Journal of Early Intervention
2023, Vol. 45(3) 227–249

DOI: 10.1177/10538151221116596
journals.sagepub.com/home/jei
SAGE

Optimizing Vocabulary Instruction for Preschool Children

Keri M. Madsen[1], Lindsey A. Peters-Sanders[1], Elizabeth Spencer Kelley[2], R. Michael Barker[1,3], Yagmur Seven[1,4], Wendy L. Olsen[1,5], Xigrid Soto-Boykin[1,6], and Howard Goldstein[1]

Abstract

A cluster randomized design was used to investigate the effects of the *Story Friends* vocabulary curriculum on learning of 84 preschoolers in 24 classrooms who were at risk for language difficulties. Children in the treatment condition received explicit vocabulary instruction of 36 words during small-group storybook listening centers with extended practice opportunities in the classroom and at home. Children in the comparison condition were exposed to the same words in stories without explicit instruction or extended practice. Children (n = 16) with average or above language skills from six treatment classrooms were included to examine the impact of extended practice. Children at risk in treatment classrooms learned significantly more words (52%) than children at risk in comparison classrooms (12%). Children not at risk learned vocabulary equal to children at risk. A tiered approach to implementing *Story Friends* appears feasible for enhancing the vocabulary learning of preschoolers with and without language delays.

Keywords

vocabulary instruction, early childhood, preschool, language development

Vocabulary knowledge is strongly associated with academic performance (Language and Reading Research Consortium [LARRC] & Logan, 2017; Quinn et al., 2015). The size of a child's vocabulary at 2 years of age has uniquely predicted kindergarten readiness (Friend et al., 2018) as well as academic achievement upon kindergarten entry even after controlling for demographic variables such as socioeconomic status (SES), gender, and parent characteristics (Morgan et al., 2015). As children progress through school, their early vocabulary knowledge plays an increasingly important role in literacy development. Vocabulary is a primary indicator of growth in

[1]University of South Florida, Tampa, USA
[2]University of Missouri, Columbia, USA
[3]The National Institute of Occupational Safety and Health, Cincinnati, OH, USA
[4]The University of Kansas, Kansas City, USA
[5]University of Florida, Gainesville, USA
[6]Arizona State University, Tempe, USA

Corresponding Author:
Howard Goldstein, University of South Florida, 13301 Bruce B. Downs Blvd., MHC 1121, Tampa, FL 33612-3807, USA.
Email: hgoldstein@usf.edu

reading comprehension (Quinn et al., 2015; Suggate et al., 2018), contributing directly as well as indirectly through word reading skill (e.g., LARRC, 2015; Tunmer & Chapman, 2012).

Children enter kindergarten with wide variation in their early language experiences and vocabulary knowledge. Many children from low SES homes begin school with limited vocabulary knowledge relative to peers from higher SES homes (Hart & Risley, 1995; Nelson et al., 2011), although there is continued debate about the magnitude and characteristics of differences related to SES (cf. Golinkoff et al., 2019; Sperry et al., 2019). Maternal education level and aspects of maternal language input (i.e., number of different words, mean length of utterance, and number of wh-questions used by mothers with their children) have been implicated as potential predictors of language development in children prior to kindergarten (Vernon-Feagans et al., 2020). Regardless of the reason, children who begin school with limited vocabulary knowledge are at greater risk for developing language and reading disabilities that impact their academic performance (Catts et al., 2006; Elwér et al., 2013).

An obvious challenge for early childhood educators is to improve the vocabulary knowledge of young children at risk (AR) for or experiencing language delays. Despite the clear importance of vocabulary, early childhood teachers rarely teach vocabulary or provide support for vocabulary learning (Dwyer & Harbaugh, 2020; Greenwood et al., 2013; Piasta et al., 2015; Wright & Neuman, 2014). Pelatti et al. (2014) reported that only a small number of teachers (28 out of 81) provided any vocabulary learning opportunities (e.g., pausing during shared reading to discuss unknown words or asking children to define words or generate synonyms and antonyms), and these opportunities averaged only 0.27 min per day. One possible explanation for the lack of vocabulary instruction is that many preschool curricula offer little guidance for teachers on how to approach vocabulary instruction (Neuman & Dwyer, 2009). Even when training and instructional tools are provided, preschool teachers may not deliver language instruction with adequate fidelity (Dickinson, 2011; Justice, Mashburn, Hamre, & Pianta, 2008).

Story Friends (Goldstein & Kelley, 2016), a supplemental preschool curriculum, was developed for easy delivery of effective, explicit vocabulary instruction. This development took into consideration the unique context of early childhood classrooms including the characteristics of children served in these settings, education and training of classroom staff, and typical classroom routines. A focus on feasibility of high-fidelity implementation informed an instructional approach that could be adopted, maintained, and scaled up in a wide variety of classrooms. A summary of the theory and evidence that guided the development of *Story Friends* follows.

Characteristics of Effective Vocabulary Instruction

Carefully selected instructional targets. Beck et al. (2013) provide a framework for the selection of vocabulary targets that categorizes words into three broad tiers. Tier 1 includes frequently used words (e.g., *ball, kick, nice*) whereas Tier 3 words are less frequent words often related to specific domains (e.g., *evaporate, peninsula).* Beck et al. argue that the best choices for explicit instruction are Tier 2 words: Sophisticated vocabulary words that are high-utility and used more frequently by adults as well as in print. Many researchers have used this framework to select Tier 2 vocabulary targets for instruction (Coyne et al., 2009; Pullen et al., 2010; Silverman et al., 2013).

Explicit teaching. Comparisons of vocabulary learning from incidental exposure and explicit teaching using storybooks indicate a learning advantage for explicit teaching with moderate to large effects (Beck & McKeown, 2007; Coyne et al., 2007; Goldstein et al., 2016; Kelley et al., 2015, 2020; Loftus-Rattan et al., 2016). For example, in Goldstein et al. (2016), children who were incidentally exposed to vocabulary words in a storybook context demonstrated little vocabulary learning, defining an average of only one new word after 24 weeks of intervention. In contrast, children who were explicitly taught the same vocabulary words in brief lessons embedded in the storybooks had much larger gains, defining an average of 10 new words after 24 weeks. However, even

with explicit instruction, young children typically learn to define less than half of the targeted words (Goldstein et al., 2016; Loftus et al., 2010; Penno et al., 2002; Wasik et al., 2016).

Multiple learning and practice opportunities. To improve learning outcomes, researchers have sought to extend vocabulary instruction and practice beyond storybook reading (Beck & McKeown, 2007; Coyne et al., 2007, 2010; Loftus-Rattan et al., 2016; McKeown & Beck, 2014; Seven et al., 2020; Silverman et al., 2013). Extended instruction often involves activities designed to provide more exposure to words and meanings, more opportunities to practice, and more contexts in which to learn words beyond the storybook. Coyne and his colleagues (2007, 2010) provided lessons after each story that reviewed taught words and asked children to answer questions, identify examples and non-examples, and generate sentences. Across studies, children have consistently demonstrated greater learning of words taught through extended instruction. In addition, home-based intervention strategies have been effective in providing increased opportunities for practice beyond the school day that led to greater early language outcomes (Heidlage et al., 2020). For example, Soto et al. (2020) found that children learned more words when opportunities for review and practice were provided at home than when words were only taught at school.

Feasibility of high-fidelity implementation. Effective vocabulary instruction benefits from high-fidelity implementation. Implementation fidelity has been shown to relate to treatment effects of classroom-delivered oral language instruction (Dickinson, 2011; Justice, Mashburn, Pence, & Wiggins, 2008; Zucker et al., 2013). Yet, interventions that are effective when delivered in highly controlled settings may not easily transfer to real-world classroom and home settings (Hulleman & Cordray, 2009). Thus, it is important to consider the feasibility of implementation in intervention design, development, and measurement (Kaderavek & Justice, 2010). To help early childhood educators and families provide extended learning opportunities, researchers need to identify strategies that can be easily incorporated and sustained within typical school and home activities. One possibility is to provide a variety of strategies that can be implemented flexibly. Neugebauer et al. (2017) found that teachers' flexible use of extended definitions and comments about target words was related to students' vocabulary performance above and beyond their use of scripted lessons, even when teachers were not provided with specific activities or a schedule to follow.

Individual differences in vocabulary learning. To address individual differences in response to vocabulary intervention, tiered approaches have been considered in which children identified as AR receive more intense intervention than children with age-appropriate vocabulary (Coyne et al., 2019; Loftus et al., 2010). For example, Pullen et al. (2010) identified first graders as AR and not at risk (NAR) based on pre-intervention vocabulary scores. All children received a first tier of instruction and children AR received a second tier of small-group intervention. Although children AR made significant gains at post-test, scores were still below the children NAR who received only the first tier of instruction. Although such tiered approaches hold promise for improving outcomes for children AR, additional research is needed to optimize effective intervention (e.g., duration, intensity, practice, and review).

