HHS Public Access

Author manuscript

J Educ Psychol. Author manuscript; available in PMC 2017 January 01.

Published in final edited form as:

J Educ Psychol. 2016 January ; 108(1): 114–129. doi:10.1037/edu0000054.

Response to Instruction in Preschool: Results of Two Randomized Studies with Children At Significant Risk of Reading Difficulties

Christopher J. Lonigan and

Department of Psychology and the Florida Center for Reading Research, Florida State University.

Beth M. Phillips

Department of Educational Psychology and Learning Systems and the Florida Center for Reading Research, Florida State University.

Abstract

Although response-to-instruction (RTI) approaches have received increased attention, few studies have evaluated the potential impacts of RTI approaches with preschool populations. This manuscript presents results of two studies examining impacts of Tier II instruction with preschool children. Participating children were identified as substantially delayed in the acquisition of early literacy skills despite exposure to high-quality, evidence-based classroom instruction. Study 1 included 93 children (*M* age = 58.2 months; *SD* = 3.62) attending 12 Title I preschools. Study 2 included 184 children (*M* age = 58.2 months; *SD* = 3.38) attending 19 Title I preschools. The majority of children were Black/African American, and about 60% were male. In both studies, eligible children were randomized to receive either 11 weeks of need-aligned, small-group instruction or just Tier I. Tier II instruction in Study 1 included variations of activities for code- and language-focused domains with prior evidence of efficacy in non-RTI contexts. Tier II instruction in Study 2 included instructional activities narrower in scope, more intensive, and delivered to smaller groups of children. Impacts of Tier II instruction in Study 1 were minimal; however, there were significant and moderate-to-large impacts in Study 2. These results identify effective Tier II instruction but indicate that the context in which children are identified may alter the nature of Tier II instruction that is required. Children identified as eligible for Tier II in an RTI framework likely require more intensive and more narrowly focused instruction than do children at general risk of later academic difficulties.

Keywords

preschool; experiment; response-to-instruction; literacy

The acquisition of well-developed reading skills is a fundamental goal of children's early educational experiences. During the initial years of formal education, children need to learn the skills associated with both reading and writing because these skills are utilized later in

Correspondence regarding this article should be directed to Christopher J. Lonigan, Department of Psychology, Florida State University, 1107 W. Call Street, Tallahassee FL 32306-4301 (lonigan@psy.fsu.edu)..

the educational process both to transmit and to evaluate knowledge. Studies of children early in their educational experiences indicate substantial stability in children's literacy-related skills. Children who develop good literacy skills early are likely to be good readers at the end of elementary school, and children who have weak literacy skills early in elementary school are likely to remain poor readers (Duncan et al., 2007; Juel, 1988; Wagner, Torgesen, & Rashotte, 1994; Wagner et al., 1997). Due, in part, to findings of substantial stability of academic-related skills from early in children's formal educational experiences, there has been increasing recognition of the significance of the preschool years for the development of later academic skills, including reading. Studies of children prior to school entry have identified the skills that are predictive of later reading skills. These emergent literacy skills are conceptualized as the basic building blocks onto which children's later literacy skills are built (Whitehurst & Lonigan, 1998). Children with well-developed emergent literacy skills are those least likely to develop later reading problems.

Because of the increased recognition of the contributions of the preschool period for children's educational successes or difficulties, there is a growing body of research concerning the efficacy of different instructional activities during preschool. Current evidence indicates that intervention during the preschool period can help shape children's educational outcomes, particularly for children at-risk of later reading difficulties (e.g., Lonigan & Cunningham, 2013; Lonigan, Schatschneider, & Westberg, 2008). Although many of the instructional activities that have been evaluated are those that could be used in a Response to Instruction (RTI) framework (e.g., Lonigan, Purpura, Wilson, Walker, & Clancy-Menchetti, 2013), we know of no between-subjects studies that have evaluated the efficacy of these preschool interventions in the context of RTI. The two studies reported herein represent evaluations of preschool early literacy interventions using a RTI approach to identifying children at risk of reading difficulties.

Response to Instruction

Response to Instruction consists of two interrelated approaches. The first represents a method of identifying children who have a learning disability, consistent with the model outlined in the reauthorization of the Individuals with Disabilities Education Improvement Act (U.S. Department of Education, 2004). In this RTI approach, children who are substantially below expected levels of academic performance in an area (e.g., reading, spelling) and are making limited or no progress in acquiring the skill-set in the context of appropriate instruction are identified as having a learning disability (Gresham, 2002; Fuchs & Fuchs, 2006; Vaughn & Fuchs, 2003). This approach involves universal screening, identification of children at risk, application of increasing intensity of instruction, and frequent assessment of the child using progress monitoring instruments, such as curriculum-based measures (Gilbert, Compton, Fuchs, & Fuchs, 2012; Shinn, 2002).

The second RTI approach represents a framework for providing increasingly intensive instruction to meet the educational needs of children who are not making adequate progress. Each level of instructional intensity is often referred to as a tier. In the standard three-tiered instructional model (Berkeley, Bender, Peaster, & Saunders, 2009; Shapiro, 2009), Tier I represents the core classroom instruction. In the RTI framework, core classroom instruction

is supposed to represent high-quality, evidence-based general education. Tier II represents an increased intensity of instruction, often administered as small-group instruction, with membership in the small-groups differentiated by students' skill levels. This small-group instruction is typically conducted by the teacher in the classroom, but it can be delivered as a supplement provided by a paraprofessional or specialist. Tier III represents another increase in instructional intensity, often achieved by providing instruction in even smaller groups or one-onone. Tier III instruction can be part of the classroom instruction, a supplement instructional approach, or replace the core classroom instruction and is often provided by specialists such as reading teachers or special education teachers (Shapiro, 2009; Strangeman et al., 2006).

Core Classroom Curriculum in Preschool

Effective instruction at the level of the classroom curriculum is a core requirement of RTI. It is generally assumed that Tier I instruction will meet the needs of 75 to 90 percent of children within a grade level (Mellard, McKnight, & Woods, 2009; Shapiro, 2009). At least with respect to reading instruction, the presumption of a high-quality, evidence-based reading curriculum as Tier I instruction in elementary school classrooms is often met. In contrast, it is not clear that this presumption is met in many preschool classrooms (Greenwood et al., 2011; 2012). Although there is evidence that preschool can have positive effects on the reading-related skills of young children (e.g., Campbell, Ramey, Pungello, Sparling, & Miller-Johnson, 2002; Gormley, Phillips, & Gayer, 2008; Wong, Cook, Barnet, & Jung, 2008), there is both limited evidence for the efficacy of most preschool curricula (Camilli, Vargas, Ryan, & Barnett, 2010; Lonigan & Cunningham, 2013) and limited evidence that most preschool settings are implementing the select number of efficacious Tier I curricula. Regarding the efficacy of curricula, the U.S. Department of Education's What Works Clearinghouse's (WWC) reported that of more than 60 commercially available literacy or comprehensive preschool curricula, only 13 had studies that met evidence standards for causal interpretation; of these, only five curricula had evidence of a positive impact on one or more literacy-related skill. Regarding the typical quality of Tier I preschool settings, results from survey studies indicate that the two most commonly used curricula are the High/Scope Curriculum and The Creative Curriculum. These curricula are used in between 59 and 70 percent of classrooms serving children with significant risk for later academic difficulties (Jackson et al., 2007; U.S. Department of Health and Human Services, 2005); yet, neither of these curricula is rated by the WWC as efficacious.

Potentially Effective Tier II Instructional Practices for Preschool

Although there is limited evidence concerning the positive benefits of most early childhood curricula, there is a body of research that has identified effective instruction practices related to preschool early literacy skills. For example, a meta-analysis reported by the National Early Literacy Panel (NELP) demonstrated that small-group instruction designed to teach children phonological awareness skills resulted in increased phonological awareness and early reading skills, particularly when combined with instruction about letters and print (Lonigan et al., 2008). A second meta-analysis reported by the NELP demonstrated a positive impact of small-group shared-reading interventions, particularly shared reading that

was interactive, to promote children's oral language skills (Lonigan, Shanahan, & Cunningham, 2008). The WWC reported similar findings concerning phonological awareness and shared-reading interventions from studies limited to center-based practices with preschool-age participants.

Results of recent studies provide additional evidence of the effectiveness of these instructional practices and demonstrate that the positive impacts of instruction are domain specific. For example, Bowyer-Crane et al. (2008) randomized 152 4-year-old children who had been identified as having poor verbal and reasoning skills to two intervention groups that received 20 weeks of small-group and one-on-one instruction. Children in one intervention group received meaning-focused instruction that included training on vocabulary, comprehension, inference generation, and narrative skills, and children in the other intervention group received code-focused instruction that included training on letter-sound knowledge, phonological awareness, and book-level reading skills. Both at immediate posttest and a 5-month follow-up, children in the code-focused group scored significantly higher than did children in the meaning-focused group on measures of letter knowledge, spelling skills, and phonological awareness, and the children in the meaning-focused group scored significantly higher than did children in the code-focused group on grammar and for the vocabulary words that had been taught as a part of the intervention.

Lonigan et al. (2013) randomized 324 preschoolers from low-income backgrounds to combinations of code-focused and meaning-focused interventions or a business-as-usual control group. Children assigned to receive the meaning-focused intervention participated in 20 minutes of daily small-group dialogic reading intervention until the middle of the preschool year (approximately 12-14 weeks) and then 10 minutes a day for the rest of the preschool year (12 weeks). Children assigned to receive the code-focused interventions participated in 10 minutes of daily small-group instruction that focused on phonological awareness, letter-knowledge, or both from the middle to the end of the preschool year (12 weeks of intervention). By midyear, children in the meaning-focused groups had higher vocabulary scores than children who had not received the dialogic reading intervention, and at posttest, significant effects were restricted to measures of the skills that were the focus of the interventions that children received (e.g., children receiving the dialogic reading intervention outscored other children on vocabulary measures, children receiving the phonological awareness intervention outscored other children on phonological awareness measures).

