analysis of the efficacy of different types of parental involvement programs for urban students. *Urban Education*, *47*(4), 706–742.

Jordan, G. E., Snow, C. E., & Porche, M. V. (2000). Project EASE: The effect of a family literacy project on kindergarten students' early literacy skills. *Reading Research Quarterly*, *35*(4), 524–546.

Justice, L. M., Kaderavek, J. N., Fan, X., Sofka, A., & Hunt, A. (2009). Accelerating preschoolers' early literacy development through classroom-based teacher–child storybook reading and explicit print referencing. *Language, Speech, and Hearing Services in Schools*, *40*(1), 67–85.

Justice, L. M., Logan, J. A., Kaderavek, J. N., & Dynia, J. M. (2015). Print-focused read-alouds in early childhood special education programs. *Exceptional Children*, *81*(3), 292–311.

Justice, L. M., McGinty, A. S., Piasta, S. B., Kaderavek, J. N., & Fan, X. (2010). Print-focused read-alouds in preschool classrooms: Intervention effectiveness and moderators of child outcomes. *Language, Speech, and Hearing Services in Schools*, *41*(4), 504–520.

Kelley, E. S., Goldstein, H., Spencer, T. D., & Sherman, A. (2015). Effects of automated Tier 2 storybook intervention on vocabulary and comprehension learning in preschool children with limited oral language skills. *Early Childhood Research Quarterly*, *31*, 47–61.

Knight, D. S., Landry, S. H., Zucker, T., Williams, J. M., Merz, E. C., & Taylor, H. B. (2016). *Cost-effectiveness of early childhood interventions to enhance head start: Evidence from a randomized experiment*. Working Papers 2. https://digitalcommons.utep.edu/cerps_wp/2.

Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. *Econometrica*, *46*, 33–50. http://dx.doi.org/10.2307/1913643

Koutsoftas, A. D., Harmon, M. T., & Gray, S. (2009). The effect of tier 2 intervention for phonemic awareness in a response-to-intervention model in low-income preschool classrooms. *Language, Speech, and Hearing Services in Schools*, *40*(2), 116–130.

Kruse, L. G., Spencer, T. D., Olszewski, A., & Goldstein, H. (2015). Small groups, big gains: Efficacy of a tier 2 phonological awareness intervention with preschoolers with early literacy deficits. *American Journal of Speech-Language Pathology*, *24*(2), 189–205.

Language and Reading Research Consortium, & Muijselaar, M. M. (2018). The dimensionality of inference making: Are local and global inferences distinguishable? *Scientific Studies of Reading*, *22*(2), 117–136.

Landry, S. H., Assel, M., Williams, J., Zucker, T. A., Swank, P. R., Gunnewig, S., & Crawford, A. (2014). *The CIRCLE progress monitoring system technical manual*. Houston, TX: University of Texas Health Science Center at Houston.

Landry, S. H., Smith, K. E., Swank, P. R., Zucker, T., Crawford, A. D., & Solari, E. F. (2012). The effects of a responsive parenting intervention on parent–child interactions during shared book reading. *Developmental Psychology*, *48*(4), 969.

Landry, S. H., Zucker, T. A., Williams, J. M., Merz, E. C., Guttentag, C. L., & Taylor, H. B. (2017). Improving school readiness of high-risk preschoolers: Combining high quality instructional strategies with responsive training for teachers and parents. *Early Childhood Research Quarterly*, *40*, 38–51.

Loftus, S. M., Coyne, M. D., McCoach, D. B., Zipoli, R., & Pullen, P. C. (2010). Effects of a supplemental vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties. *Learning Disabilities Research & Practice*, *25*(3), 124–136.

Long, Y. (2012). *The impact of parental involvement on preschool children's later language development in low-income hispanic english language learners. Open Access Theses and Dissertations from the University of Nebraska College of Education and Human Sciences*. Paper 144. http://digitalcommons.unl.edu/cehsdiss/144.

Lonigan, C. J., Allan, N. P., & Lerner, M. D. (2011). Assessment of preschool early literacy skills: Linking children's educational needs with empirically supported instructional activities. *Psychology in the Schools*, *48*(5), 488–501.

Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of educational psychology*, *108*(1), 114–129.
Lonigan, C. J., Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (2007). *Test of preschool early literacy*. Austin, TX: Pro-Ed.
Mahoney, A. D., McConnell, S. R., Larson, A. L., Becklenberg, A., & Stapel-Wax, J. L. (2020). Where do we go from here? Examining pediatric and population-level interventions to improve child outcomes. *Early Childhood Research Quarterly*, *50*, 205–220.
Marrapodi, M. (2016). Transmedia meets the digital divide: adapting transmedia approaches to reach underserved Hispanic families. *Journal of Children and Media*, *10*(2), 276–284. http://dx.doi.org/10.1080/17482798.2016.1140492
McIntyre, L. L., & Phaneuf, L. K. (2008). A three-tier model of parent education in early childhood: Applying a problem-solving model. *Topics in Early Childhood Special Education*, *27*(4), 214–222.
Mendelsohn, A. L., Huberman, H. S., Berkule, S. B., Brockmeyer, C. A., Morrow, L. M., & Dreyer, B. P. (2011). Primary care strategies for promoting parent–child interactions and school readiness in at-risk families: The Bellevue Project for Early Language, Literacy, and Education Success. *Archives of Pediatrics & Adolescent Medicine*, *165*(1), 33–41.
Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research*, *79*(2), 979–1007.
Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. (2008). Added value of dialogic parent–child book readings: A meta-analysis. *Early Education and Development*, *19*(1), 7–26. http://dx.doi.org/10.1080/10409280701838603
National Academies of Sciences, Engineering, and Medicine. (2017). *Promoting the Educational Success of Children and Youth Learning English: Promising Futures*. Washington, DC: The National Academies Press. http://dx.doi.org/10.17226/24677
National Early Literacy Panel [NELP]. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
Neuman, S. B. (2017). The information book flood: Is additional exposure enough to support early literacy development? *The Elementary School Journal*, *118*(1), 1–27.
Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literacy Research*, *43*(2), 103–129.
Neuman, S. B., & Knapczyk, J. J. (2018). Reaching families where they are: Examining an innovative book distribution program. *Urban Education*, *55*(4), 542–569. http://dx.doi.org/10.1177/0042085918770722
Oakhill, J. V., & Cain, K. (2012). The precursors of reading ability in young readers: Evidence from a four-year longitudinal study. *Scientific Studies of Reading*, *16*(2), 91–121. http://dx.doi.org/10.1080/10888438.2010.529219
Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*(1), 23–33.
Petscher, Y. (2016). Do our means of inquiry match our intentions? *Frontiers in Psychology*, *7*, 1–6. http://dx.doi.org/10.3389/fpsyg.2016.01048
Petersen, D. B., & Spencer, T. D. (2012). The narrative language measures: Tools for language screening, progress monitoring, and intervention planning. *Perspectives on Language Learning and Education*, *19*(4), 119–129.
Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., . . . & Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children*, *77*(2), 161–183.
Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disability. *Learning Disabilities Research & Practice*, *25*(3), 110–123.
Reese, E., & Cox, A. (1999). Quality of adult book reading affects children's emergent literacy. *Developmental Psychology*, *35*(1), 20–28.
Roberts, T. A. (2008). Home storybook reading in primary or second language with preschool children: Evidence of equal effectiveness for second-language vocabulary acquisition. *Reading Research Quarterly*, *43*(2), 103–130.
Rogde, K., Melby-Lervåg, M., & Lervåg, A. (2016). Improving the general language skills of second-language learners in kindergarten: A randomized controlled trial. *Journal of Research on Educational Effectiveness*, *9*, 150–170. http://dx.doi.org/10.1080/19345747.2016.1171935
Sanders, M. R., Kirby, J. N., Tellegen, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychology Review*, *34*(4), 337–357.
Sénéchal, M., & Young, L. (2008). The effect of family literacy interventions on children's acquisition of reading from kindergarten to grade 3: A meta-analytic review. *Review of Educational Research*, *78*(4), 880–907.
Scheffner Hammer, C., 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*, 715–733. http://dx.doi.org/10.1016/j.ecresq.2014.05.008
Schiller, P. A. (2013). *Getting a jump on head start readiness: Frog street pre-k and closing the early achievement gap*. Retrieved from: http://www.frogstreet.com/wp-content/uploads/2015/03/Frog-Street-White-Paper-II-Final-Schiller-Patterson.pdf.