Story Friends Development

The *Story Friends* program was created through an iterative development process whereby data from preliminary studies and feedback from key stakeholders were used to revise and refine the design, delivery, and measurement strategies (see Kelley & Goldstein, 2014). The program delivers explicit vocabulary instruction embedded in recorded narration of children's storybooks. Children listen using headphones and follow along in their own books during small-group listening centers. The program includes two series, *Forest Friends and Jungle Friends*, each with four

recurrent characters and written around events relatable to preschool children (e.g., attending the first day of school, making new friends). Each series includes an introductory book, nine instructional books, and three review books. Children listen to each book three times, thereby increasing their exposure to the word lessons and the number of opportunities to respond. The embedded lessons follow a carefully designed instructional format based on the principles of robust vocabulary instruction (Beck et al., 2013) and provide explicit instruction including a child-friendly definition, examples, and multiple opportunities for children to respond (e.g., repeat the definition, use a gesture). The teacher or paraeducator facilitates the listening center, helping children keep their place in their book and encouraging them to respond.

Target vocabulary words were selected to be relevant for future reading and academic success. Similar to many studies of vocabulary instruction within the context of storybooks (e.g., Coyne et al., 2009; Pullen et al., 2010; Silverman et al., 2013), the approach for selecting words was aligned with the recommendations for Tier 2 word selection (Beck et al., 2013). First, the selected words are likely to be new and more sophisticated words for concepts already familiar to many preschoolers (e.g., *enormous, ridiculous, construct*). Second, the words are likely to have utility across multiple contexts. Third, words were selected that could be taught effectively through recorded, explicit instruction embedded within the context of the stories. The research team selected words using a collaborative and iterative process; information from children's learning, teacher feedback, and observations were used to revise the target words.

In previous studies of *Story Friends*, children in treatment classrooms learned significantly more words than children in comparison classrooms with large treatment effects (Goldstein et al., 2016, 2017; Kelley et al., 2015). These studies provided evidence of efficacy and feasibility of high-fidelity implementation, but also identified opportunities to improve student learning. Thus, revisions resulted in several substantial changes. First, two additional words per book were selected, doubling the number of instructional targets to 36 words per series (Peters-Sanders et al., 2020). Second, additional components were developed to promote review and practice of the target vocabulary beyond the small-group listening center: (a) Classwide extension materials were designed to promote opportunities for practice during the school day (see Seven et al., 2020) and (b) home extension materials were designed to encourage parents or caregivers to review the words with children at home (see Soto et al., 2020). Third, adding classwide and home review not only extended instruction and practice beyond the listening center for children receiving the small-group intervention but also created opportunities for other children in the classroom to learn target words and meanings.

A randomized control trial of the revised *Forest Friends* series resulted in large treatment effects on learning of target vocabulary words in preschoolers who were AR (Kelley et al., 2020). Social validity measures allowed teachers and parents to provide valuable input that informed additional revisions to the classroom and home review and practice materials. The current study is the first randomized group design investigating the revision of *Jungle Friends* and its implementation. Teachers in participating classrooms were asked to implement revised classroom and home review and practice procedures, and we assessed the acceptability and feasibility of these changes. We also examined the effects of the revised program on children who were NAR and did not receive small-group instruction but were exposed to target vocabulary via classwide and/or home review of the words. The performance of these children was of interest because word learning in this group would provide evidence of classwide review and practice of words beyond the listening center, which was difficult to measure directly.

Purpose

This study sought to evaluate the efficacy of the *Jungle Friends* series of *Story Friends* and to gather information about acceptability and feasibility of classwide and home learning

opportunities. Like previous studies, we were interested in the effects of this revised intervention on children who were AR. However, because the classwide and home extensions were provided to all children in the classroom, we also sought to examine the effects on children NAR (i.e., who did not receive small-group instruction). This study is an effectiveness-implementation Hybrid Type-1 study (Curran et al., 2012) with the goal of accelerating the translation of research findings into practice. The following research questions were addressed:

Research Question 1 (RQ1): What are the effects of *Jungle Friends* embedded vocabulary instruction plus classroom and home extension activities versus exposure to the same books without vocabulary instruction on the learning of target vocabulary words by preschool children who are AR for language and learning difficulties? Are intervention group effects moderated by classroom or child characteristics (e.g., attendance and language status)?
Research Question 2 (RQ2): To what extent do children AR for language and learning difficulties retain their knowledge of target vocabulary at 2 months follow-up?
Research Question 3 (RQ3): How do baseline knowledge and learning of challenging vocabulary words demonstrated by children NAR compare to children AR in treatment classrooms?
Research Question 4 (RQ4): What are the teachers' and parents' perceptions of the acceptability and feasibility of implementing the various components of the intervention?

Method

Design

A cluster-randomized design was used to address RQ1, RQ2, and RQ4. Teachers and children in 24 classrooms were randomly assigned to treatment and comparison conditions. Sixteen teachers had participated in the study of the *Forest Friends* book series during the prior school year; of these, 7 teachers who had been randomly assigned to the treatment condition the year prior remained in the treatment condition for the current study. To maintain equal group sizes, classrooms of the other 9 returning teachers and 8 newly-recruited teachers were randomly assigned to 5 treatment and 12 comparison openings.

The treatment condition entailed 13 weeks of *Jungle Friends* stories with vocabulary instruction. Children AR in these classrooms received small-group explicit instruction of 36 vocabulary words embedded in pre-recorded storybooks. In addition, classwide and home extension materials were provided to treatment teachers and caregivers to promote practice with target vocabulary beyond the listening center with all children in the classroom (i.e., children AR and NAR).

The comparison condition also entailed 13 weeks of *Jungle Friends* stories, but without vocabulary instruction. Children in these classrooms were exposed to the target words in the same stories, but did not receive explicit instruction or practice target vocabulary. In lieu of the vocabulary focus, comparison teachers were provided with brief lessons and home extension materials related to story comprehension.

To address RQ3, a sample of children identified as NAR were selected from six treatment classrooms to examine their knowledge of target words before and after the classwide and home extensions were used. Children AR from the same classrooms who received small-group instruction served as a comparison for this exploratory analysis.

Participants and Setting

Participants were 84 children identified as AR and 16 children identified as NAR from participating classrooms. In addition, 22 teachers and 5 paraeducators participated in the study. Classrooms

were early childhood programs from both private and public school settings. In Missouri, classrooms comprised a combination of Head Start and Title I programs. In Florida, classrooms were a combination of School Readiness (a statewide program offering financial assistance to low-income families for early childhood education), Florida Voluntary Prekindergarten (a free program for 4- to 5-year-old children), and Title I. These 24 classrooms were expected to serve a large proportion of children AR for language and literacy difficulties based on socioeconomic and demographic factors.

All children in these classrooms were invited to take part in the study. Children who received parental consent to participate were assessed in the fall by research staff with two norm-referenced measures: the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007), which is a measure of receptive vocabulary, and the Core Language Score subtests of the Clinical Evaluation of Language Fundamentals—Preschool (CELF P-2, Wiig et al., 2004), a comprehensive measure of receptive and expressive language. Three or four children from each classroom were selected to form the AR sample that received the small-group listening center instruction. Based on prior studies of *Story Friends*, children were selected for the AR group if they had standard scores between .5 and 2 standard deviations below the normative mean (i.e., 92-70) on either the PPVT-4 or CELF P-2. In classrooms with more than four children scoring within this range, we asked teachers to help us determine which of the qualifying children would be appropriate candidates (e.g., regular attendance). A total of 85 children were selected. One child withdrew from school at the beginning of intervention, resulting in 45 children in treatment and 39 children in comparison conditions. Figure S1 in the Supplemental Materials provides a CONSORT flow of clusters and child participants for the cluster randomized study. An additional 16 consented students from six of the treatment classrooms in Tampa who scored 100 or greater on the PPVT and the CELF comprised a NAR group for an exploratory analysis.

Demographic information was collected for all participating children through a survey sent home to their parents. A total of 62 (74%) surveys were returned for the AR group and 12 (75%) were returned for the NAR group. Not all items were completed for every child. Table 1 provides demographic characteristics for the child participants. Among the returned surveys, five parents indicated concerns about their child's development, with one reporting their child had been assessed recently for speech, behavior, and attention difficulties. No other parents reported that their child was eligible for or receiving special education services at the time of the survey. To determine dual language learner (DLL) status for the exploratory analysis, parents were asked to report all languages other than English spoken in the home and to rate how often those languages were used (i.e., 100%, 75%, 50%, or $<$50% of the time). All DLL children included in the exploratory analysis reportedly used a second language in the home at least 50% of the time. Although Spanish was the language most frequently reported by these parents, other reported languages were Arabic, Creole, French, Hindi, Jamaican Patois, Karen, Portuguese, Tamil, Telugu, Vietnamese, and Yoruba.

Materials and Procedures Common to Both Conditions

Teachers in both conditions received an MP3 player, five headphone sets with a splitter, four copies of each *Jungle Friends* book, and shortened, consumable versions of each storybook (i.e., mini-books) to send home. The *Jungle Friends* series included 13 books (1 introductory and 9 instructional storybooks, and 3 review books). The introductory storybook was used to introduce characters and establish procedures at the listening center. The other storybooks were divided into three units of three instructional books and one review book per unit.

All participating teachers and paraeducators received approximately 1 hr of training in the fall to learn how to implement the *Story Friends* program. Research staff delivered trainings to individual teachers or teacher/paraeducator teams at their schools using the procedural manual for each

Table 1. Demographic Characteristics for Child Participants.