Whereas these and similar studies (e.g., Farver, Lonigan, & Eppe, 2009; Fricke, Bowyer-Crane, Haley, Hulme, & Snowling, 2013; Gettinger & Stoiber, 2007) demonstrate that preschool children who receive focused interventions gain more skills than children who receive only the standard curriculum in place in their classrooms, the children in most of these studies were selected based on general risk for educational difficulties rather than a specific and measured risk as exemplified by the RTI approach of identifying children who are not making adequate progress with exposure solely to the classroom curriculum. Consequently, with the exception of the Bowyer-Crane et al. (2008) study, which did not include a business-as-usual control group, and the single-subject design study of phonological awareness instruction by Koutsoftos, Harmon, & Gray (2009), it is unknown if

children in studies examining effectiveness of intervention programs meet an identification standard for Tier II instruction. Moreover, in most cases, the curriculum in use in the classrooms from which children were drawn was unlikely to meet the RTI standard of a high-quality, evidence-based curriculum. Therefore, it is unclear the degree to which instructional practices with evidence of efficacy, which could be used as Tier II or Tier III instruction, would be effective for children who fail to make progress in the context of core classroom instruction that represents high-quality evidence-based general education. Children who fail to make adequate progress with core classroom instruction that meets the requirements of the RTI model may represent a population of preschoolers with more significant and more difficult-to-remediate weaknesses in academic skills. As a result, the nature and intensity of instruction that would be required to be effective in accomplishing this remediation goal for these particular children is largely unknown. This question motivated the current studies. To date, whereas other work has addressed benefits of Tier II preschool instruction in single-subject designs (e.g., Spencer et al., 2012), or has compared RTI framework implementation to traditional, non-tiered preschool instruction (Gettinger & Stoiber, 2012), no published study to our knowledge has addressed the specific focal question of the efficacy of Tier II preschool interventions in in the context of an effective Tier I core preschool curriculum.

Current Studies

Given the increased focus on preschool as a key developmental period for children's educational success and the increased acceptance of RTI approaches as a way to reduce the incidence of academic difficulties, the goals of the current studies were to evaluate the efficacy of Tier II instructional programs for children at risk of educational difficulties associated with reading in the context of an RTI model of identification. These two studies were conducted within a local school district's Title I preschool program. Prior to this project, many of the district's preschools had participated in a randomized trial of different preschool curricula as a part of the Preschool Curriculum Evaluation Research Consortium (PCERC; 2008). In that study, the combination of DLM Early Childhood Express and Open Court Pre-K curricula yielded significant positive results on children's language, reading, phonological awareness, and mathematics outcomes, with significant effect sizes ranging from .32 to .68, when compared to the curriculum that the district was using (High/Scope). Following the PCERC study, the district's superintendent required that all preschools in the district start using the combined DLM/Open Court curriculum. Consequently, the district's preschool curriculum met the RTI model's requirement of a high-quality, evidence-based Tier I curriculum.

For the current studies, children in the district's preschool were assessed after three months of exposure to the classroom curriculum to identify children who had significant weaknesses in their early literacy skills despite exposure to high-quality, evidence-based Tier I instruction. Identified children were randomized to either receive or not receive supplemental Tier II early-literacy interventions that were based on instructional activities shown to promote significant growth in children's language and code-based skills in prior randomized evaluations (e.g., Lonigan et al., 2013). Given prior evidence of the modularity of intervention outcomes, separate code-focused and language-focused strands of

intervention were developed to insure that children received direct attention to the one or both skill domains for which they demonstrated need. Once identified, children assigned to receive Tier II instruction participated in small-group language-focused instruction, code-focused instruction, or both types of instruction for 11 weeks, and children assigned to the comparison group received only their Tier I classroom-based instruction. Based on the RTI model and results of prior studies of supplemental instruction, it was hypothesized that children receiving Tier II instruction would end the year with substantially better early literacy skills than identified children who received only Tier I instruction.

STUDY 1

Methods

Participants—Children attending a Title I-funded prekindergarten program in the local school district were recruited for participation. Enrolled children included students with identified developmental delay in language and those with family backgrounds that indicated risk factors. Enrolled children also included some with diagnosed developmental disabilities. Exclusion criteria included frank sensory impairment (i.e., children with severely impaired visual or auditory abilities) and children with no expressive language ability. Children were qualified for the intervention based on the average of two scores in each skill domain of phonological awareness (i.e., Blending and Elision subtests of Test of Preschool Early Literacy [TOPEL]), print knowledge (i.e., Print Knowledge subtest of TOPEL and Test of Early Reading Ability), or oral language (i.e., Definitional Vocabulary subtest of TOPEL and Clinical Evaluation of Language Fundamentals-Preschool). Within domain, the average standard score had to be 90 or below (25th percentile) to qualify a child for randomization. All children who met this criterion in at least one of three domains (i.e., language, phonological awareness, print knowledge) were entered into the randomization, which occurred within school with students matched into pairs, to the extent possible, on qualifying domain and average standard scores on language measures (e.g.., children with adjacent average scores on language measures who qualified for code instruction were paired for randomization).

Across 14 schools, 167 children were recruited in the fall. Midyear assessments, which were used to qualify children for the Tier II instruction, were conducted in January after the children had been exposed to between three and four months of the Tier I curriculum in their classrooms. Of the 167 children recruited in the fall, seven attended two schools in which there were too few consented and qualified children to allow for randomization. Of the other 160 students, representing 12 schools, 93 qualified in at least one domain, and 67 did not qualify. The 93 qualifying participants included 58 boys (63%). Their average age at the midyear assessment was 58.2 months (*SD* = 3.62, range = 51-65 months). Most children were Black/African American (56%); 28 percent of the children were White, and the remainder of the children were Latino/Hispanic (5%), bi-racial (3%), Asian (1%), or unknown (7%).

Measures

Preschool Comprehensive Test of Phonological and Print Processing: (Pre-CTOPPP; Lonigan, Wagner, Torgesen, & Rashotte, 2002). Children were administered the Pre-CTOPPP, which was the development version of the TOPEL (Lonigan, Wagner, Torgesen, & Rashotte, 2007). The Pre-CTOPPP has excellent psychometric properties for 3- to 5-year-olds (i.e., αs = .86 to .96) and substantial evidence of validity (e.g., concurrent validity coefficients of .59 to .77 with other measures of similar constructs). Pre-CTOPPP Blending included 21 items that spanned the levels of complexity from compound word to phoneme levels and included both multiple choice items with pictures and free response items. Likewise, Pre-CTOPPP Elision included 18 items that also spanned the range of linguistic complexity and included both multiple choice and free response items. The Pre-CTOPPP Print Knowledge subtest included 36 items covering content including print concepts (e.g., "which one is a letter," "which can you read"), letter-name recognition, letter-sound recognition, and letter-name and letter-sound production. Pre-CTOPPP Receptive Vocabulary was designed as a receptive vocabulary measure in a common format with four pictures in a grid, where children had to point to the picture named. There were 40 items. Similarly, the Pre-CTOPPP Definitional Vocabulary measure included 40 items; these each included two parts, one in which the child had to label the single image or group of images shown and then respond to a follow-up question regarding function or relevant context for the item. These second questions enable the measure to tap a more definitional, depth of vocabulary dimension rather than simply confrontational naming. The maximum possible raw score was 80. All subtests included at least one practice item. All items were administered to all children without ceiling criteria.

Test of Early Reading Ability-Third Edition: (TERA-3; Reid, Hresko, & Hammill, 2001). Children received all three subtests of the TERA-3 (i.e., Alphabet, Conventions, and Meaning). Each subtest was administered according to established basal and ceiling rules. The composite Reading Quotient standard score was used for qualifying. The maximum possible combined raw score was 80. The TERA-3 has excellent psychometric properties, including coefficient alphas above .95 for all subtests (Reid et al., 2001).

Clinical Evaluation of Language Fundamentals–Preschool: (CELF-P; Wiig, Secord, & Semel, 1992). Five of six subtests from the CELF-P were administered to the children, including three receptive measures and two of three expressive language measures; the third was estimated from the scale scores on other measures such that the standard Expressive Language score could be calculated, along with the Receptive Language and the total Standard Score (this latter score was used in the qualifying process only). All subtests include basal and ceiling rules which were followed. Internal consistency reliability for children in this sample's age span ranged from .81-.96 for the composite scores, and scales have robust evidence of concurrent validity with other norm-referenced measures (Wiig et al., 1992).

Code and Language Intervention Post-test (CLIP): A curriculum-aligned assessment was developed to measure letter knowledge and oral language concepts targeted within these two intervention strands. The CLIP included expressive naming of all 26 letter names, and,

separately, naming of all 26 letter sounds, when shown individual cards presenting an upper-case letter. Both long and short vowel sounds were accepted as correct. On each separate assessment, if a student provided the letter name/sound when the other was requested, they were prompted with 'that is the name/sound of the letter, can you tell me the sound/name?', but all items were scored as correct or incorrect based on the requested label. Internal consistency reliabilities for both the letter-name (α = .96) and letter-sound (α = .95) assessments were high. For vocabulary, 12 expressive items including both naming items (e.g., 'what is this') and cloze items (e.g., 'Here are three —') represented a selection of the nouns and adjectives targeted during the intervention. We also used five items from the Preschool Language Scales to assess auditory comprehension and expressive language related to categories, object functions, and part/whole relations. The overall reliability of the CLIP language measure was adequate (α = .70), but the reliabilities for the two subscales was lower (CLIP Vocabulary α = .59, CLIP Language α = .60), likely due to the limited number of items on each subscale that were selected to map onto to specific vocabulary and language forms taught in the intervention.

Qualifying Children for the RTI Intervention—All child assessments were administered by research staff specifically trained to criterion performance on the measures used in this study through didactic presentations, modeling, a performance assessment of test administration to an adult posing as a child, and live observations of assessment with feedback. To facilitate the qualifying process, scores on the separately administered Pre-CTOPPP Blending and Elision subtests were converted into a joint standard score based on the scoring rules for the same items used in the TOPEL norm-referenced measure. Similarly, the Pre-CTOPPP Print Knowledge subtest was standard scored using the scoring rules (e.g., applying ceiling criteria) for the same items used in the TOPEL norm-referenced measure. This standard score was then averaged with the standard score from the TERA-3. In comparable fashion, the Definitional Vocabulary subtest was converted to its corresponding TOPEL standard score and averaged with the CELF-P Total Standard Score.

Tier I Instruction in Classrooms—All classrooms from which children for this study were recruited were part of the local school district's Title I program. Prior to this study, the district had switched its preschool curriculum to the combination of DLM Early Childhood Express and Open Court Pre-K as a result of a randomized evaluation that showed this curriculum combination to improve children's language, literacy, and mathematics scores more than the curriculum that had been in use in the district's preschools. Although the district mandated the switch to this curriculum combination, it had insufficient funds to supply all preschools with all materials for both curricula; therefore, for any preschool classroom participating in this study, we provided any missing materials for the classrooms. We had conducted the professional development for the prior curriculum evaluation study, and we also provided all of the district's preschool teachers moving to the new curriculum with professional development concerning the use of the combined curriculum. For this study, we did not collect fidelity of implementation data for teachers use of the combined curriculum; however, as part of a larger project in which this study was embedded, we conducted observations of all classrooms multiple times each year, and we simultaneously observed preschool classrooms that were not a part of the school district and that were not

using the combined curriculum. Details of these observations are reported in the on-line supplemental materials (see Appendix S1). Results of these observations indicated that there was a significantly higher frequency and/or quality of instructional activities occurring in the district's preschools than in the preschools that were not a part of the district.