Solari, E. J., Denton, C. A., Petscher, Y., & Haring, C. (2018). Examining the effects and feasibility of a teacher-implemented Tier 1 and Tier 2 intervention in word reading, fluency, and comprehension. *Journal of Research on Educational Effectiveness*, *11*(2), 163–191. http://dx.doi.org/10.1080/19345747.2017.1375582
Suskind, D. L., Leffel, K. R., Graf, E., Hernandez, M. W., Gunderson, E. A., Sapolich, S. G., & Levine, S. C. (2016). A parent-directed language intervention for children of low socioeconomic status: A randomized controlled pilot study. *Journal of child language*, *43*(2), 366–406.
Swanson, E., Vaughn, S., Wanzek, J., Petscher, Y., Heckert, J., Cavanaugh, C., . . . & Tackett, K. (2011). A synthesis of read-aloud interventions on early reading outcomes among preschool through third graders at risk for reading difficulties. *Journal of Learning Disabilities*, *44*(3), 258–275.
U.S. Department of Health and Human Services. (2010). *The Head Start child development and early learning framework*. Washington, DC: Office of Head Start.
U.S. Department of Health and Human Services. (2016). *Head Start program facts fiscal year 2016*. Retrieved from https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/hs-program-fact-sheet-2016.pdf.
van den Broek, P., Kendeou, P., Lousberg, S., & Visser, G. (2017). Preparing for reading comprehension: Fostering text comprehension skills in preschool and early elementary school children. *International Electronic Journal of Elementary Education*, *4*(1), 259–268.
van Kleeck, A. (2008). Providing preschool foundations for later reading comprehension: The importance of and ideas for targeting inferencing in storybook-sharing interventions. *Psychology in the Schools*, *45*(7), 627–643. http://dx.doi.org/10.1002/pits.20314
van Kleeck, A., Vander Woude, J., & Hammett, L. (2006). Fostering literal and inferential language skills in Head Start preschoolers with language impairment using scripted book-sharing discussions. *American Journal of Speech-Language Pathology*, *15*(1), 85–95.
Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
Wanzek, J., Petscher, Y., Al Otaiba, S., Kent, S. C., Schatschneider, C., Haynes, M., . . . & Jones, F. G. (2016). Examining the average and local effects of a standardized treatment for fourth graders with reading difficulties. *Journal of Research on Educational Effectiveness*, *9*(Supp1), 45–66.
Wasik, B. A., & Hindman, A. H. (2015). Talk alone won't close the 30-million word gap. *Phi Delta Kappan*, *96*, 50–54. http://dx.doi.org/10.1177/0031721715575300
Wasik, B. A., & Hindman, A. H. (2018). Increasing preschoolers' vocabulary development through a streamlined teacher professional development intervention. *Early Childhood Research Quarterly*.
What Works Clearinghouse. (2019). *WWC standards brief: Baseline equivalence*. Washington, DC: Institute of Education Sciences.
White, A. (2019). Engaging families through observation and reflection: A collaborative approach to making learning visible. *Journal of Children and Poverty*, *25*(1), 45–56.
Whitehurst, G. J., & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, *69*(3), 848–872.
Wiig, E. H., Secord, W. A., & Semel, E. (2006). *Clinical evaluation of language fundamentals, second edition preschool (CELF preschool-2)*. San Antonio, TX: The Psychological Corporation.
Wilder, S. (2014). Effects of parental involvement on academic achievement: A meta-synthesis. *Educational Review*, *66*(3), 377–397.
Wong, K., Boben, M., & Thomas, M. C. (2018). *Disrupting the Early Learning Status Quo: Providence Talks as Innovative Policy in Diverse Urban Communities*.. Accessed on 9/15/20 from: http://www.providencetalks.org/wp-content/uploads/2018/07/updated-brown-eval.pdf
York, B. N., Loeb, S., & Doss, C. (2018). One step at a time: The effects of an early literacy text messaging program for parents of preschoolers. The Journal of Human Resources. *Advance Online Publication*, http://dx.doi.org/10.3368/jhr.54.3.0517-8756R
Zill, N., Resnick, G., Kim, K., O'Donnell, K., Sorongon, A., Ziv, Y., D'Elio, M. A., et al. (2000). *Head Start Performance Measures Center, Family and Child Experiences Survey (FACES): Technical Report*. Washington, DC: Office of Planning. Research and Evaluation, Administration for Children and Families, US Department of Health and Human Services. Accessed from https://www.acf.hhs.gov/opre/resource/head-sart-performance-measures-center-family-and-child-experiences.
Zucker, T. A., Carlo, M. C., Landry, S. H., Masood-Saleem, S. S., Williams, J. M., & Bhavsar, V. (2019). Iterative Design and Pilot Testing of the Developing Talkers Tiered Academic Language Curriculum for Pre-Kindergarten and Kindergarten., *12*(2), 274-306. *Journal of Research on Educational Effectiveness*, *12*(2), 274–306.
Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a brief tiered language intervention for pre-kindergartners at risk. *Early Education & Development*, *24*(3), 366–392.