	Group		
	Total AR sample	AR treatment	AR comparison
Variable	$n = 84$	$n = 45$	$n = 39$
Age in months at pretest (*SD*)	55.6 (3.5)	55.3 (3.3)	55.9 (3.7)
CELF P-2 pretest *M* (*SD*)	81.4 (9.1)	81.9 (8.8)	80.9 (9.4)
PPVT-4 pretest *M* (*SD*)	87.6 (8.2)	87.7 (9.2)	87.5 (7.4)
Demographic survey respondents	$n = 62$ (74%)	$n = 30$ (67%)	$n = 32$ (82%)
Ethnicity			
Hispanic/Latino	27 (44%)	11 (37%)	16 (50%)
Non-Hispanic	26 (42%)	17 (57%)	9 (28%)
Item not reported	9 (15%)	2 (7%)	7 (22%)
Race			
American Indian/Alaskan	2 (3%)	1 (3%)	1 (3%)
Asian	6 (10%)	1 (3%)	5 (16%)
Black/African American	20 (32%)	15 (50%)	5 (16%)
White	14 (23%)	4 (13%)	10 (31%)
Multiple races reported	0%	0%	0%
Unknown/item not reported	20 (32%)	9 (30%)	11 (34%)
Languages spoken in the home			
English only	21 (34%)	12 (40%)	9 (28%)
Spanish	24 (39%)	9 (30%)	15 (47%)
Arabic, French Creole, Other	14 (23%)	6 (20%)	8 (25%)
Item not reported	3 (4%)	3 (10%)	0%
Education			
Some high school, GED, or less	6 (10%)	2 (7%)	4 (13%)
High school	22 (35%)	8 (27%)	14 (44%)
Some 4-year college	4 (6%)	2 (7%)	2 (6%)
Associate degree	4 (6%)	1 (3%)	3 (9%)
Bachelor's degree	13 (21%)	10 (33%)	3 (9%)
Graduate/professional degree	4 (6%)	2 (7%)	2 (6%)
Unknown/item not reported	9 (15%)	5 (17%)	4 (13%)
Annual household income			
<US$20,000	19 (31%)	9 (30%)	10 (31%)
US$20,000-US$50,000	22 (35%)	8 (27%)	14 (44%)
>US$50,000	14 (23%)	10 (33%)	4 (13%)
Unknown/item not reported	7 (11%)	3 (10%)	4 (13%)

Note. AR = at risk; CELF P-2 = Clinical Evaluation of Language Fundamentals Preschool-2; PPVT-4 = Peabody Picture Vocabulary Test-4; NAR = not at risk.

condition. The trainings consisted of an overview of the importance of facilitating language development tailored to their treatment condition, basic procedures for delivering the program (i.e., how often to schedule these activities, how to configure students at the listening center and encourage their responses, how to use the MP3 player); and instructions for how to distribute home extension materials. Teachers in both conditions received home extension materials and a letter for caregivers briefly explaining the program and providing URL links to additional information.

Educators were responsible for delivering the *Jungle Friends* program to their students. Teachers were asked to read the weekly storybook once to the whole class. Children in the AR

group in both conditions participated in the small-group listening center three times per week. Each child listened using headphones connected to an MP3 player while following along in their own copy of the book. The teacher or a paraeducator facilitated the small-group listening center by preparing materials, operating the MP3 player, and helping children turn pages. The teacher was asked to encourage children to respond by providing children with non-verbal cues (i.e., giving a thumbs up, nodding, etc.). Table S1 in the Supplemental Materials provides a summary of the similarities and differences across conditions.

Treatment Condition

During small-group listening centers, two embedded lessons were provided for each of the four target words per book: One after the word occurred in the story and one at the end of the story. Each target word was used 11 times and definitions were used 7 times during explicit instruction, examples, and feedback. For each word, children had approximately six opportunities to respond (e.g., to repeat the word, to repeat the definition, to respond to questions about the word, or to respond with gestures) per listen. The unit review books provided brief lessons for all 12 words taught during a unit. Each target word was used four times, defined once, and children had two opportunities to respond for each word (e.g., repeat the word, say the definition) per listen. Table S2 in the Supplemental Materials provides a sample of an embedded lesson script.

Materials to encourage extended review and practice of target vocabulary were provided to teachers. These included teacher prompt cards, a word jar, and a review poster. Prompt cards contained the word, the definition, and examples of how the word might be used during the school day. To build classwide interest in word learning, teachers could place a pompom in the word jar when noting a child's use of a target word, or by drawing attention to their own use of a target word. The review poster included a set of 36 word cards for teachers to attach using Velcro. Each word card contained a target word, the definition, and corresponding pictures from the story. Because preschool children do not yet read, the words and definitions were provided for the teachers to recall target words and meanings while the pictures helped children recall how the words were used in the stories. During training, we explained that teachers may choose to use the poster to review the words as a regularly scheduled activity (e.g., during circle time) or they may choose to refer to the poster more flexibly as opportunities to talk about the words arise. We recommended the use of these materials throughout the school day but permitted flexible implementation of classroom extension activities.

Teachers sent home materials to facilitate home review and practice of vocabulary beyond the school day. A mini-book was sent home with every child in the classroom at the beginning of the week to enable families to read the stories and review the words. Vocabulary necklaces and stickers contained an illustration from the book, all four target words, and the prompt, "*Ask me what these words mean.*" Children attending the small-group listening center received a necklace to wear home at the end of the week. The following week, stickers were placed on the shirts of all children as a reminder to practice the words.

Comparison Condition

Children in the comparison condition listened to the same stories as children in the treatment condition, but the embedded instruction of vocabulary was not provided. These children were exposed to target vocabulary words within the stories during the whole-class read aloud and when listening to the story at the small-group listening center. They also were exposed to the target vocabulary words when they listened to the review books, which contained all 12 target vocabulary words from the unit but without any explicit instruction.

Comparison teachers were not made aware of the target vocabulary. In lieu of the vocabulary focus, these teachers delivered brief lessons that focused on story comprehension. Each lesson comprised a set of cards with pictured scenes for the children and brief scripts on the back for the teacher. The scripts were structured around questions that tapped into children's literal and inferential understanding of common story elements: characters (who?), setting (where?), events (what happened?), problem, solution, character motivation (why?), and character emotion (how did he or she feel?). The scripts provided prompts to use when children did not respond correctly to questions. Teachers were free to deliver these lessons at any time following the initial introduction to the weekly book. Teachers in the comparison condition also sent home mini-books each week. The mini-books included three comprehension questions related to the story to encourage discussion of the story.

Screening and Descriptive Measures

Clinical Evaluation of Language Fundamentals-Preschool, second edition. The CELF P-2 (Wiig et al., 2004) is a norm-referenced measure of oral language ability. The Core Language Score (CLS) subtests (i.e., Sentence Structure, Word Structure, and Expressive Vocabulary) were administered to identify children eligible for inclusion in the AR and NAR groups and to assess preintervention group equivalence. The split-half reliability for the CLS ranges from .92 to .94 for children between 4 and 5 years of age, and test–retest reliability is reported to be .89.

Peabody Picture Vocabulary Test, fourth edition. The PPVT-4 (Dunn & Dunn, 2007), a measure of receptive vocabulary was administered to identify children and to assess group equivalence before treatment. Split-half reliability reported for the PPVT-4 is .94 and test–retest reliability is .93 across age and grade levels.

Outcome Measures

Assessments of target vocabulary. The primary outcome measures were word scores derived from researcher-created curriculum-based assessments of the words taught. These assessments were designed to measure definitional knowledge of the words. Measures of definitional knowledge offer an advantage over frequently used receptive measures of word knowledge (see Hadley et al., 2016), which most often assess recognition of words through a forced-choice (i.e., pointing) task. By using a measure of definitional knowledge, more nuanced levels of word knowledge may be revealed based on semantic and contextual information provided. Due to the nature and duration of the intervention in this study, we would not expect to see changes on measures of general vocabulary; thus, more distal norm-referenced measures were not used as an outcome.

Children were asked to respond to one open-ended item per word (e.g., *Tell me, what does enormous mean?*). If a correct response was not provided, a single contextual prompt from the story was provided *(In the story, Ellie was Enormous. Enormous means____?).* For each item, a taught definition, a reasonable definition in the child's own words, or synonym was considered a correct response worth 2 points. A partial definition or use of the word in a meaningful sentence without defining the word was considered a partially correct response worth 1 point. An incorrect, unrelated, "I don't know" response, or repetition of the contextual prompt was scored with 0 point. Thus, scores could vary from 0 to 8 word points per book.

As a baseline measure of target vocabulary knowledge, children were assessed using a randomly selected sample of eight target words. We did not pretest all 36 words from the *Jungle Friends* program because, historically, children have demonstrated little or no knowledge of these words at pretest (Kelley et al., 2015; Peters-Sanders et al., 2020; Seven et al., 2020).

In addition, assessing pretest knowledge of 36 words can be taxing on young children, particularly when they have had little or no prior exposure to the words. To reduce this testing effect, pretesting of target words was assessed using one of three forms, each containing a random sample of eight of the target words. These forms were then randomly assigned within each classroom to participating children. Internal consistency for pretest is much lower than post-tests (Cronbach's $\alpha = .49$), which is likely due to a floor effect and restricted range of scores.