Intervention Design and Implementation—The Tier II instruction was designed to include in-depth, teacher-directed, small-group instruction in each of three key emergent literacy strands of phonological awareness (PA), print knowledge (PK), and oral language, including both vocabulary and broader language skills. Daily activities for each instructional strand were described in detailed written lesson plans that included descriptions of the instructional activities for each day, the duration of the activity, and, when appropriate, scripts for the instructional activities. In the PA strand, instruction targeted rhyme matching, compound words, two-and three-syllable words, and onset-rime. In the latter four levels of linguistic complexity, both blending and elision activities were included. Materials included sets of picture cards, tubs of rhyming words, and ‘puzzle cards’ that divided pictures to represent the syllable or onset-rime breaks in words (e.g., a card picturing an elephant divided into three pieces to represent its three syllables). The intervention included explicit introductory and review activities designed to provide clear models and repeated, scaffolded practice of the task (e.g., blending syllables) being taught in each lesson. All lessons included use of pictures and objects to reduce demands on children's memories and to make the abstract phonological concepts more concrete (e.g., visually replacing the two separate images for component words with the single image representing a new compound word). These lessons were modified from those implemented with significant results in a prior small-group intervention (Lonigan et al., 2013).

For the PK strand, the intervention began with a focus on the first letters in participating children's names, and continued to cover the names of all 26 alphabet letters in a sequence generally designed to move from letters likely already known by some children (e.g., A, B) to letters less likely known (e.g., U, V). Letter-sound instruction occurred at the same time as letter-name instruction. Materials included letter cards of various sizes, upper- and lower-case magnet letters, tubs containing objects beginning with the letter sound, and sets of alliterative picture cards. Occasionally used materials in both strands included bingo cards, a board game, and word-family beach balls.

The oral language strand targeted vocabulary in semantic categories related to basic concepts appropriate for the preschool period (e.g., colors, size attributes, shapes, comparing and contrasting concrete attributes such as size and color, body parts), syntax such as plurals, present progressive, ‘wh-’ questions, and regular past tense, prepositions, and comprehension skills such as following two- and three- step directions, sequencing, and oral retelling of stories. Five themes were used to organize and teach the concepts. These themes included fairy tales, transportation, bugs, animals, and ocean/summer. Intervention activities in this strand incorporated theme-related storybooks, from which vocabulary targets were selected. Story-related materials also were used to teach language concepts (e.g., during the Fairy Tales theme the three bears were used to teach relative size concepts and prepositions).

A scope and sequence as well as exemplar instructional activities for each intervention strand is provided in the on-line supplemental materials (see Appendix S2), The intervention was 11 weeks and was conducted in small groups of 2 to 6 children four days per week. Fridays were used as make-up days for students absent one or more days earlier in the week. As described above, children qualified for the intervention based on midyear scores. Children qualified on PA, PK, and language independently. However, given their developmental overlap, children who qualified for either PA or PK received both strands, such that PK was taught for 20 minutes on Mondays and Wednesdays and PA for 20 minutes on Tuesdays and Thursdays, for 40 minutes of each strand per week. The language strand occurred Monday - Thursday for 20 minutes each day. Depending on how they qualified, children randomized to treatment received one or three strands, such that they received either 80 minutes of either code-focused or language-focused intervention or 160 minutes total of both code-focused and language-focused intervention per week. That is, if a child qualified on PK and language, he or she received 80 minutes of language intervention, and 40 minutes each of PK and PA per week. If a child qualified on only PA, he or she received 40 minutes of PA and 40 minutes of PK per week. A child who qualified only on language received just the 80 minutes of language intervention per week.

Interventionist Training and Fidelity Monitoring—Six interventionists, employed by the project, with education ranging from bachelor's to master's degrees taught the intervention groups. Interventionists received approximately six hours of initial professional development that included an overview of the intervention design, the rationale for the selected content skills, and modeling and practice of each intervention strand based on the written lesson plans. Intervention coordinators, who also were substantially responsible for developing the daily lesson plans, provided guidance and feedback on the lessons, and on how to differentiate instruction and pacing for individual groups and children within the structured lesson format. Interventionists received a 2-1/2 hour booster training partway through the intervention window. To insure that intervention strand was not confounded with interventionist, each interventionist taught both code-focused and language-focused groups.

To increase fidelity of the interventions, each daily session was audio-recorded, and one lesson per strand per interventionist per week was randomly selected and reviewed by an intervention coordinator, who then provided weekly feedback to the interventionists during the implementation. For each tape reviewed for fidelity, the intervention coordinator followed along with the printed lesson guide for that lesson. For all types of lessons, tapes were coded for eight adherence items and five quality items. For code-focused lessons, an additional seven adherence and three quality items were coded. All items were coded as present or absent, representing that 80% or more of the lesson included the item. Average adherence was high for the 78 language-focused lessons (M = .91, SD = .12; range: .38 - 1.0) and the 92 code-focused lessons (M = .90, SD = .13, range: .38 - 1.0) for which fidelity ratings were completed. Average quality ratings for language-focused (M = .96, SD = .12, range: .40 - 1.0) and code-focused (M = .94, SD = .13, range: .25 - 1.0) lessons were similarly high. For both adherence and quality ratings, there was a non-significant (ps > .11) trend for higher ratings across weeks of intervention. Because intervention coordinators

were the experts for the content and implementation of the lessons, audio recordings were not double-coded to compute inter-rater reliability; however, in mixed models in which intervention coordinator was treated as a random factor (i.e., interventionists nested within the intervention coordinator who coded the tape), intervention coordinator did not account for any reliable variance (*p*s > .38), indicating that the coding scheme was used consistently across intervention coordinators.

Results

Preliminary analyses—Based on their midyear scores, 93 children qualified for the RTI intervention based on their code-related skills (*n* = 30), language skills (*n* = 9), or both (*n* = 54). Of these children, 90 (97%) completed posttests following the intervention period. For children who qualified for the code-related instruction, 96% completed posttests (100% control, 93% intervention), and for children who qualified for the language instruction, 95% completed posttests (100% control, 91% intervention). There were no differences between control and intervention participants in terms of sex or race/ethnicity for those who qualified for the language instruction (*p*s > .59) or those who qualified for the code-related instruction (*p*s > .84). Standard score equivalents for the measures indicated that these children scored in the low-average to below-average range (i.e., average standard scores ranged from 78 to 86). Analysis of midyear scores by how children qualified for the RTI intervention (i.e., code-only, language-only, both) revealed expected patterns (see Table 1). In general, children who qualified based on language skills had lower scores on the language measures than did children who qualified based on code skills; children who qualified on code-skills had lower scores on the code-related measures than did children who qualified based on language skills; and children who qualified for both types of intervention had lower scores on both language and code-related measures than did children who qualified in only one domain.

The primary analysis of outcomes was conducted separately for children who qualified for the language instruction and for children who qualified for the code-related instruction on the relevant outcome measures (e.g., language measures for language intervention). Because children were nested within schools, all analyses were conducted as multi-level models in SPSS (Version 19) with school as a random factor and using restricted maximum likelihood estimation. Prior to conducting the primary analyses, two sets of preliminary analysis were conducted. First, for each intervention, we examined if the effects of the instruction were dependent on whether children qualified for just the target instruction or whether they qualified for both instruction types by including a term for how children qualified (i.e., just target skill vs. both target and other skill) and a treatment-by-qualification interaction term in the models. Second, for each outcome, we examined if the homogeneity of regression assumption was met by including treatment-by-covariate (i.e., midyear score on outcome, age) terms in the model.

In the first set of preliminary analyses, the effect of the code-related instruction was not moderated by whether children qualified for just the target or both instruction types for the code-related outcomes (all *p*s > .16). Consequently, these terms were not included in the models for the primary analyses of code-related outcomes. Because all but one child in the

completer sample assigned to the control group who qualified for the language instruction also qualified for the code instruction, the moderation analyses could not be conducted. Analyses that compared children who qualified for the language instruction only versus children who qualified for both language and code instruction to children in the control group did not reveal a consistent pattern, but these analyses indicated an effect of qualifying status on the Definitional Vocabulary subtest of the Pre-CTOPPP and the Expressive Language Subscale of the CELF. These effects are reported as a part of the primary analysis of language outcomes below. In the second set of preliminary analyses, children's scores at midyear on the outcome variable did not interact significantly with the treatment factor for any of the language outcomes (*p*s > .12). There was a marginally significant interaction for the Pre-CTOPPP Blending subtest ($p < .06$), but there were no significant interactions between midyear scores and treatment for any other code-related outcome (*p*s > .22). Age did not interact with treatment for any outcome (all *p*s > .15). Consequently, with the exception of the analysis of blending outcomes, these terms were not included in the models for the primary analysis.

Language-focused instruction—Results of the primary analysis for the children who qualified for the language instruction are shown in Table 2. Average raw scores, adjusted for age, on the language measures administered before the start of the intervention are shown on the left side of the table, and average raw scores at posttest, adjusted for age and the midyear measure on the same or most related outcome, are shown on the right side of the table. With the exception that children assigned to the control group scored marginally higher than did children assigned to the intervention group on the Expressive Language subscale of the CELF ($p < .09$), there were no statistically significant differences between children who received the language instruction and children who were assigned to the control group on scores pre-intervention (*p*s > .18), and children assigned to control and intervention groups did not differ in age ($p = .13$).

Children who received the language instruction scored marginally higher than did control group children on the Expressive Language subscale of the CELF at posttest, $F(1, 44.98) = 3.07$, $p < .09$; however, as noted above, whether children qualified for just language or both language and code instruction moderated this outcome. Children who qualified for both the language and code-related instruction (Adj. $M = 20.44$, $SD = 6.67$) scored significantly higher than did children in the control group (Adj. $M = 17.66$, $SD = 6.64$) at posttest on the Expressive Language subscale of the CELF, $t(44.17) = 2.50$, $p < .02$ (ES = .42), but there was no difference between children who qualified for just the language instruction (Adj. $M = 17.24$, $SD = 6.50$) and children in the control group, $t(47.37) = 0.28$, $p = .78$. Children's qualification status also affected outcomes on the Definitional Vocabulary subtest of the Pre-CTOPPP. Children who qualified for just the language instruction (Adj. $M = 34.41$, $SD = 9.60$) scored significantly lower than did children in the control group (Adj. $M = 45.59$, $SD = 13.49$) at posttest on this outcome, $t(53.02) = 2.79$, $p < .01$ (ES = -.87), but there was no difference between children who qualified for both the language and code-related instruction (Adj. $M = 45.34$, $SD = 13.71$) and children in the control group, $t(50.86) = -0.09$, $p = .93$. There were no other main effects of the language-focused instruction on the language outcomes (*p*s > .13). To investigate whether language-instruction group size affected

children's outcomes, multi-level regressions were conducted for children receiving the language instruction. In these models, children's posttest scores on the language measures were regressed on age, midyear scores on the same measure (or Receptive Vocabulary subtest of Pre-CTOPPP for the CLIP), and group size (i.e., number of children in a child's small group). Across outcomes, group size did not significantly predict posttest scores (all *p*s > .12).

Code-focused instruction—Results of the primary analysis for the children who qualified for the code-related instruction are shown in Table 3. Average raw scores, adjusted for age, on the code-related measures administered before the start of the intervention are shown on the left side of the table, and average raw scores at posttest, adjusted for age and the midyear measure on the same or most related outcome, are shown on the right side of the table. There were no statistically significant differences between children who received the code-related instruction and children who were assigned to the control group on scores pre-intervention (all *p*s > .27), but children assigned to control group were marginally younger than children assigned to the code-focused instruction group ($p = .08$).

At posttest, children who received the code-focused instruction scored marginally higher than did children in the BAU control group on the CLIP Letter Names test, $F(1, 77.00) = 3.52$, $p < .07$, but there were no significant differences between groups on any of the other code-related measures (*p*s > .18). Follow-up analyses of the interaction between the midyear score and instruction group for the Blending subtest of the Pre-CTOPPP revealed that for children 1 *SD* above the mean on the Blending subtest at midyear, children who received the code-focused instruction (Adj. $M = 14.59$) scored significantly lower ($p < .03$) than did children in the BAU control group (Adj. $M = 17.33$), $F(1, 76.00) = 5.16$, $p < .03$; however, there were no differences between intervention and control groups at the mean or 1 *SD* below the mean on the Blending subtest at midyear (*p*s > .24). As with the language instruction outcomes, multi-level regressions were used to investigate whether code-instruction group size affected children's outcomes by regressing children's posttest scores on the code-related measures on age, midyear scores on the same measure (or Pre-CTOPPP Print Knowledge subtest for the CLIP), and group size. Across outcomes, group size did not significantly predict posttest scores (all *p*s > .23).

Discussion

Results of Study 1 revealed the challenges of providing an RTI-based supplementary instruction model within classrooms providing relatively high-quality Tier I instruction to an at-risk group of preschool children. That is, the subset of children who qualified at the midyear assessment for the interventions were the lowest performing children among peers who also were, in general, below average in code skills, language skills, or both at the beginning of the year. As a result, the code-focused instruction we provided, which had been successful in improving PA and print-related skills for children enrolled in similar classrooms earlier (Lonigan et al., 2013) when the Tier I curriculum was less robust and had less of an instructional focus, was arguably much less successful under the new context and with a more targeted group of participants. Likewise, the language-focused instruction was not sufficiently robust to yield significant outcomes. Although the number of participants in

the study was modest, inspection of the effect sizes across outcomes does not suggest that low statistical power explains the absence of positive impacts of the Tier II instruction. These findings suggested that an efficacious intervention would need to be more intensive, more focused, and more explicit. All of these elements of the interventions were revised and refined for the Study 2 intervention protocols.

STUDY 2

Methods

Participants—Recruitment for Study 2 followed comparable procedures as in Study 1 and included those 14 schools plus an additional seven schools from within the same school district. In the fall, 277 children were recruited for participation. Midyear assessments, which were used to qualify children for the Tier II instruction, were conducted in January after the children had been exposed to between three and four months of the Tier I curriculum in their classrooms. Of the 277 children recruited in the fall, seven attended two schools where there were too few consented and qualified children for randomization at the school, and 14 other children were not available for midyear assessment. Of the remaining 256 children, 184 qualified in at least one domain and were randomized, and 72 students did not qualify. Qualifying participants included 109 boys (59%). Their average age at the midyear assessment was 58.2 months (*SD* = 3.38, range = 52-64 months). The majority of the qualifying children were Black/African American (66%); 28% of the children were White, and the remaining children were Latino/Hispanic (5%), Asian (1%), or unknown (8%).

Measures

Standardized language- and code-related measures: The same measures used in Study 1, with the exception of the CLIP, were used in Study 2. The qualifying procedures (i.e., using standard scores) and randomization procedures (i.e., matching students with adjacent average standard scores on language within qualification domain to the extent possible) described for Study 1 also were applied for the new cohort of participants.

CLIP-2: A revised version of the CLIP was administered at posttest. The letter knowledge measures were unchanged from Study 1. As in Study 1, internal consistency reliabilities for both the letter-name (α = .96) and letter-sound (α = .95) assessments were high, including the subset of letter names (α = .92) and letter sounds (α = .91) that were the targets of the intervention. Entirely new vocabulary and oral language items were developed to be aligned with the Year 2 content described below, including 15 expressive vocabulary items selected from the basic concepts vocabulary targets of the intervention (e.g., colors, shapes, body parts, animals) and 12 oral language items, all of which were receptive (e.g., “point to the building that is *taller*”). The overall reliability of the CLIP language measure was slightly less than adequate (α = .65); reliabilities for the two subscales were lower (CLIP Vocabulary α = .55, CLIP Language α = .43), likely due to the limited number of items on each subscale that were selected to map onto to specific vocabulary and language forms taught in the intervention.

Intervention Design and Implementation—Results from the first year of intervention suggested a need to intensify the intervention delivery. Changes were made in multiple aspects of the intervention content and implementation. Regarding content, we narrowed the focus of intervention within each strand to allow more time on topic across fewer distinct content targets. In the PK strand, we reduced the coverage of letter names and sounds from the full alphabet to the letters in each child's name and 11 target letters (A, U, B, M, P, T, D, N, F, J, G), which were selected both to represent the range of more and less likely to be known letters and to include two vowels. Each week of intervention included opportunities for cumulative review of all previously taught letters. In the PA strand, we limited the focus to compound words, two-syllable words, and onset-rime, with activities targeting both blending and elision at each level of linguistic complexity. The oral language and vocabulary strand underwent the most substantive changes. Whereas we retained the focus on vocabulary and basic concepts, we removed the focus on thematic storybook reading and instead used the basic concepts as the organizing principle for instruction of language and vocabulary targets (e.g., animal and body part names were taught both as vocabulary and to provide language contexts with which to teach verbs, prepositions, and spatial relations). Another substantive change was that lesson plans explicitly included upward and downward scaffolding instructions for interventionists so that within-group instruction could be better differentiated.

Daily activities for each instructional strand were described in detailed written lesson plans that included descriptions of the instructional activities for each day, the duration of the activity, and, when appropriate, scripts for the instructional activities. A scope and sequence as well as exemplar instructional activities for each intervention strand is provided in the on-line supplemental materials (see Appendix S3). As in Study 1, the intervention was 11 weeks, with sessions conducted Monday-Thursday and Friday make-up sessions. One element of our intensification was to reduce the maximum small-group size from 6 to 4 children. To further intensify the amount of need-matched intervention content children received, we separated the PA and PK intervention into two distinct implementation strands to allow students who qualified for either or both code-focused strands to receive a greater intensity of instruction in the appropriate skill areas. Children who qualified on both PA and PK received combined sessions of greater length. Specifically, any child that qualified on PA or PK but not both received four 15-minute sessions of the one strand area per week. In contrast, students who qualified on both PA and PK received four 20-minute sessions that included 10 minutes of each strand per week. Any child who qualified on language received four 20-minute language sessions per week, either exclusively or in addition to the single or dual code-focused sessions they received. Thus, weekly intervention doses ranged from 60 to 160 minutes per week, depending on whether a child qualified for one, two, or all three strands.

Another key change from Study 1 was the incorporation of regular progress monitoring assessments for the PA strand. These were brief, individually-administered curriculum-linked assessments used to determine differentiated pacing for each small group, at the several junctures in this strand where there were transitions in the scope and sequence of lessons (e.g., moving from compound-word to syllable manipulation). Groups in which

approximately 70% of the participating children achieved a passing score on the formative assessment progressed to the next content area in the scope and sequence, whereas groups demonstrating lesser mastery repeated the prior week's lessons once more before moving forward.

Interventionist Training and Fidelity Monitoring—Intervention groups were led by 12 female instructors employed by the project, including two who had led groups in Study 1. Initial interventionist training covered the same content but was lengthier than in Study 1, with additional booster training sessions, mid-implementation, for a total of approximately 20 hours of training. As in Study 1, each session was audiotaped, and randomly selected sessions per interventionist per week were listened to by intervention coordinators who then provided weekly feedback to each interventionist. Furthermore, each interventionist was observed live during the initial weeks of implementation and provided with individualized feedback regarding group management, pacing, and fidelity.

As in Study 1, for each tape reviewed for fidelity, the intervention coordinator followed along with the printed lesson guide for that lesson. For all types of lessons, tapes were coded for eight adherence items and five quality items. For code-focused lessons, an additional seven adherence and three quality items were coded. All items were coded as present or absent for each item, representing that 80% or more of the lesson included the item. Average adherence was high for the 201 language-focused lessons (M = .86, SD = .16; range: .38 - 1.0) and the 200 code-focused lessons (M = .87, SD = .14, range: .44 - 1.0) for which fidelity ratings were completed. Average quality ratings for language-focused (M = .91, SD = .15, range: .40 - 1.0) and code-focused (M = .91, SD = .13, range: .25 - 1.0) lessons also were high. Both adherence and quality ratings were constant across weeks of intervention, with no significant differences between or trends across weeks. Because intervention coordinators were the experts for the content and implementation of the lessons, audio tapes were not double-coded to compute inter-rater reliability; however, in mixed models in which intervention coordinator was treated as a random factor (i.e., interventionists nested within the intervention coordinator who coded the tape), intervention coordinator did not account for any reliable variance (ps > .30), indicating that the coding scheme was used consistently across intervention coordinators.

Results

Preliminary analyses—Based on their midyear scores, 184 children qualified for the RTI intervention based on their code-related skills (n = 42), language skills (n = 23), or both (n = 119). Of these children, 171 (93%) completed posttests following the intervention period. For children who qualified based on code-related skills, 92% completed posttests (90% control, 94% intervention), and for children who qualified for the language intervention, 93% completed posttests (92% control, 94% intervention). Of children who qualified for the code-related interventions and completed posttests, 44 qualified on the basis of PA scores only, 32 qualified on the basis of PK scores only, and 72 qualified on the basis of both PA and PK scores. There were no differences between control and intervention participants in terms of sex for those who qualified for the language instruction or those who qualified for the code-related instruction (ps > .32). There was a trend for a difference in race/ethnicity

between treatment and control groups for children who qualified for the language instruction (p = .06) and the code-related intervention (p = .09), which was the result of there being more children of Latino background in the control groups than in the intervention groups (7 vs. 1 and 5 vs. 1 in language-qualified and code-qualified children, respectively). Comparisons across control and treatment groups for children who were and were not non-Latino white, however, revealed no significant differences (ps > 14). Standard score equivalents for the measures indicated that the qualified children scored in the low-average to below-average range (i.e., average standard scores ranged from 80 to 86).