Posttest assessment of 12 target vocabulary words occurred at the end of each of three units of instruction. This typically occurred 1 or 2 days after unit completion. Together, the three end-of-unit assessments of target vocabulary measured posttest knowledge of all 36 *Jungle Friends* words. Internal consistency for the posttest assessment of target vocabulary is high (Cronbach's $\alpha = .92$, .90, and .90, for Units 1, 2, and 3, respectively).

To assess retention of vocabulary knowledge, target words from Unit 1 also were included on the Unit 3 post-test. To identify items for the retention measure, we examined learning from Unit 1 and chose the six words most frequently learned.

Scoring Reliability

Trained members of the research team completed scoring for all measures. Each CELF P-2 and PPVT-4 protocol was scored by a primary scorer and checked by a second scorer. The few disagreements were resolved by a third scorer. The assessments of target vocabulary were scored by a primary scorer, and approximately one third of these measures were scored by a second scorer. Scorers were blind to condition and to pretest versus post-test. Detailed scoring guides were used to facilitate reliable scoring and included specific criteria for assigning a score and several examples for each word in each score category. Prior to scoring, research assistants at both sites reviewed scoring criteria and guides and completed a training set. Item-level inter-scorer agreement was 100% at pretest and 99.4%, 99.1%, and 100% for the three unit post-tests.

Implementation Fidelity

Implementation fidelity was measured along five dimensions of treatment integrity (adherence, exposure, quality of delivery, participant responsiveness, and program differentiation), as described by Dane and Schneider (1998). Program exposure, participant responsiveness, and adherence were measured using teachers' session logs and research staff conducted live observations of adherence, quality of delivery, and program differentiation.

Teachers recorded the dates they completed the read alouds and listening centers on the session logs as well as child attendance. Teachers also rated the behavior of each child during the listening centers using a 3-point scale (i.e., mostly on task, sometimes distracted, or off-task behavior that interfered with learning). These data were used to measure exposure to the program and participant responsiveness. The session log also contained a procedural checklist for teachers to complete during each listening center session. The checklist included six items used to determine whether instruction was delivered as intended (i.e., each child had a book, each child had headphones, correct audio was played, entire audio was played, children were provided with reinforcement, and an adult was present in the listening center). These checklists provided teacher-reported adherence data.

Trained research staff completed observations in each classroom at three points during the 13-week program (approximately once per unit). Data from these observations were used to measure adherence and program differentiation during the listening center as well as aspects of quality of delivery. Research staff also recorded evidence of classroom review and practice of target words. Observations were scheduled in advance with teachers and lasted 20 to 30 min. The observation form included six items similar to the teacher procedural checklist (e.g., each child

used headphones, each child had a book, facilitator wore headphones, correct audio was played, entire audio was played, environment was quiet with few distractions) and items pertaining to the quality of implementation (e.g., the adult provided encouragement or positive feedback, cued for attention or redirected children, modeled appropriate verbal responses, assisted in page-turning, did not stop or interrupt the story, etc.) and whether evidence of review or practice of target vocabulary was observed (i.e., teachers discussed target words with children, referred to the review board, used Word Jar, etc.). In addition, research staff made anecdotal notes describing any classroom review or practice observed.

Social Validity Measures

Social validity surveys were sent to participating parents and teachers to collect feasibility and implementation information about the *Story Friends* program. At the end of the study, teachers were asked to rate their level of agreement with statements that addressed their impressions (e.g., "The intervention is a good way to address language delays") and their implementation ("Story Friends can be implemented as frequently as recommended") of the program. Open-ended questions gauged students' reaction to the program, teachers' use of the classroom extension activities, and what changes, if any, they would make to the program. All teachers also participated in a structured interview after completing the program.

To collect social validity data on home review materials, a satisfaction survey was sent home to parents or caregivers after the first and third units of instruction to prompt return of the survey. This survey asked parents or caregivers to rate the frequency in which their child brought home extension materials (e.g., My child came home wearing a vocabulary necklace each week; my child brought home a mini book each week), to note the frequency with which they practiced words at home, and to answer open-ended questions about their satisfaction with the program.

Data Analysis

To examine the effects of the *Story Friends* intervention, the primary outcome was word scores at pre- and post-test for each of the nine instructional books on the unit assessments of target vocabulary. For each book children could score a maximum of eight points (two points per word for four words). Because children were nested within classrooms, we used a multi-level modeling (MLM) approach to analyze our data. The approach accounts for missing data, allowing us to analyze our data without removing participants who have one or more missing data points (Baayen et al., 2008). All analyses were conducted using JMP Pro 15.

Data from the assessments of target vocabulary were examined using the missing data function in JMP Pro 15. In the treatment group, 96% of the participants had complete data at pre- and post-test (43 out of 45). Two participants from two classrooms withdrew from school and were missing data for the last three books at post-test (7, 8, 9). In the comparison condition, 95% of the participants had complete data at pre- and post-test (37 out of 39). Two children from one classroom withdrew from school and were missing data for the last three books at post-test (7, 8, 9). Thus, data were missing for 6 out of 810 observations in the treatment group (0.74% of 9 books $\times$ 2 observations each $\times$ 45 participants), and 6 out of 702 observations in the comparison group (0.85% of 9 books $\times$ 2 observations $\times$ 39 participants).

We conducted a $2 \times 2 \times 9$ mixed between- and within-subjects analysis and modeled several variables as fixed effects. The first factor was between subjects and evaluated differences between the treatment and comparison groups. The second factor was within subjects and modeled participants' word scores at pre-intervention and post-intervention (time). The last factor was within subjects and modeled differences among the nine instructional books.

Table 2. Group Comparisons for Independent Variables.

	Treatment		Comparison				
Variable	*M*	*SD*	*M*	*SD*	*t*	*p*	*d*
CELF P-2	81.89	9.37	81.3	9.03	−0.29	0.38	0.06
Pre-PPVT-4	87.69	7.35	87.75	9.23	0.03	0.51	0.01
Pre-Word Score	0.16	0.25	0.11	0.19	−1.09	0.14	0.23
Post-Word Score	4.35	2.15	1.12	0.88	−9.23	.0001	1.92

Note. CELF P2 = Clinical Evaluation of Language Fundamentals Preschool-2 (Wiig et al., 2004); PPVT-4 = Peabody Picture Vocabulary Test-4 (Dunn & Dunn, 2007); Pre Word Score = pretest score per book out of eight possible points.

Results

Data Screening and Descriptive Results

The distribution for measures was analyzed for normalcy using the Distribution function in JMP Pro 15. Visual inspection of the distributions, their QQ plots, and skew of the CELF P-2, pre-intervention PPVT-4, pre-intervention word score, and post-intervention word score indicated all the variables were approximately normally distributed. We tested for pre-intervention differences between treatment and comparison groups using several between-subjects *t*-tests reported in Table 2. There were no significant differences between groups on the CELF P-2, PPVT-4, or targeted word scores. There was a small, not significant effect for pre-intervention word scores; because this variable was an outcome in all our models it was not necessary to include it as a covariate in the analysis.

Effects on Vocabulary Learning in Children AR

To account for the clustered structure of the data, the random effects for differences between children and for differences between classrooms were included in the analysis. The intra-class correlation (ICC) for children was .015, indicating that 1.5% of the differences in word scores were attributed to differences between children. The ICC for classroom was .188 indicating that 18.8% of the differences in word scores were attributed to differences between classrooms. An additional 79.7% of the variance in word scores remained to be explained by the fixed effects that were included in the model next.

The final model fitted using restricted maximum likelihood is presented in Table 3. The model containing fixed effects accounted for 71% of the variance in word learning outcomes. There were significant main effects for treatment condition $F(1, 23) = 32.25$, $p < .01$, time $F(1, 1{,}383) = 1{,}421.22$, $p < .01$, and book $F(8, 1{,}382) = 5.24$, $p < .01$. In addition, there were significant two-way interactions for each of the possible combinations of factors, treatment condition and book $F(8, 1{,}382) = 4.87$, $p < .01$, treatment condition and time $F(1, 1{,}383) = 538.7$, $p < .01$, and book and time $F(8, 1{,}382) = 4.6$, $p < .01$; and a significant three-way interaction between treatment condition, time, and book $F(8, 1{,}382) = 4.52$, $p < .01$.

Figure 1 presents the estimated marginal means for the treatment and comparison groups at pre- and post-intervention across all nine books. Each mean is presented with a 95% confidence interval. The pre-intervention scores were consistently low and did not differ between treatment conditions for any of the books. The figure clearly depicts the significant three-way interaction between condition, book, and time. Children in the treatment group had significantly higher post-intervention word scores ranging from 5.71 to 3.04 out of eight across books. However, as children progressed through the intervention, vocabulary learning tended to decrease across books.

Table 3. Multilevel Results for Final Model.

Variable	Parameter estimate	*SE*	DFDen	*t* Ratio	Variance component	*SE*
Intercept	1.42*	0.145	22.65	9.83		
Fixed effects						
Condition	0.82*	0.145	22.65	5.68		
Book 1	0.49*	0.097	1,382	5.01		
Book 2	0.03	0.097	1,382	0.34		
Book 3	−0.02	0.097	1,382	−0.25		
Book 4	0.12	0.097	1,382	1.25		
Book 5	−0.13	0.097	1,382	−1.37		
Book 6	−0.07	0.097	1,382	−0.74		
Book 7	−0.16	0.099	1,383	−1.57		
Book 8	0.11	0.099	1,383	1.14		
Time	−1.31*	0.035	1,383	−37.70		
Condition × Book 1	0.20*	0.097	1,382	2.04		
Condition × Book 2	0.32*	0.097	1,382	3.29		
Condition × Book 3	−0.04	0.097	1,382	−0.45		
Condition × Book 4	0.17	0.097	1,382	1.70		
Condition × Book 5	−0.18	0.097	1,382	−1.84		
Condition × Book 6	0.20*	0.097	1,382	2.09		
Condition × Book 7	−0.09	0.099	1,383	−0.88		
Condition × Book 8	−0.28*	0.099	1,383	−2.80		
Condition × Time	−0.80*	0.035	1,383	−23.21		
Book 1 × Time	−0.48*	0.097	1,382	−4.88		
Book 2 × Time	−0.07	0.097	1,382	−0.68		
Book 3 × Time	0.11	0.097	1,382	1.14		
Book 4 × Time	−0.09	0.097	1,382	−0.92		
Book 5 × Time	0.12	0.097	1,382	1.28		
Book 6 × Time	0.04	0.097	1,382	0.38		
Book 7 × Time	0.10	0.099	1,383	0.99		
Book 8 × Time	−0.07	0.099	1,383	−0.75		
Condition × Book 1 × Time	−0.20*	0.097	1,382	−2.00		
Condition × Book 2 × Time	−0.34*	0.097	1,382	−3.46		
Condition × Book 3 × Time	0.02	0.097	1,382	0.20		
Condition × Book 4 × Time	−0.09	0.097	1,382	−0.94		
Condition × Book 5 × Time	0.16	0.097	1,382	1.65		
Condition × Book 6 × Time	−0.20*	0.097	1,382	−2.01		
Condition × Book 7 × Time	0.07	0.099	1,383	0.74		
Condition × Book 8 × Time	0.26*	0.099	1,383	2.58		
Random effects						
Children within classroom					0.16	0.05
Classroom					0.38	0.14

*$p < .01$.

To assess the size of the treatment effect, we compared the post-intervention word scores of the treatment and comparison groups for each instructional book and calculated Hedges' *g*. Effect size estimates ranged from 1.05 to 2.12 across the nine books, with a mean effect size of $g = 1.55$ all of which exceeded the benchmark of .80 for a large effect as established by Cohen (1988). Thus, the treatment consistently demonstrated a large impact on vocabulary learning.

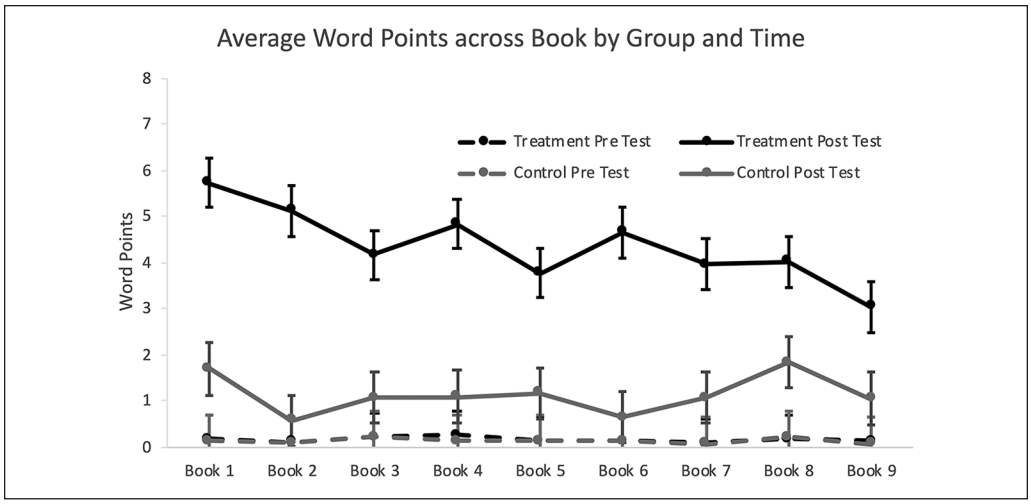

Figure 1. Average word points earned across books by group at pre- and post-intervention. Error bars are 95% confidence intervals.

Moderation in the Treatment AR Group

Due to the large effect of *Story Friends* on word learning, we explored two potential moderating variables within the treatment group: the number of instructional listens and students' pretest CELF P-2 scores. We used MLM analysis to evaluate the moderating effects these variables had on the relation between book and word learning. There was no main effect of instructional listens on word learning $F(1, 39) = 0.30$, $p = .6$, and no interaction effect between instructional listens and book $F(8, 743) = 0.46$, $p = .88$. There was a significant main effect for the CELF P-2 on word score, $F(1, 33) = 14.48$, $p < .01$. The unstandardized regression coefficient associated with that effect, $B = 0.05$, $B\ SE = 0.01$, indicated that for every one-point increase in CELF P-2 scores there was a 0.05 increase in word points across books. There was no significant interaction between CELF P-2 and book, $F(8, 744) = 0.22$, $p = .99$.

Because only 27 parents from treatment classrooms reported languages used in the home on a family demographic survey, this information was used in a preliminary comparison of word learning between 15 DLLs and 12 native English speakers from 11 treatment classrooms. T-tests revealed no significant differences between groups on pre-intervention word scores, $t(25) = .54$, $p = .59$; average word scores per book were .19 ($SD = .06$) for DDLs and .14 ($SD = .18$) for native English speakers. More importantly, there were no significant differences on post-intervention word scores, $t(25) = 1.10$, $p = .28$; the average post-intervention word score per book for the DLL group was 4.81 ($SD = 2.6$) and 3.94 ($SD = 1.5$) for the native English-speaking group.

Retention of Vocabulary in the Treatment AR Group

We assessed children's retention of vocabulary words at the end of the intervention. Six of the best-learned words from Unit 1 were included in the posttest for Unit 3, which was administered approximately 2 months after the words were taught. Word scores at pre-intervention, post-intervention, and retention for children in the treatment group were included in a one-way within-subjects ANOVA. Results indicated that there was a significant difference among the three assessment points, $F(2, 117) = 57.17$, $p < .001$. Based on a Tukey's test HSD post hoc analysis, the pre-intervention word score of 8% ($M = 0.63$, $SD = 0.93$) was significantly lower than 71% at post-intervention ($M = 8.48$, $SD = 3.55$) representing a large effect ($g = 2.77$) and

56% at follow-up ($M = 6.73$ $SD = 3.71$) also representing a large effect ($g = 2.06$). However, there was a significant difference between word scores at post-intervention vs. at follow-up. Children demonstrated a decrement in word knowledge retention (1.75 out of 12 word points) from the first unit of instruction to the end of intervention, representing a medium effect ($g = 0.48$).

Effects on Vocabulary Learning in Children NAR

T-tests were used to analyze the differences between children in the AR and NAR groups. We compared the performance of 16 children NAR in six treatment classrooms to 24 of their peers AR from the same classrooms. Importantly, there were no significant differences between groups on pre-intervention word scores, $t(39) = 1.9$, $p = .06$. The average post-intervention word score per book for the AR group was 5.6 ($SD = 2.0$) and the average word score per book for the NAR group was 6.4 ($SD = 1.5$; out of eight possible points). Both groups made similar gains, as they did not differ significantly post-intervention, $t(39) = 1.67$, $p = .10$.

Implementation Fidelity

Session logs included child attendance data for 97% of the listening center sessions. Children in the treatment condition completed 97% (range 78-100%) and children in the comparison group completed 97% (range 86-100%) of the expected listens. Across the three units, the average number of reported listens of a possible 12 was 11.5, 11.3, and 11.0, respectively, for children in the treatment group, and 11.3, 11.6, and 10.9, respectively, for the comparison condition. Teachers completed a mean of 83% of read alouds (range 33-100% treatment; 25-100% comparison). Teacher ratings of child responsiveness were completed less frequently than attendance, with 8% and 7% missing for treatment and comparison classrooms, respectively. Teachers rated children "mostly on task" with means of 83% and 88% in treatment and comparison classrooms, respectively. Off-task behavior that interfered with learning was rarely reported (3% and 1% in treatment and comparison classrooms). Teachers completed a mean of 93% of procedural checklists. Adherence was high, averaging 97% (range 92-100%) in treatment and 95% (range 53-100%) in comparison classrooms.

Scheduling conflicts occasionally interfered with the three planned observations per classroom, resulting in averages of 2.75 and 2.67 observations in treatment and comparison classrooms, respectively. Observed adherence to the program was high, averaging 96% (range 67-100%) in treatment and 97% (range 83-100%) in comparison classrooms. In treatment classrooms, teachers modeled verbal responses and provided encouragement (i.e., nodding and smiling, giving a thumbs up during child responses) during 80% and 87% of the observed lessons, respectively. Treatment teachers assisted children (i.e., helping them find the right page, turning pages, redirecting them to the task) during 100% of the observations. Teachers in the comparison classroom provided praise and encouragement to children while listening to the books during 90% of the observations and redirected children to the task and provided assistance to children during 97% of the observations.