Analysis of midyear scores by how children qualified for the RTI intervention (i.e., code-only, language-only, both) revealed expected patterns (see Table 4). In general, children who qualified based language skills had lower scores on language measures than did children who qualified based on code skills; children who qualified on code-skills had lower scores on code-related measures that did children who qualified based on language skills; and children who qualified for both types of instruction had lower scores on both language and code-related measures than did children who qualified for only one type of instruction. Because children were nested within schools, all analyses were conducted as multi-level models in SPSS (Version 19) with school as a random factor and using restricted maximum likelihood estimation.

As with Study 1, the primary analysis of outcomes was conducted separately for children who qualified for the language instruction and for children who qualified for the code-related instruction on the relevant outcome measures (e.g., language measures for language instruction). Although children who qualified for the code-related instruction received just the code-related instruction for which they qualified (i.e., PA, PK, both), they were combined for the primary analyses on code-related outcomes to simplify the analyses. Follow-up analyses were used to examine if receiving one or two of the code-related intervention strands yielded different effects. We conducted two sets of preliminary analysis to determine if the effects of the interventions were dependent on whether children qualified for just the target instruction and to determine if the homogeneity of regression assumption was met. In the first set of preliminary analyses, the effect of the instruction was not moderated by whether children qualified for just the target or both instruction domains for any of the language outcomes (all ps > .30) or any of the code-related outcomes (all ps > .24). Consequently, these terms were not included in the models for the primary analyses. In the second set of preliminary analyses, children's scores at midyear on the outcome variable did not interact with the treatment factor for any of the code-related variables (p > .16). There was a marginally significant interaction for the Receptive Vocabulary subtest of the Pre-CTOPPP (p < .06), but there were no significant interactions between midyear scores and treatment for any other language outcome (ps > .23). Age did not interact with treatment for any outcomes (all ps > .12). Consequently, only the interaction between midyear scores and treatment group for the Receptive Vocabulary subtest of the Pre-CTOPPP was included in the models for the primary analysis.

Language-focused instruction—Results of the primary analysis for the children who qualified for the language instruction are shown in Table 5. Average raw scores, adjusted for age, on the language measures administered before the start of the instruction are shown on

the left side of the table, and average raw scores at posttest, adjusted for age and the midyear measure on the same or most related outcome, are shown on the right side of the table. There were no statistically significant differences between children who were assigned to the control or intervention groups on scores pre-intervention (all *p*s > .74), and children assigned to control and intervention groups did not differ in age ($p = .43$).

At posttest, children who received the language-focused instruction scored significantly higher than did children in the BAU control group on the Receptive Vocabulary subtest of the Pre-CTOPPP, $F(1, 124.00) = 4.79$, $p = .03$, as well as the targeted vocabulary, $F(1, 113.65)$, $p < .001$, and targeted language components, $F(1, 113.53) = 3.93$, $p = .05$, as measured by the CLIP-2. There were no significant differences between groups on any of the other language measures (all *p*s > .53). Follow-up analyses of the interaction between the midyear score and intervention group for the Receptive Vocabulary subtest of the Pre-CTOPPP revealed that for children 1 *SD* below the mean at midyear, children who received the language-focused instruction (Adj. $M = 28.21$) scored significantly higher (ES = .50, $p < .02$) than did children in the BAU control group (Adj. $M = 25.78$), but there was no difference in scores for children 1 *SD* above the mean at midyear who did (Adj. $M = 32.35$) or did not (Adj. $M = 32.75$) receive the language-focused instruction ($p = .70$). As in Study 1, multi-level regressions were conducted for children receiving the language instruction to investigate whether language-instruction group size affected children's outcomes. When posttest scores for children who received the language-focused instruction were regressed on age, midyear scores on the same measure (or measure of similar construct for the CLIP-2), and group size (i.e., number of children included in a child's small group), group size significantly predicted posttest scores on the Definitional Vocabulary subtest of the Pre-CTOPPP, $F(1, 47.68) = 6.86$, $p < .02$, but no other language measure (all *p*s > .25). For Definitional Vocabulary, larger group sizes for the instruction were associated with lower scores (unstandardized coefficient = -2.58).

Code-focused instruction—Results of the primary analysis for the children who qualified for the code-focused instruction are shown in Table 6. Average raw scores, adjusted for age, on the code-related measures administered before the start of the intervention are shown on the left side of the table, and average raw scores at posttest, adjusted for age and the midyear measure on the same or most related outcome, are shown on the right side of the table. There were no statistically significant differences between children who were assigned to the control or intervention groups on scores pre-intervention (all *p*s > .16), and children assigned to control and intervention groups did not differ in age ($p = .33$).

At posttest, children who received the code-focused instruction scored significantly higher than did children in the BAU control group on the Elision subtest of the Pre-CTOPPP, $F(1, 125.73) = 10.52$, $p = .002$, and the Print Knowledge subtest of the Pre-CTOPPP, $F(1, 131.88) = 4.19$, $p < .05$, as well as the targeted letter names, $F(1, 139.00) = 8.10$, $p = .005$, and targeted letter sounds, $F(1, 139.00) = 9.87$, $p = .002$, as measured by the CLIP-2. There were no significant differences between groups on any of the other code-related measures (all *p*s > .12). When posttest scores for children who received the code-focused instruction were regressed on age, midyear scores on the same measure (or measure of related construct

for the CLIP-2), and group size, group size did not significantly predict posttest scores (all *p*s > .15). Given the separated implementation of PA and PK lessons in Study 2, we also tested if the impacts of the code-focused interventions were moderated by whether children received one or both of the code-focused strands. Results indicated that none of the outcomes for which there was a significant main effect were moderated by this factor. However, there was a marginally significant moderation effect on Blending ($p = .09$). Specifically, there was a positive and marginally significant effect (ES = .41, $p = .06$) of the code-focused intervention for children who only qualified for one strand, whereas the effect for children who qualified for both code strands was not significant (ES = -.12, $p = .59$).

Discussion

Results in Study 2 were substantially stronger than in Study 1, supporting our initial hypotheses regarding the benefits of Tier II instruction. In Study 2, there were significant or marginally significant results for outcomes including phonological awareness, print knowledge, receptive vocabulary, and both receptive and expressive language. Outcomes were particularly strong for the targeted letters and focal language and vocabulary measures, for which the effect sizes were moderate to large. Effects were smaller for the more distal standardized measures of language- and code-related skills. Several potential explanations for the stronger observed effects of Study 2 include the smaller group size, more focused instruction, and more explicit guidance on differentiating instruction for the interventionists.

GENERAL DISCUSSION

The purpose of these studies was to evaluate the effectiveness of explicit, modular Tier II interventions for preschool children who were nonresponsive to high-quality, evidence-based Tier I classroom instruction. The results indicated that intensive Tier II instruction in small groups can have a significant impact on both the code-related and the language skills of preschool children who are at high risk of academic difficulties. Particularly in Study 2, the narrowly focused, highly differentiated instruction led to educationally meaningful, moderate-to-large improvements in proximal outcomes as well as effects on more distal outcomes on standardized assessments. The results of Study 1 revealed that intervention activities that had been shown previously to be effective with preschoolers at risk of academic difficulties generally were substantially less effective with preschoolers who continued to have low scores on key literacy-related skills despite exposure to high-quality, evidence-based classroom instruction. Additionally, a substantial percentage of children in these preschool classrooms were identified as eligible for Tier II instruction. These results have significant implications for implementing RTI models in preschool settings.

Robust evidence supports strong predictive relations between emergent literacy skills assessed in preschool and conventional literacy skills assessed in kindergarten and beyond (e.g., Furnes & Samuelsson, 2010; Lonigan & Shanahan, 2010). This evidence demonstrates that children with well-developed emergent literacy skills in preschool are those who learn to read sooner, and children who learn to read sooner develop greater mastery in both reading comprehension and general academic skills (e.g., Kaplan & Walpole, 2005; Sparks, Patton, & Murdoch, 2014). Conversely, children who have difficulty learning to read have substantial and sustained risk for continued problems in reading and general academic skills.

Given this evidence, the imperative to ameliorate the weaknesses in early skills of preschool students and to alter their learning trajectories is compelling. Results from the current studies indicate that explicit, small-group instruction can be an effective tool to support skill gains during the preschool period for children with significant risk of reading acquisition difficulties.

The current studies, which met criteria for RTI by identifying children for Tier II instruction from classrooms that were providing strong Tier I instruction, highlights the differences with results from prior studies in which children were selected from a group with high general risk but who were receiving Tier I instruction of lesser quality (e.g., Lonigan et al., 2013). The results of the current studies suggest that it is harder to remediate the skill deficits of children who continue to experience such weaknesses despite exposure to effective classroom practices than those of children with general risk for later academic difficulties. In large measure, this is likely because the children identified as qualifying for our Tier II interventions had lower skills, on average, than many who might be identified as qualifying in the context of a weaker Tier I environment. In studies of intervention with children at general risk of academic difficulties--often because of conditions associated with poverty, at least some of the children make gains as a result of the interventions because these interventions are these children's first sustained exposure to learning opportunities that promote literacy-related skills. In contrast, prior to being evaluated for inclusion in the Tier II subgroup, children in the two studies reported herein had exposure to between three and four months of an effective literacy-focused curriculum. Consequently, identified children's low scores were not a function of lack of exposure to the types of learning opportunities that promote literacy-related skills.

The improved results for Study 2 relative to Study 1 aligned with our supposition that to have a significant impact on students who are identified as having weak skill development even in the context of an evidence-based Tier I classroom curriculum, highly explicit and differentiated interventions must be provided. In Study 1, despite the fact that the interventions used were variations of ones that had previously proven effective for promoting the development of the language and code-related skills of children at general risk for later academic difficulties (e.g., Lonigan et al., 2013), there were minimal effects on the language and literacy skills of the children in the study. In contrast to the results of Study 1, statistically significant positive effects of the more focused and intensified interventions in Study 2 were observed both on all of the proximal assessments of intervention targets and on several standardized measures of the constructs that were the focus of the interventions.

The results of these studies also suggest that the intensity and duration of intervention for the particularly intransigent group of preschool children identified within an RTI model may need to be even greater than what we provided in Study 2, particularly in the language domain That is, the finding in Study 2 of only one significant impact on the standardized language measures suggests that a longer intervention period was warranted. Findings of substantially weaker impacts of interventions on standardized measures than those observed on proximal measures are common. In a meta-analysis of the effects of vocabulary interventions with preschool and kindergarten children, Marulis and Neuman (2010) reported that the average effect for experimenter-created measures of vocabulary was

significantly larger than the average effect for standardized measures of vocabulary. Overall, effect sizes were 40% smaller on standardized measures than they were on experimenter-created measures. Whereas experimenter-created measures may be more sensitive to intervention effects, such measures essentially demonstrate that children learned the content of what they were taught. In contrast, the identification of specific risk for later academic difficulties and the established linkage between vocabulary and later academic difficulties is based on standardized measures, representing the more general skill construct. Consequently, although it is important to demonstrate gains on measures of the content of an intervention, an actual reduction in risk for later academic difficulties likely depends on an intervention's capacity to affect children's scores on standardized measures, demonstrating generalized change in the underlying construct measured. Results of these studies suggest that achieving this goal likely will require a sustained instructional effort and should be the focus of future studies that explore and compare both intervention content and intensity.

More than half of the children in Study 1 and more than two-thirds of the children in Study 2 qualified for both code- and language-focused interventions. This result suggests two key ideas. First, it provides further support for the conclusion that code and language skills, even in the preschool period, are highly related (e.g., Storch & Whitehurst, 2002). Students weak in one area, even in the context of high-quality Tier I instruction, are also likely to be weak in the other area. Second, given that the modular code and language interventions were provided to children during the same intensive 4-month time period and analyses revealed no significant difference in effects for children who received just one versus both interventions, these results indicate that children can be supported in developing skills in multiple domains of emergent literacy at the same time without compromising the efficacy of the alternative module.

Implications for RTI in Preschool Settings

Overall, the results of these studies likely signifies that expected gains as a result of a Tier II intervention in preschool will be difficult to achieve, limited in scope, largest for proximal outcomes, and require intervention of longer duration with explicit focus on multiple targets to achieve broad effects on distal outcomes. In both Study 1 and even in Study 2 despite impacts on all targeted skill areas, average end-of-year standard scores for treatment and control students remained in the low-average to below-average range in both code-related and language domains (i.e., posttest standard scores = 84 - 91 across both studies). Consequently, even effective Tier II interventions do not result in complete “catch-up” or normalization of children's skills. Of course, the effect of intervention on children's skills from the current studies was the result of only 11 weeks of Tier II intervention; however, studies of substantially longer supplemental interventions with at-risk students also report positive effects that do not result in attainment of “average” scores (e.g., Mathes et al., 2005). Therefore, it is unlikely that an RTI approach that is limited in scope or duration will provide lasting benefit to preschool children who make limited progress despite exposure to high-quality, evidence-based Tier I instruction and effective Tier II interventions. Broad and sustained interventions are likely what will be needed. Such a conclusion also highlights the importance of providing Tier I instruction that is actually of high quality, and the need for substantial improvements in many preschool settings. These improvements likely will

require both the adoption of evidence-based curricula and high quality, equally robust professional development to support effective implementation.

Unlike the assumption that classroom Tier I instruction will meet the needs of 75% or more of the children in elementary school (Mellard et al., 2009; Shapiro, 2009), our results suggest that many children in preschools serving at risk populations will need additional instructional support. Despite the better than typical quality of the Tier I instruction provided in the preschools in our studies, more than half of participating children qualified for the interventions. In Study 1, 58% of consented and assessed students met the qualifying criteria; whereas in Study 2, a remarkable 72% did so. Greenwood et al. (2012) reported higher percentages of children they classified as Tier II and Tier III from Head Start, state-funded preschools, and Title I preschools than from tuition-based programs. Although the percentage of Tier II and III classified children was lower in their study than in our study, their Tier II identification cut-point was equal to about the 10th percentile on their screening measure.

High proportions of children eligible for Tier II instruction indicate that the preschool landscape for RTI may be quite distinct from what is typically found in later grades. That is, the eligibility criteria for students to Title I-funded preschool programs such as those in the current studies as well as to many state-funded preschool programs, Head Start, and federally subsidized child-care centers intentionally increase the likelihood that enrolled students will be from lower-SES family backgrounds and delayed in the acquisition of early skills relative to their more advantaged peers. The resulting concentrating of higher-need children in classrooms means that expectations for how many students might qualify for Tier II or Tier III instruction that are derived from studies with older student populations are likely to be substantially incorrect. This concentration of risk status in some types of preschools raises questions about preschool teachers' abilities to utilize an RTI model with identification criteria that identifies children on the basis of absolute risk (i.e., national norms). At a minimum, the efforts toward raising the quality of Tier I instruction in preschool settings are a prerequisite for widespread implementation of RTI models in preschool settings. Furthermore, future research is likely needed to explore the benefits and constraints of applying RTI models in these settings wherein children are identified for supplemental Tier II instruction based on relative risk (e.g., the bottom 15% of enrolled students). Alternative "reverse RTI" models in which all children in such settings initially receive small-group, explicit instruction typical of Tier II and then exit or receive less intense instruction based on progress monitoring results also are worth exploring. Additionally, future research is needed to better understand the characteristics of children who are unlikely to benefit substantially from Tier II instruction (i.e., non-responders), which may lead to the development of effective Tier II or Tier III instruction for these children.

Limitations

Although these studies had a number of strengths, including the identification of children eligible for Tier II instruction who were exposed to high-quality evidence-based classroom instruction, the use of multiple standardized measures to qualify children in each skill area

rather than relying on a broad screening measure that might have over-identified children (i.e., false positives from poor specificity; Gilbert et al., 2012; Johnson, Jenkins, & Petscher, 2010), and a randomized design that allowed unambiguous attribution of effects to Tier II instruction, there were a number of limitations worth noting. First, we could not estimate the effectiveness of the individual strands of Tier II because the majority of children qualified for both meaning-focused and code-focused strands. Second, although beneficial in screening-out false positives, delaying identification to mid-year also meant that the intervention period was restricted to the spring and that some children likely did not receive a sufficient dose of intervention (Compton et al., 2012; Gilbert et al., 2013; Harn, Linan-Thompson, & Roberts, 2008). Third, our studies were specifically focused on evaluating the impacts of Tier II interventions for preschool children exposed to high quality Tier I instruction, a context that includes most, but not all the elements of a full RTI model as it did not include the more regular progress monitoring and Tier III instruction typically seen in full RTI frameworks. Future investigation of our Tier II interventions from Study 2 in the full RTI context are warranted, alongside ongoing research on similar preschool RTI models such as the Recognition and Response Model (e.g., Coleman, Buysse, & Neitzel, 2006) and the Tier II and Tier III interventions developed by the Center for Response to Instruction in Early Childhood (e.g., Noe, Spencer, Kruse, & Goldstein, 2014; Spencer et al., 2012). Finally, the results of this study only address short-term gains as a result of effective Tier II instruction. Questions of the durability of effects beyond the end of the preschool year were not addressed.

Summary and Conclusions

Results of these studies demonstrate the potential for effective Tier II instruction in the context of an RTI model of identification for preschool children at substantial risk of reading difficulties. Notably, however, it appears that children who qualify for Tier II instruction in the context of high-quality, evidence-based Tier I instruction are likely to have more intractable delays in pre-academic skills than children who are identified only as having generally high risk for academic difficulties. Consequently, effective Tier II instruction needs to be intense and relatively narrow in scope. Moreover, even with effective Tier II instruction, the gap between preschool children who fail to make significant progress when exposed to high-quality, evidence-based classroom instruction and preschool children with typical development remains substantial. Additional research is needed that explores adaptations in the intensity, focus, and duration of Tier II instruction in the context of high-quality Tier I curricula. Finally, the results of this study and others suggest that the organization structure of preschools (i.e., density of children at risk of academic difficulties) results in a substantial proportion of children who qualify for Tier II and Tier III instruction. Hence, it may be particularly challenging to implement traditional RTI models of instruction in many preschools.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research and report was supported by grants from the Eunice Kennedy Schriver National Institute of Child Health and Human Development (HD052120 & HD060292). The views expressed herein are those of the authors and have not been reviewed or approved by the granting agency. We thank the children, families, and preschool program staff for their participation in this research, and we are indebted to Kylie S. Flynn, Amy Augustyn, and Lindsay Schmerbeck for their work on the development of the interventions used in this project as well as Marcy Wyatt and the many research assistants who worked on the project.

References

*included in supplemental material

Berkeley S, Bender WN, Peaster LG, Saunders L. Implementation of response to intervention: A snapshot of progress. Journal of Learning Disabilities. 2009; 42:85–95. [PubMed: 19103800]

Bowyer-Crane C, Snowling MJ, Duff FJ, Fieldsend E, Carroll JM, Miles J, Götz K, Hulme C. Improving early language and literacy skills: Differential effects of an oral language versus phonology with reading intervention. Journal of Child Psychology and Psychiatry. 2008; 49:422–432. [PubMed: 18081756]

Camilli G, Vargas S, Ryan S, Barnett WS. Meta-analysis of the effects of early education interventions on cognitive and social development. Teachers College Record. 2010; 112:579–620.

Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early childhood education: Young adult outcomes from the abecedarian project. Applied Developmental Science. 2002; 6:42–57.

Coleman, MR.; Buysse, V.; Neitzel, J. Full report. The University of North Carolina at Chapel Hill, FPG Child Development Institute; Chapel Hill: 2006. Recognition and response: An early intervening system for young children at risk for learning disabilities..

Coleman, MR.; Roth, FP.; West, T. Roadmap to pre-K RtI: Applying response to intervention in preschool settings. National Center for Learning Disabilities; New York, NY: 2009.

Compton DL, Gilbert JK, Jenkins JR, Fuchs D, Fuchs LS, Cho E, Bouton B. Accelerating chronically unresponsive children to tier 3 instruction: What level of data is necessary to ensure selection accuracy? Journal of Learning Disabilities. 2012; 45:204–216. [PubMed: 22491810]

Connor CM, Piasta SB, Fishman B, Glasney S, Schatschneider C, Crowe E, Morrison FJ. Individualizing student instruction precisely: Effects of child × instruction interactions on first graders' literacy development. Child Development. 2009; 80:77–100. [PubMed: 19236394]

Duncan GJ, Dowsett CJ, Claessens A, Magnuson K, Huston AC, Klebanov P, Japel C. School readiness and later achievement. Developmental Psychology. 2007; 43:1428–1446. [PubMed: 18020822]

Ehri LC, Nunes SR, Willows DM, Schuster BV, Yaghoub-Zadeh Z, Shanahan T. Phonemic awareness instruction helps children learn to read: Evidence from the national reading panel's meta-analysis. Reading Research Quarterly. 2001; 36:250–287.