Research staff recorded whether they observed teachers reviewing or practicing target words and took anecdotal notes about what they observed. Evidence of review or practice was recorded when researchers observed teachers discussing any of the target words and their meanings with children. No comparison teachers were observed to provide instruction of target vocabulary words. Review or practice of target words was observed in 8 out the 12 treatment classrooms during 16 of the 33 total observations completed across all classrooms (i.e., during 48% of all observations). No review or practice was observed in four classrooms. Teachers used a variety of strategies to extend practice beyond the storybook: discussing target words and their meanings prior to or after the story, placing word cards on the review board and discussing the words,

modeling the use of a word in a sentence, asking children to provide the meaning of a target word, asking children to say or repeat a word, and asking questions about the words. From the four classrooms where review and practice were not observed, three of these teachers self-reported discussing specific target words during circle time or with individual children at various times during the day. For example, one teacher explained that she recognized a child's effort for using the word *gorgeous* to describe another child while standing in line. Another teacher reported that "the words *announce*, *select*, and *gorgeous* were easy to use in the classroom."

Social Validity

Only 9 of 24 teachers returned social validity surveys (6 treatment, 2 comparison). This 38% rate of return was low compared with a prior study in which 75% of teachers responded (Kelley et al., 2020). As a result, we completed live interviews with all teachers. Overall, survey results were positive. Teachers felt that the program was easy to implement ($M = 5.67$ out of 6) and that it was a good way to address language delays ($M = 5.88$ out of 6). Teachers in the treatment classrooms rated the classroom and home extension materials highly as well. Teachers reported sending home necklaces, stickers, and mini books each week ($M = 6$ for each component), whereas the use of the word jars received a mean rating of 4.25 out of 6. Teachers indicated high interest in using the *Story Friends* program in the future ($M = 5.38$ out of 6).

The teacher interviews were conducted individually. All teachers stated they would use the program in the future and that it was easy to implement. Several said the stories and the characters were engaging and relatable to children and that their students began to use target words in the classroom. All treatment teachers valued the review board and the mini-books, but few reported using the word jar consistently and some indicated that it was difficult to remember to use the jar every time they discussed a word. Teachers also indicated that parental response to the program was largely positive. In several treatment classrooms, teachers reported that parents were "amazed" that their children had learned these words. One teacher stated, "a parent told me that the conversation was elevated with her child." Treatment teachers offered some suggestions for change (e.g., add more words, design an option to play the audio for the whole class, eliminate cumbersome necklaces).

Out of 84 parents who received social validity surveys, 31 returned a survey (20 from 8 different treatment classrooms, and 11 from 8 different comparison classrooms). Although the parent response rate was only 37%, this was an improvement from the 13% returned in the previous study (Kelley et al., 2020). Parents rated the use of the home extension program using a 3-point rating scale (1 = never, 2 = sometimes, and 3 = often). Parents of children in the treatment condition indicated that their child came home wearing a vocabulary necklace often, ($M = 2.6$ out of 3) and vocabulary stickers sometimes ($M = 2.0$ out of 3). Several parents commented that their child wore the necklace at home and got family members or friends to ask them the meaning of words. They stated that their child brought home the mini book often ($M = 2.9$ out of 3). Parents noted that the materials sent home were easy to use, did not require a lot of time to implement, and were a good reminder to review the vocabulary words and definitions at home. Several commented that their children liked to re-read the stories at home and that the books were a good way to review the vocabulary words. Parents reported practicing the vocabulary words and their definitions at home most days of the week (3-5 days).

Discussion

This study extends prior research on *Story Friends* by examining the vocabulary learning of preschoolers AR for language and learning difficulties who completed the revised *Jungle Friends* book series. The revised program teaches twice as many words as the original program

and adds materials to promote review and practice of words beyond the listening center. As in previous studies, children AR in the treatment group learned to define many more target words (54%) than children AR in the comparison group (14%), representing a large treatment effect. This is consistent with findings of other studies examining the benefits of explicit vocabulary instruction over vocabulary exposure alone (e.g., Coyne et al., 2007). Learning was evident for all nine books in this series, but there was a gradual decrease in the amount of learning per book across the study. For the six best-learned words from Unit 1, learning averaged 71% at post-test, and retention of knowledge was 56% approximately 2 months later. Other studies of vocabulary intervention have observed similar decrements in definitional knowledge from post-test to maintenance testing (e.g., Coyne et al., 2009; Pullen et al., 2010), although in a previous study of *Story Friends* there was no significant difference between post-intervention and retention scores (Kelley et al., 2020).

Pretest language scores on the CELF P-2 had a small, significant moderating effect on learning. This finding is not surprising given that children with higher language scores tend to benefit more from vocabulary instruction (Coyne et al., 2019; Goldstein et al., 2017). In addition, an analysis of the results from a subset of children with reported home language data revealed that DLLs did not differ from monolingual English speakers at pretest or posttest. The DLLs were able to define a mean of 60% of words post-intervention versus a mean of 49% of words by the monolingual English speakers; this difference was not significant. These findings align with other studies that examined vocabulary instruction with DLLs. For example, Silverman (2007) found no differences in target word learning between monolingual English speakers and DLLs after 14 weeks of vocabulary instruction. In addition, other researchers have found that English learners acquired more words when they received instruction including definitions and rich explanations of words (Collins, 2010) as well as extended instruction of words beyond the storybook (Loftus-Ratan et al., 2016). Although some researchers have found that DLLs benefit most when they receive vocabulary instruction in both their native language and English (Mendez et al., 2015; Restrepo et al., 2013), many early childhood classrooms may lack the bilingual staff to provide instruction in both languages. Although the samples in these studies have consisted largely of Spanish-speaking DLLs, our sample is more linguistically heterogenous, yet our preliminary findings indicate that vocabulary instruction can be beneficial to young DLLs, even when provided in English.

Although it is difficult to compare results across vocabulary intervention studies due to differences in participant characteristics, words taught, and assessment methods, it appears that we continue to make incremental progress in providing effective vocabulary instruction to preschool children. Across studies, children typically learn only a portion of the words taught. In Loftus et al. (2010), at-risk children who received small-group and classroom-based vocabulary instruction knew 27.5% of the taught words on the definitional task at post-test. In Coyne et al. (2010), at-risk children who received vocabulary instruction during interactive storybook reading knew 37.5% of target words on the definitional task at posttest. Kelley et al. (2020) found that children AR who received instruction using the revised *Forest Friends* series knew 42% of taught words at posttest (15 of a possible 36). Children AR in the current study who received instruction using *Jungle Friends* knew 54% of taught words at posttest. Several factors could be responsible for this incremental improvement in learning including possible differences in a new cohort of child participants or teachers and their program delivery, changes to home extension materials, and differences in words. We cannot affirm that this is a reliable improvement without further replications, but this encouraging finding exposes factors to be considered for future Hybrid Designs that combine effectiveness and implementation evaluations.

The revised *Story Friends* program was designed to promote the extended practice of target vocabulary beyond the listening center. We expected that encouraging teachers to use classroom extensions would increase the incidental learning opportunities initiated by teachers across the

day. The extended practice appears to have benefited all children in treatment classrooms. Children NAR from a subsample of six classrooms knew an average of 80% of target words without receiving the intensive and highly structured explicit instruction provided during small-group listening centers. Children AR receiving the explicit instruction in the same classrooms knew an average of 70% of words at post-test. These six classrooms were slightly higher performing than our average for treatment classrooms in general, but the difference between children AR and NAR is not significant. It is challenging to directly measure the extended practice used throughout the school day. The learning among children not AR provides evidence that review and practice occurred in these classrooms. In addition, this finding adds support to the hypothesis that children with higher language abilities require less intense instruction to learn new words (Coyne et al., 2019; Pullen et al., 2010; Seven et al., 2020).

Limitations and Future Directions

The gradual decrease in word learning over time and the loss of word knowledge observed at 2-month follow-up are areas that warrant further examination. This decrement in learning and retention may relate to fatigue with the program or with repeated testing. It also is possible that, as children encounter more words and definitions, they are prone to confusion among words with similar meanings or used in similar contexts. Future studies should be designed to explore variations in testing and intervention delivery to examine these diminishing learning and maintenance effects. In addition, future studies are needed to examine the long-term effects of vocabulary learning from programs like *Story Friends* on later reading development.

Although we collected some data on the classroom and home extension activities, we were unable to directly measure the extent to which these learning opportunities occurred. As such, we cannot reliably estimate the additional time devoted to review and practice events interspersed across the day, but it likely varied among children. We were able to examine whether other children in a sample of the classrooms who were not receiving small-group instruction were learning the target words. These children with typical language development made gains similar to the at-risk children. Although it is not possible to isolate the effects of the classroom and home extensions as components of the larger program, small-scale studies have demonstrated the effects of these components (Seven et al., 2020; Soto et al., 2020). Future research should aim to capture more precise information on use of classroom and home extensions to understand how these aspects of the program contribute to children's learning.