Farver JM, Lonigan CJ, Eppe S. Effective early literacy skill development for young English language learners: An experimental study of two methods. Child Development. 2009; 80:703–719. [PubMed: 19489898]

Foorman BR, Torgesen J. Critical elements of classroom and small-group instruction promote reading success in all children. Learning Disabilities Research & Practice. 2001; 16:203–212.

Fricke S, Bowyer-Crane C, Haley AJ, Hulme C, Snowling MJ. Efficacy of language intervention in the early years. Journal of Child Psychology and Psychiatry. 2013; 54:280–290. [PubMed: 23176547]

Fuchs D, Fuchs LS. Introduction to response to intervention: What, why, and how valid is it? Reading Research Quarterly. 2006; 41:93–99.

Furnes B, Samuelsson S. Predicting reading and spelling difficulties in transparent and opaque orthographies: A comparison between Scandinavian and US/Australian children. Dyslexia: An International Journal of Research and Practice. 2010; 16:119–142.

Gettinger M, Stoiber KC. Applying a response-to-intervention model for early literacy development in low-income children. Topics in Early Childhood Special Education. 2007; 27:198–213.

Gettinger M, Stoiber KC. Curriculum-based early literacy assessment and differentiated instruction with high-risk preschoolers. Reading Psychology. 2012; 33:11–46.

Gilbert JK, Compton DL, Fuchs D, Fuchs LS, Bouton B, Barquero LA, Cho E. Efficacy of a first-grade responsiveness-to-intervention prevention model for struggling readers. Reading Research Quarterly. 2013; 48:135–154.

Gormley WT Jr. Phillips D, Gayer T. Preschool programs can boost school readiness. Science. 2008; 320:1723–1724. [PubMed: 18583597]

Greenwood CR, Carta JJ, Atwater J, Goldstein H, Kaminski R, McConnell S. Is a response to intervention (RTI) approach to preschool language and early literacy instruction needed? Topics in Early Childhood Special Education. 2012; 33:48–64. [PubMed: 24899769]

Greenwood CR, Bradfield T, Kaminski R, Linas M, Carta JJ, Nylander D. The response to intervention (RTI) approach in early childhood. Focus on Exceptional Children. 2011; 43:1–22.

Gresham, FM. Responsiveness to intervention: An alternative approach to the identification of learning disabilities.. In: Bradley, R.; Danielson, L.; Hallahan, DP., editors. Identification of learning disabilities: Response to treatment. Erlbaum; Mahwah, NJ: 2002. p. 467-519.

Gilbert JK, Compton DL, Fuchs D, Fuchs LS. Early screening for risk of reading disabilities: Recommendations for a four-step screening system. Assessment for Effective Intervention. 2012; 38:6–14. [PubMed: 24478613]

Harn BA, Linan-Thompson S, Roberts G. Intensifying instruction: Does additional instructional time make a difference for the most at-risk first graders? Journal of Learning Disabilities. 2008; 41:115–125. [PubMed: 18354932]

Jackson, R.; McCoy, A.; Pistorino, C.; Wilkinson, A.; Burghardt, J.; Clark, M.; Swank, P. Final Report to Congress. NCEE 2007-4007. National Center for Education Evaluation and Regional Assistance; 2007. National Evaluation of Early Reading First..

Johnson ES, Jenkins JR, Petscher Y. Improving the accuracy of a direct route screening process. Assessment for Effective Intervention. 2010; 35:131–140.

Juel C. Learning to read and write: A longitudinal study of 54 children from first through fourth grades. Journal of Educational Psychology. 1988; 80:437–447.

Kaplan D, Walpole S. A stage-sequential model of reading transitions: Evidence from the early childhood longitudinal study. Journal of Educational Psychology. 2005; 97:551–563.

Landry SH, Crawford A, Gunnewig S, Swank PR. The CIRCLE-Teacher Behavior Rating Scale. 2000 Unpublished research.

Lonigan, CJ.; Cunningham, AE. Significant differences: Identifying the evidence base for promoting children's early literacy skills in early childhood education.. In: Shanahan, T.; Lonigan, CJ., editors. Literacy in Preschool and Kindergarten Children: The National Early Literacy Panel and Beyond. Brookes; Baltimore, MD: 2013. p. 161-193.

Lonigan CJ, Purpura DJ, Wilson SB, Walker PM, Clancy-Menchetti J. Evaluating the components of an emergent literacy intervention for preschool children at risk for reading difficulties. Journal of Experimental Child Psychology. 2013; 114:111–130. [PubMed: 23073367]

Lonigan, CJ.; Schatschneider, C.; Westberg, L. Developing Early Literacy: Report of the National Early Literacy Panel. National Institute for Literacy; Washington, DC: 2008. Impact of code-focused interventions on young children's early literacy skills.; p. 107-151.

Lonigan CJ, Shanahan T. Developing early literacy skills: Things we know we know and things we know we don't know. Educational Researcher. 2010; 39:340–346. [PubMed: 22294802]

Lonigan, CJ.; Shanahan, T.; Cunningham, A. Developing Early Literacy: Report of the National Early Literacy Panel. National Institute for Literacy; Washington, DC: 2008. Impact of shared-reading interventions on young children's early literacy skills.; p. 153-171.

Lonigan, CJ.; Wagner, RK.; Torgesen, JK.; Rashotte, C. Preschool Comprehensive Test of Phonological and Print Processing. Authors; Tallahassee, FL: 2002.

Lonigan, CJ.; Wagner, RK.; Torgesen, JK.; Rashotte, C. Test of Preschool Early Literacy. ProEd; Austin, TX: 2007.

Marulis LM, Neuman SB. The effects of vocabulary intervention on young children's word learning: A meta-analysis. Review of Educational Research. 2010; 80:300–335.

Mathes PG, Denton CA, Fletcher JM, Anthony JL, Francis DJ, Schatschneider C. The effects of theoretically different instruction and student characteristics on the skills of struggling readers. Reading Research Quarterly. 2005; 40:148–182.

Mellard D, McKnight M, Woods K. Response to intervention screening and progress-monitoring practices in 41 local schools. Learning Disabilities Research & Practice. 2009; 24:186–195.

Noe S, Spencer TD, Kruse L, Gooldstein H. Effects of a tier 3 phonological awareness intervention on preschoolers; emergent literacy. Topics in Early Childhood Special Education. 2014; 34:27–39.

Preschool Curriculum Evaluation Research Consortium. Effects of preschool curriculum programs on school readiness (NCER 2008-2009). National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Washington, DC.; Washington, DC: 2008.

Reid, DK.; Hresko, WP.; Hammill, DD. Test of Early Reading Ability-Third Edition. Austin Texas Pro Ed.; 2001.

Shapiro, ES. Tiered instruction and intervention in a response-to-intervention model. RTI Action Network; 2009. Retrieved from http://www.rtinetwork.org/essential/tieredinstruction/tiered-instruction-and-intervention-rti-model

Shinn, MR. Best practices in school psychology IV. Vol. 1. Vol. 2. National Association of School Psychologists; Washington, DC.: 2002. Best practices in using curriculum-based measurement in a problem-solving model.; p. 671-697.

Sparks RL, Patton J, Murdoch A. Early reading success and its relationship to reading achievement and reading volume: Replication of ‘10 years later’. Reading and Writing. 2014; 27:189–211.

Spencer EJ, Goldstein H, Sherman A, Noe S, Tabbah R, Ziolkowski R,R, Schneider N. Effects of an automated vocabulary and comprehsnion intervention: An early efficacy study. Journal of Early Intervention. 2012; 34:195–221.

Storch SA, Whitehurst GJ. Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. Developmental Psychology. 2002; 18:934–947. [PubMed: 12428705]

Strangeman, N.; Hitchcock, C.; Hall, T.; Meo, G., et al. Response-to- instruction and universal design for learning: How might they intersect in the general education classroom?. 2006. Retrieved from http://www.ldonline.org/article/13002

United States Department of Education. Public Law 108-446. Individuals with disabilities education improvement act. 2004.

United States Department of Health and Human Services, Administration for Children and Families. Head Start Impact Study: First Year Findings. Author; Washington, DC: 2005.

Vaughn S, Fuchs LS. Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. Learning Disabilities Research & Practice. 2003; 18:137–146.

Wagner RK, Torgesen JK, Rashotte CA. Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. Developmental Psychology. 1994; 30:73–87.

Wagner RK, Torgesen JK, Rashotte CA, Hecht SA, Barker TA, Burgess SR, Donahue J, Garon T. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. Developmental Psychology. 1997; 33:468–479. [PubMed: 9149925]

Whitehurst GJ, Lonigan CJ. Child development and emergent literacy. Child Development. 1998; 69:848–872. [PubMed: 9680688]

Wilig, E.; Secord, W.; Semel, E. Clinical Evaluation of Language Fundamentals-Preschool. The Psychological Corporation; New York: 1992.

Wong VC, Cook TD, Barnett WS, Jung K. An effectiveness-based evaluation of five state pre-kindergarten programs. Journal of Policy Analysis and Management. 2008; 27:122–154.

Table 1

Pre-intervention scores by qualifying status for children identified as eligible for RTI intervention in Study 1

	Qualifying Domain						Effect Size for Comparison		
	Code (*n* = 30)		Language (*n* = 9)		Both (*n* = 54)				
Outcome Measure	**Adj.-M**	**(*SD*)**	**Adj.-M**	**(*SD*)**	**Adj.-M**	**(*SD*)**	**Code vs. Language**	**Code vs. Both**	**Language vs. Both**
Chronological Age (Months)	58.03	(3.95)	57.67	(3.97)	58.37	(3.42)	.09	−.09	−.20
Pre-CTOPPP Def. Vocabulary	56.10	(8.22)	48.92	(5.57)	37.11	(13.39)	.93^{+}	1.61***	.93**
Pre-CTOPPP Rec. Vocabulary	32.33	(4.57)	30.12	(4.25)	28.65	(5.69)	.49	.69***	.27
CELF-P Receptive Language	25.32	(5.22)	19.29	(6.44)	17.15	(5.11)	1.10^{+}	1.59***	.40
CELF-P Expressive Language	26.87	(4.36)	18.23	(5.28)	18.11	(5.97)	1.89***	1.61***	.02
Pre-CTOPPP Elision	11.28	(3.48)	8.72	(3.77)	8.04	(3.19)	.72^{+}	.98***	.21
Pre-CTOPPP Blending	15.64	(4.01)	15.17	(1.39)	13.24	(4.04)	.13	.60**	.51
Pre-CTOPPP Print	20.74	(8.91)	24.46	(8.12)	17.93	(7.53)	−.43	.35	.86*
TERA-3 Total Score	16.42	(6.82)	17.73	(4.13)	13.76	(5.75)	−.21	.43^{+}	.71

Notes. *N* = 93; Adj-*M* = raw score on measure adjusted for children's chronological ages; Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; CELF-P = Clinical Evaluation of Language Fundamentals-Preschool; Def. Vocabulary = Definitional Vocabulary; Rec. Vocabulary = Receptive Vocabulary; TERA-3 = Test of Early Reading Achievement, third edition.