The preliminary results of children who were DLLs in this study are promising. Our data were limited because not all families reported language use in the home, which lowered the sample size for a comparison of learning between DLLs and monolingual English speakers. In addition, the language assessments used in this study were not designed specifically for DLLs. However, parents of 8 of the 15 DLLs included in this exploratory analysis reported that their child was most comfortable using English despite the use of another language at home, and 4 reported their child was equally comfortable using English and their other language. Given the linguistic diversity of this subsample and that the primary purpose of the study was not to evaluate the effects of the intervention on DLLs, assessing each child in their home language was not within the scope of the study. In exploring the results for these children, we anticipated that *Story Friends* could have a positive impact, but did not expect to see such remarkable growth in their vocabulary knowledge. Additional studies should be designed specifically to investigate the effects of this program on this population of children.

Finally, capturing changes in vocabulary with norm-referenced measures after short periods of time, even after vocabulary instruction, is challenging (Elleman et al., 2009; National Reading Panel, 2000). Standardized measures of general vocabulary are not sensitive enough to detect

subtle changes over a short period, nor are they designed to capture learning of specific words (Hadley & Dickinson, 2020). Across intervention studies, few programs that target the meaning of specific words have a significant impact on general vocabulary knowledge (e.g., Coyne et al., 2010, 2019; Elleman et al., 2009; Silverman et al., 2013). *Story Friends* targets specific words and does not focus on increasing more generative word learning; therefore, we would not expect to see changes on these more distal measures after 3 months of intervention. Thus, like other vocabulary intervention researchers (e.g., Coyne et al., 2007; Loftus-Rattan et al., 2016; Restrepo et al., 2013).), we did not examine distal outcomes of general vocabulary. In the context of this limitation, the question often arises as to whether teaching a small set of vocabulary words is worth pursuing. Our position is that young children with limited vocabulary are at a long-term disadvantage. Advancing their knowledge of sophisticated vocabulary (i.e., increasing their lexicon) can only benefit them. The more words a child knows, the better they will understand the context of the story. From a theoretical perspective, having more word knowledge also will prepare them for later reading and comprehension tasks. In addition, knowing more words makes learning new words possible, as word learning involves forming strong semantic connections with known words (Hadley & Dickinson, 2020; Hadley et al., 2018). The positive effects of learning specific words may not be readily seen in the preschool child, but they can be anticipated later when children are learning to read and reading to learn. Because it has been established that early childhood education tends to neglect vocabulary development in young children, this program attempts to help early childhood teachers address this paucity by providing easy-to-deliver explicit vocabulary instruction that results in word learning. Future research should strive to examine the long-term impact that programs like *Story Friends* have on later vocabulary, reading, and comprehension outcomes. In addition, research should explore the potential impact of using these programs on preschool teachers' instructional practices surrounding vocabulary instruction.

Conclusion

In this study, *Jungle Friends* produced large treatment effects on vocabulary learning of children when preschool educators implemented the program with intended dosage. Children AR receiving explicit instruction during small-group listening centers and NAR receiving only classroom and home review and practice made notable gains. This study contributes to continued work on the *Story Friends* intervention, providing an example of the utility of time-intensive and rigorous development research to produce effective and feasible interventions.

Our findings suggest that *Story Friends* provides a viable option for delivering effective, explicit vocabulary instruction to children in preschool classrooms. Procedural fidelity as recorded by both teachers and research staff was high, with children receiving more than 90% of the intended listens. The implementation of the *Story Friends* program appears to place minimal demands on teachers, with almost no planning required and a listening center activity that fits into existing center rotations. The automated instruction alleviates the potentially unrealistic expectation that preschool teachers and parents have the time and knowledge to carefully select words, generate child-friendly definitions, identify multiple contexts for explaining meanings, and repeat instruction with sufficient intensity. The program ensures preschool children have access to valuable vocabulary instruction that they otherwise are unlikely to receive.

Authors' Note

R.M.B. contributed to this article in his personal capacity. The views expressed are his own and do not necessarily represent the views of the National Institute of Occupational Safety and Health, the Centers for Disease Control and Prevention, or the U.S. Government.

Declaration of Conflicting Interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: Beyond the authors' salaries at their respective institutions, Howard Goldstein and Elizabeth Kelley also declared a financial interest through the publication of the *Story Friends* curriculum for which they receive royalties from Brookes Publishing.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grants R324A150132 and R324A200179 to the University of South Florida. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education. The research was approved by the University of South Florida and the University of Missouri Institutional Review Boards. These institutions manage the Conflicts of Interest for the authors of *Story Friends*. Elizabeth Spencer Kelley and Howard Goldstein are the authors of *Story Friends*, published by Paul Brookes, and they receive royalties.

ORCID iD

Keri M. Madsen https://orcid.org/0000-0001-6008-9017

Supplemental Material

Supplemental material for this article is available online.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal*, *107*(3), 251–271. https://doi.org/10.1086/511706

Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*(2), 278–293. https://doi.org/10.1044/1092-4388(2006/023)

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Collins, M. F. (2010). ELL preschoolers' English vocabulary acquisition from storybook reading. *Early Childhood Research Quarterly*, *25*(1), 84–97. https://doi.org/10.1016/j.ecresq.2009.07.009

Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly*, *30*(2), 74–88. https://doi.org/10.2307/30035543

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, *110*(1), 1–18. https://doi.org/10.1086/598840

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Ruby, M., Crevecoeur, Y. C., & Kapp, S. (2010). Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness*, *3*(2), 93–120. https://doi.org/10.1080/19345741003592410

Coyne, M. D., McCoach, D. B., Ware, S., Austin, C. R., Loftus-Rattan, S. M., & Baker, D. L. (2019). Racing against the vocabulary gap: Matthew effects in early vocabulary instruction and intervention. *Exceptional Children*, *85*(2), 163–179. https://doi.org/10.1177/0014402918789162

Curran, G., Bauer, M., Mittman, B., Pyne, J. M., & Stetler, C. (2012). Effectiveness-implementation hybrid designs: Combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical Care*, *50*(3), 217–226. https://doi.org/10.1097/MLR.0b013e3182408812

Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and earlysecondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, *18*(1), 23–45. https://doi.org/10.1016/S0272-7358(97)00043-3

Dickinson, D. K. (2011). Teachers' language practices and academic outcomes of preschool children. *Science*, *333*(6045), 964–967. https://doi.org/10.1126/science.1204526

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test - 4* (4th ed.). NCS Pearson Assessments.

Dwyer, J., & Harbaugh, A. G. (2020). Where and when is support for vocabulary development occurring in preschool classrooms? *Journal of Early Childhood Literacy*, *20*(2), 252–295. https://doi.org/10.1177/1468798418763990

Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1–44. https://doi.org/10.1080/19345740802539200

Elwér, Å., Keenan, J. M., Olson, R. K., Byrne, B., & Samuelsson, S. (2013). Longitudinal stability and predictors of poor oral comprehenders and poor decoders. *Journal of Experimental Child Psychology*, *115*(3), 497–516. https://doi.org/10.1016/j.jecp.2012.12.001

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kinergarten readiness. *Developmental Psychology*, *54*(7), 1317–1333. https://doi.org/10.1037/dev0000514

Goldstein, H., & Kelley, E. S. (2016). *Story friends teacher guide*. Paul H. Brookes.

Goldstein, H., Kelley, E. S., Greenwood, C., McCune, L., Carta, J., Atwater, J., Guerrero, G., McCarthy, T., Schneider, N., & Spencer, T. (2016). Embedded instruction improves vocabulary learning during automated storybook reading among high-risk preschoolers. *Journal of Speech, Language, and Hearing Research*, *59*(3), 484–500. https://doi.org/10.1044/2015_JSLHR-L-15-0227

Goldstein, H., Ziolkowski, R. A., Bojczyk, K. E., Marty, A., Schneider, N., Harpring, J., & Haring, C. D. (2017). Academic vocabulary learning in first through third grade in low-income schools: Effects of automated supplemental instruction. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3237–3258. https://doi.org/10.1044/2017_JSLHR-L-17-0100

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development*, *90*(3), 985–992. https://doi.org/10.1111/cdev.13128

Greenwood, C. R., Carta, J. J., Atwater, J., Goldstein, H., Kaminski, R., & McConnell, S. (2013). Is a response to intervention (RTI) approach to preschool language and early literacy instruction needed? *Topics Early Childhood Special Education*, *33*(1), 48–64. https://doi.org/10.1177/0271121412455438

Hadley, E. B., & Dickinson, D. K. (2020). Measuring young children's word knowledge: A conceptual review. *Journal of Early Childhood Literacy*, *20*(2), 223–251. https://doi.org/10.1177/1468798417753713

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2018). Building semantic networks: The impact of a vocabulary intervention on preschoolers' depth of word knowledge. *Reading Research Quarterly*, *54*(1), 41–61. https://doi.org/10.1002/rrq.225

Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181–198. https://doi.org/10.1002/rrq.130

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Paul H. Brookes.

Heidlage, J. K., Cunningham, J. E., Kaiser, A. P., Trivette, C. M., Barton, E. E., Frey, J. R., & Roberts, M. Y. (2020). The effects of parent-implemented language interventions on child linguistic outcomes: A meta-analysis. *Early Childhood Research Quarterly*, *50*(1), 6–23. https://doi.org/10.1016/j.ecresq.2018.12.006

Hulleman, C. S., & Cordray, D. S. (2009). Moving from the lab to the field: The role of fidelity and achieved relative intervention strength. *Journal of Research on Educational Effectiveness*, *2*(1), 88–110. https://doi.org/10.1080/19345740802539325

Justice, L. M., Mashburn, A. J., Hamre, B. K., & Pianta, R. C. (2008). Quality of language and literacy instruction in preschool classrooms serving at-risk pupils. *Early Childhood Research Quarterly*, *23*(1), 51–68. https://doi.org/10.1016/j.ecresq.2007.09.004

Justice, L. M., Mashburn, A. J., Pence, K. L., & Wiggins, A. (2008). Experimental evaluation of a preschool language curriculum: Influence on children's expressive language skills. *Journal of Speech, Language, and Hearing Research*, *51*(4), 983–1001. https://doi.org/10.1044/1092-4388(2008/072)

Kaderavek, J. N., & Justice, L. (2010). Fidelity: An essential component of evidence-based practice in speech-language pathology. *American Journal of Speech Language Pathology*, *19*(4), 369–379. https://doi.org/10.1044/1058-0360(2010/09-0097)

Kelley, E. S., Barker, R. M., Peters-Sanders, L., Madsen, K., Seven, Y., Soto, X., Olsen, W., Hull, K., & Goldstein, H. (2020). Feasible implementation strategies for improving vocabulary knowledge of high-risk preschoolers: Results from a cluster-randomized trial. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4000–4017. https://doi.org/10.1044/2020_JSLHR-20-00316

Kelley, E. S., & Goldstein, H. (2014). Building a tier 2 intervention: A glimpse behind the data. *Journal of Early Intervention*, *36*(4), 292–312. https://doi.org/10.1177/1053815115581657

Kelley, E. S., Goldstein, H., Spencer, T. D., & Sherman, A. (2015). Effects of automated Tier 2 storybook intervention on vocabulary and comprehension learning in preschool children with limited oral language skills. *Early Childhood Research Quarterly*, *31*, 47–61. https://doi.org/10.1016/j.ecresq.2014.12.004

Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly*, *50*(2), 151–169. https://doi.org/10.1002/rrq.99

Language Reading Research Consortium & Logan, J. (2017). Pressure points in reading comprehension: A quantile multiple regression analysis. *Journal of Educational Psychology*, *109*(4), 451–464. https://doi.org/10.1037/edu0000150

Loftus, S. M., Coyne, M. D., McCoach, D. B., Zipoli, R., & Pullen, P. C. (2010). Effects of a supplemental vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties. *Learning Disabilities Research & Practice*, *25*(3), 124–136. https://doi.org/10.1111/j.1540-5826.2010.00310.x

Loftus-Rattan, S. M., Mitchell, A. M., & Coyne, M. D. (2016). Direct vocabulary instruction in preschool: A comparison of extended instruction, embedded instruction, and incidental exposure. *The Elementary School Journal*, *116*(3), 391–410. https://doi.org/10.1086/684828

McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly*, *29*(4), 520–530. https://doi.org/10.1016/j.ecresq.2014.06.002

Mendez, L. I., Crais, E. R., Castro, D. C., & Kainz, K. (2015). A culturally and linguistically responsive vocabulary approach for young Latino dual language learners. *Journal of Speech, Language, and Hearing Research*, *58*(1), 93–106. https://doi.org/10.1044/2014_JSLHR-L-12-0221

Morgan, P. L., Farkas, G., Hillemeier, M. M., Hammer, C. S., & Maczuga, S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, *86*(5), 1351–1370. https://doi.org/10.1111/cdev.12398

National Reading Panel. (2000). *Report of the National reading panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. US Government Printing Office.

Nelson, K. E., Welsh, J. A., Trup, E. M. V., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. *First Language*, *31*(2), 164–194. https://doi.org/10.1177/0142723710391887

Neugebauer, S., Coyne, M., McCoach, B., & Ware, S. (2017). Teaching beyond the intervention: The contribution of teacher language extensions to vocabulary learning in urban kindergarten classrooms. *Reading and Writing*, *30*, 543–567. https://doi.org/10.1007/s11145-016-9689-x

Neuman, S., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-k. *The Reading Teacher*, *62*(5), 384–392. https://doi.org/10.1598/RT.62.5.2

Pelatti, C. Y., Piasta, S. B., Justice, L. M., & O'Connell, A. A. (2014). Language- and literacy-learning opportunities in early childhood classrooms: Children's typical experiences and within-classroom variability. *Early Childhood Research Quarterly*, *29*(4), 445–456. https://doi.org/10.1016/j.ecresq.2014.05.004

Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, *94*(1), 23–33. https://doi.org/10.1037//0022-0663.94.1.23

Peters-Sanders, L., Kelley, E., Haring, C., Madsen, K., Soto, X., Seven, Y., Hull, K., & Goldstein, H. (2020). Moving forward four words at a time: Effects of a supplemental preschool vocabulary intervention. *Language, Speech, and Hearing Services in Schools*, *51*(1), 165–175. https://doi.org/10.1044/2019_LSHSS-19-00029

Piasta, S. B., Justice, L. M., McGinty, A., Mashburn, A., & Slocum, L. (2015). A comprehensive examination of preschool teachers' implementation fidelity when using a supplemental language and literacy curriculum. *Child Youth Care Forum*, *44*, 731–755. https://doi.org/10.1007/s10566-015-9305-2

Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disabilities. *Learning Disabilities Research & Practice*, *25*(3), 110–123. https://doi.org/10.1111/j.1540-5826.2010.00309.x

Quinn, J. M., Wagner, R. K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development*, *86*(1), 159–175. https://doi.org/10.1111/cdev.12292

Restrepo, M. A., Morgan, G. P., & Thompson, M. S. (2013). The efficacy of a vocabulary intervention for dual language learners with language impairment. *Journal of Speech, Language, and Hearing Research*, *56*(2), 748–765. https://doi.org/10.1044/1092-4388(2012/11-0173)

Seven, Y., Hull, K., Madsen, K., Ferron, J., Peters-Sanders, L., Soto, X., Kelley, E. S., & Goldstein, H. (2020). Classwide extensions of vocabulary intervention improve learning of academic vocabulary by preschoolers. *Journal of Speech, Language, and Hearing Research*, *63*(1), 173–189. https://doi.org/10.1044/2019_JSLHR-19-00052

Silverman, R. (2007). A comparison of three methods of vocabulary instruction during read-alouds in kindergarten. *The Elementary School Journal*, *108*(2), 97–113. https://doi.org/10.1086/525549

Silverman, R., Crandell, J. D., & Carlis, L. (2013). Read alouds and beyond: The effects of read aloud extension activities on vocabulary in Head Start classrooms. *Early Education and Development*, *24*(2), 98–122. https://doi.org/10.1080/10409289.2011.649679

Soto, X., Seven, Y., McKenna, M., Madsen, K., Peters-Sanders, L., Kelley, E. S., & Goldstein, H. (2020). Iterative development of a home review program to promote preschoolers' vocabulary skills: Social validity and learning outcomes. *Language, Speech, and Hearing Services in Schools*, *51*(2), 371–389. https://doi.org/10.1044/2019_LSHSS-19-00011

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development*, *90*(4), 1303–1318. https://doi.org/10.1111/cdev.13072

Suggate, S., Schaughency, E., McAnally, H., & Reese, E. (2018). From infancy to adolescence: The longitudinal links between vocabulary, early literacy skills, oral narrative, and reading comprehension. *Cognitive Development*, *47*, 82–95. https://doi.org/10.1016/j.cogdev.2018.04.005

Tunmer, W. E., & Chapman, J. W. (2012). The Simple View of Reading redux: Vocabulary knowledge and the independent components hypothesis. *Journal of Learning Disabilities*, *45*(5), 453–466. https://doi.org/10.1177/0022219411432685

Vernon-Feagans, L., Bratsch-Hines, M., Reynolds, E., & Willoughby, M. (2020). How early language input varies by race and education and predicts later child language. *Child Development*, *91*(4), 1098–1115. https://doi.org/10.1111/cdev.13281

Wasik, B. A., Hindman, A. H., & Snell, E. K. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly*, *37*, 39–57. https://doi.org/10.1016/j.ecresq.2016.04.003

Wiig, E. H., Secord, W. A., & Semel, E. (2004). *Clinical Evaluation of Language Fundamentals Preschool-2* (2nd ed.). Harcourt Assessment.

Wright, T. S., & Neuman, S. B. (2014). Paucity and disparity in kindergarten oral vocabulary instruction. *Journal of Literacy Research*, *46*(3), 330–357. https://doi.org/10.1177/1086296X14551474

Zucker, T. A., Solari, E. J., Landry, S. H., & Swank, P. R. (2013). Effects of a brief tiered language intervention for prekindergartners at risk. *Early Education & Development*, *24*(3), 366–392. https://doi.org/10.1080/10409289.2012.664763