$^{+}$ $p < .10$

* $p < .05$

** $p < .01$

*** $p < .001$.

Table 2

Pre- and post-intervention scores for completer children in control and intervention groups who received Tier II language intervention in Study 1

	Pretest					Posttest						
	Control Group		Intervention Group			Unstandardized Parameter for Covariates		Control Group		Intervention Group		
Outcome Measure	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**	**CA**	**Mid.**	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**
Chronological Age (Months)	57.59	(3.39)	58.94	(3.39)	.40	---	---	---	---	---	---	---
Pre-CTOPPP Def. Vocabulary	39.08	(10.32)	39.15	(16.04)	−.01	.22^{+}	.42***	45.75	(13.49)	42.41	(12.63)	−.25
Pre-CTOPPP Rec. Vocabulary	25.11	(4.86)	24.86	(5.96)	−.05	.15	.21	29.55	(5.25)	28.21	(4.90)	−.26
CELF-P Receptive Language	15.81	(5.20)	13.89	(5.30)	−.37	.08	.92***	16.57	(6.62)	18.65	(7.53)	.29
CELF-P Expressive Language	17.55	(5.58)	15.00	(5.24)	−.47^{+}	.32^{+}	.75***	17.65	(6.64)	19.45	(6.52)	.27^{+}
CLIP Vocabulary Assessment	---	---	---	---	---	−.14	.87***	3.58	(2.01)	3.41	(2.05)	−.08
CLIP Language Assessment	---	---	---	---	---	−.16	.63***	3.90	(1.30)	3.97	(1.00)	.06

Notes. $N = 60$ ($n = 29$ control, $n = 31$ intervention). Adj-*M* = raw score on measure adjusted for children's chronological ages for pretest and chronological ages and pre-intervention score on measure for posttest (for CLIP, scores on Receptive Vocabulary subtest used); Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; CELF-P = Clinical Evaluation of Language Fundamentals-Preschool; Def. Vocabulary = Definitional Vocabulary; Rec. Vocabulary = Receptive Vocabulary; CLIP = Code and Language Intervention Posttest.

* $p < .01$.

$^{+}$ $p < .10$

Table 3

Pre- and post-intervention scores for completer children in control and intervention groups who received Tier II code-related intervention in Study 1

	Pretest					Posttest						
	Control Group		Intervention Group			Unstandardized Parameter for Covariates		Control Group		Intervention Group		
Outcome Measure	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**	**CA**	**Mid.**	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**
Chronological Age (Months)	57.61	(3.47)	59.00	(3.58)	.39^{+}	---	---	---	---	---	---	---
Pre-CTOPPP Elision	7.51	(3.60)	6.68	(3.24)	−.24	.45	.58***	9.07	(3.95)	9.70	(4.40)	.15
Pre-CTOPPP Blending	10.57	(3.94)	11.33	(4.58)	.18	.04	.54***	14.75	(4.71)	13.69	(3.66)	−.25
Pre-CTOPPP Print	15.65	(8.25)	14.52	(8.51)	−.13	−.02	.89***	18.63	(9.57)	19.53	(8.95)	.10
TERA-3 Total Score	11.91	(6.65)	11.98	(6.42)	.01	.24	.83***	15.23	(7.52)	14.78	(7.11)	−.06
CLIP Letter Names	---	---	---	---	---	.02	.19***	12.66	(9.45)	14.89	(8.08)	.25^{+}
CLIP Letter Sounds	---	---	---	---	---	.08*	.08***	5.13	(6.54)	6.43	(7.25)	.19

Notes. $N = 81$ ($n = 43$ control, $n = 38$ intervention). Adj-*M* = raw score on measure adjusted for children's chronological ages for pretest and chronological ages and pre-intervention score on measure for posttest (for CLIP, scores on Print Knowledge subtest used); Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; TERA-3 = Test of Early Reading Achievement, third edition; CLIP = Code and Language Intervention Posttest.

$^{+}p < .10$

$^{*}p < .05$

$^{***}p < .001$.

Table 4

Pre-intervention scores by qualifying status for children identified as eligible for RTI intervention in Study 2

	Qualifying Domain						Effect Size for Comparison		
	Code (*n* = 42)		Language (*n* = 23)		Both (*n* = 119)				
Outcome Measure	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**Code vs. Language**	**Code vs. Both**	**Language vs. Both**
Chronological Age (Months)	57.88	(3.80)	58.48	(3.55)	58.01	(3.37)	−.16	−.04	.14
Pre-CTOPPP Def. Vocabulary	55.45	(7.07)	44.69	(8.37)	41.43	(10.35)	1.43***	1.46***	.32+
Pre-CTOPPP Rec. Vocabulary	31.51	(3.11)	30.03	(2.77)	27.00	(4.97)	.49	.99***	.65***
CELF-P Receptive Language	25.73	(5.79)	20.84	(4.26)	16.21	(5.97)	.92***	1.61***	.81***
CELF-P Expressive Language	26.21	(5.07)	21.97	(4.17)	18.10	(5.50)	.89**	1.50***	.73***
Pre-CTOPPP Elision	9.08	(3.02)	10.17	(3.60)	6.22	(2.51)	−.34	1.08***	1.46***
Pre-CTOPPP Blending	12.98	(3.99)	15.40	(1.66)	11.94	(3.74)	−.72**	.27+	.99***
Pre-CTOPPP Print	17.79	(8.50)	27.98	(4.92)	14.54	(8.07)	−1.37***	.40*	1.75***
TERA-3 Total Score	14.85	(10.00)	20.49	(8.49)	10.82	(5.63)	−.59**	.57***	1.57***

Notes. *N* = 184; Adj-M = raw score on measure adjusted for children's chronological ages; Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; CELF-P = Clinical Evaluation of Language Fundamentals-Preschool; Def. Vocabulary = Definitional Vocabulary; Rec. Vocabulary = Receptive Vocabulary; TERA-3 = Test of Early Reading Achievement, third edition.

+ $p < .10$

* $p < .05$

** $p < .01$

*** $p < .001$.

Table 5

Pre- and post-intervention scores for completer children in control and intervention groups who received Tier II language intervention in Study 2

	Pretest					Posttest						
	Control Group		Intervention Group			Unstandardized Parameter for Covariates		Control Group		Intervention Group		
Outcome Measure	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**	**CA**	**Mid.**	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**
Chronological Age (Months)	57.92	(3.56)	58.39	(3.24)	.14	---	---	---	---	---	---	---
Pre-CTOPPP Def. Vocabulary	41.96	(11.09)	41.78	(9.54)	−.02	−.02	.66***	47.90	(9.26)	47.54	(9.56)	−.04
Pre-CTOPPP Rec. Vocabulary	27.45	(4.91)	27.40	(4.71)	−.01	−.05	.43***	29.23	(5.48)	30.26	(4.19)	.21*
CELF-P Receptive Language	16.86	(6.05)	16.53	(5.61)	−.06	.10	.87***	19.19	(7.02)	19.68	(7.64)	.07
CELF-P Expressive Language	18.52	(5.48)	18.51	(5.44)	.00	−.12	.66***	20.09	(5.57)	20.07	(6.48)	.00
CLIP Vocabulary Assessment	---	---	---	---	---	−.05	.08***	8.93	(2.42)	10.27	(2.24)	.57***
CLIP Language Assessment	---	---	---	---	---	.06	.05**	8.13	(2.32)	8.79	(1.88)	.31*

Notes. $N = 132$ ($n = 66$ control, $n = 66$ intervention). Adj-*M* = raw score on measure adjusted for children's chronological ages for pretest and chronological ages and pre-intervention score on measure for posttest (for CLIP-2, scores on Receptive Vocabulary subtest used); CA = Chronological Age; Mid. = MidYear Score; Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; CELF-P = Clinical Evaluation of Language Fundamentals-Preschool; Def. Vocabulary = Definitional Vocabulary; Rec. Vocabulary = Receptive Vocabulary; CLIP = Code and Language Intervention Posttest.

* $p < .05$

** $p < .01$

*** $p < .001$.

Table 6

Pre- and post-intervention scores for completer children in control and intervention groups who received Tier II code-related intervention in Study 2

	Pretest					Posttest						
	Control Group		Intervention Group			Unstandardized Parameter for Covariates		Control Group		Intervention Group		
Outcome Measure	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**	**CA**	**Mid.**	**Adj.-*M***	**(*SD*)**	**Adj.-*M***	**(*SD*)**	**ES**
Chronological Age (Months)	57.88	(3.49)	58.43	(3.38)	.16	---	---	---	---	---	---	---
Pre-CTOPPP Elision	7.36	(2.94)	6.70	(2.94)	−.22	.36**	.47**	8.60	(3.77)	10.27	(3.68)	.45**
Pre-CTOPPP Blending	12.15	(4.02)	12.12	(3.64)	−.01	.04	.30**	13.26	(4.10)	13.81	(4.05)	.13
Pre-CTOPPP Print	15.54	(7.88)	14.96	(8.72)	−.07	.01	.83**	20.26	(8.79)	22.02	(8.33)	.21*
TERA-3 Total Score	11.56	(5.91)	12.08	(8.62)	.07	.27*	.88**	15.73	(6.90)	14.92	(9.77)	−.10
CLIP2 Letter Names	---	---	---	---	---	−.14	.74**	14.59	(8.54)	15.77	(7.83)	.14
CLIP2 Letter Sounds	---	---	---	---	---	−.20	.71**	9.74	(7.93)	11.09	(7.39)	.18
CLIP2 Targeted Letter Names	---	---	---	---	---	−.07	.30**	6.18	(3.92)	7.54	(3.50)	.37**
CLIP2 Targeted Letter Sounds	---	---	---	---	---	−.10	.31**	4.36	(3.80)	5.79	(3.55)	.39**

Notes. N = 148 (n = 75 control, n = 73 intervention). Adj-M = raw score on measure adjusted for children's chronological ages for pretest and chronological ages and pre-intervention score on measure for posttest (for CLIP2, scores on Print Knowledge subtest used); CA = Chronological Age; Mid. = MidYear Score; Pre-CTOPPP = Preschool Comprehensive Test of Phonological and Print Processing; TERA-3 = Test of Early Reading Achievement, third edition; CLIP2 = Code and Language Intervention Posttest.

* $p < .05$

** $p < .01$

** $p < .001$.