

Preschool Teachers' Fidelity in Implementing a Vocabulary Intervention

Variation Across Settings and Strategies

Elizabeth Burke Hadley, Molly Scott, Matthew E. Foster, David K. Dickinson, Kathy Hirsh-Pasek, and Roberta Michnick Golinkoff

This study examined preschool teachers' fidelity in implementing a vocabulary intervention. The purpose of the study is to inform the scaling up of vocabulary interventions, identifying strategies that are both feasible for teachers and effective for vocabulary learning. We analyzed data from a vocabulary intervention in which teachers ($n = 10$) taught 80 new target words to children ($n = 138$) during shared book reading (BR) and playful learning experiences (PLEs). Teachers were asked to use core intervention strategies, which included both teacher- and child-focused practices. Results showed that teachers had higher adherence to strategies in BR than in PLEs and for teacher-focused strategies versus child-focused practices. Across settings, teachers' use of core strategies and the use of child-focused practices were significantly related to children's vocabulary outcomes. Results not only suggest the importance of child-focused vocabulary teaching strategies but also indicate the need for additional coaching in this area for teachers. **Key words:** *early childhood, fidelity, oral language, pre-k, shared book reading, vocabulary*

Author Affiliations: Department of Language, Literacy, EdD, Exceptional Education, and Physical Education, University of South Florida, Tampa (Dr Hadley); Teachers College, Columbia University, New York, New York (Dr Scott); Child & Family Studies, University of South Florida, Tampa (Dr Foster); Department of Teaching, Learning, & Diversity, Vanderbilt University, Nashville, Tennessee (Dr Dickinson); Department of Psychology, Temple University, Philadelphia, Pennsylvania (Dr Hirsh-Pasek); and School of Education, University of Delaware, Newark (Dr Golinkoff).

This study was funded by grant R305A150435 from the Institute of Education Sciences, US Department of Education. The opinions expressed are those of the authors and do not represent the views of the Institute or the US Department of Education. The authors thank Molly F. Collins, Emily Hopkins, Rebecca Dore, and Jessica Lawson-Adams for their work on this grant.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.topicsinlanguageorders.com).

VOCABULARY INTERVENTION researchers have found that young children can and do learn new words from explicit instruction in early childhood settings (Marullis & Neuman, 2010). However, it has proven difficult to translate smaller-scale vocabulary interventions implemented by researchers under favorable conditions (i.e., efficacy trials) into

Author disclosures can be found at <http://links.lww.com/TLD/A99>.

Corresponding Author: Elizabeth Burke Hadley, PhD, Department of Language, Literacy, EdD, Exceptional Education, and Physical Education, College of Education, University of South Florida, 4202 E Fowler Ave., EDU 105 Tampa, FL 33620 (hadleye@usf.edu).

DOI: 10.1097/TLD.0000000000000294

real-world classroom contexts in which teachers use intervention methods over the long term (i.e., effectiveness trials). Although many vocabulary interventions show robust effect sizes in the short term on proximal measures of word learning (Marulis & Neuman, 2010), the complexity of methods and time requirements of many interventions may not be feasible for teachers (Dickinson, 2011; Neuman & Danielson, 2021), making it difficult to obtain longer-term effects on general word knowledge (e.g., Mendive et al., 2016). Researchers need clear information about how to optimize instructional methods and where to direct resources to provide targeted support for teachers.

Based on a recent study of a preschool vocabulary intervention in which children showed significant growth in their knowledge of target words (Scott et al., under review), we sought to investigate implementation fidelity (IF) to better understand both the feasibility and the efficacy of specific practices. By investigating which practices teachers choose to implement, and which practices are especially effective for word learning, we aim to inform the design of efficient and effective interventions that teachers can use as part of their daily classroom routine.

A FINE-GRAINED APPROACH TO MEASURING TREATMENT ADHERENCE

Implementation fidelity refers to the degree to which an intervention is delivered as intended. Implementation fidelity is most often operationalized as treatment adherence, or the extent to which the core practices that are essential to treatment effects are implemented (e.g., Capin et al., 2018). Other dimensions of IF include quality (i.e., how well the intervention is implemented), dosage, program differentiation, and participant responsiveness (Dane & Schneider, 1998). Although treatment adherence is an essential element of high-quality studies (Gersten et al.,

2005), few early literacy interventions report data on this dimension of IF (Capin et al., 2018).

Interventions are typically designed with the assumption that, if teachers adhere to intervention procedures, there will be growth in children's skills. However, although early childhood language interventions have indeed found that better treatment adherence is associated with better child outcomes (Justice et al., 2008; Wasik & Hindman, 2011), others have found no relation between adherence and child outcomes (Hamre et al., 2010; Mendive et al., 2016). There are several potential explanations for this lack of relation. First, the intervention procedures may not be effective, and so better adherence makes no difference. Second, strict adherence (i.e., "sticking to the script") may not be optimal, as higher-quality instruction may necessitate adaptations to accommodate students' needs (Harn et al., 2013). A final potential explanation, and one we explore here, is that many measures of treatment adherence do not adequately capture variation in teachers' enactment of intervention practices, instead relying on brief checklists that identify whether central intervention components were implemented (e.g., Nelson et al., 2012). For example, Hamre et al. (2010) measured IF overall using several dimensions, but treatment adherence was measured by just four items: whether the teacher generally followed the script, had the materials available, used the materials appropriately, and completed the lesson components. There was no relationship found between teachers' adherence and children's outcomes, perhaps because this checklist was not specific to the intervention strategies. Using such general adherence measures may not capture some of the more subtle variations in teachers' implementation of core intervention practices that relate to children's outcomes. Tracking more fine-grained patterns of adherence is likely to reveal stronger relations to child outcomes. For example, Wasik and Hindman (2011) measured specific practices used in

the intervention (e.g., open-ended questions, vocabulary definitions), which explained 29% of the variance in children's receptive vocabulary.

In addition to tracking more proximal, fine-grained practices, Neugebauer et al. (2021) suggest reporting adherence scores for individual practices or groups of practices rather than a single, aggregated score. Using aggregated fidelity scores is typical for intervention studies, but they may mask how specific practices contribute to the overall success of the intervention. Furthermore, given research that suggests certain language practices may be more or less successful depending on the activity in which they are used (Hadley et al., 2022), aggregating across activity settings may obscure practices that are feasible and/or successful in one setting (e.g., one-on-one shared book reading [BR]) but not in another (e.g., group read alouds). Detailed adherence data also may identify practices that teachers consistently choose not to implement. If researchers identify practices that teachers consistently omit, they can then dig deeper to understand potential barriers. For example, the strategies may be impractical, too difficult, or conflict with teachers' beliefs about learning (Pence et al., 2008; Zucker et al., 2021). In contrast, if teachers use certain practices with nearly perfect adherence, they may need less intensive coaching and support to use them.

Finally, it is important to note that there is some overlap in how we conceptualize adherence with how others have conceptualized the related IF dimension of quality. However, operationalizing what "quality" instruction looks like in relation to a specific intervention has proved to be an amorphous and elusive task (Darrow, 2013). Instead, we break down the construct of adherence, focusing on the feasibility and efficacy of specific, contextualized practices, which may in turn help pinpoint the behaviors that serve as markers of high-quality implementation.

COMPONENTS OF THE LANGUAGE FOR READING INTERVENTION THAT SUPPORT VOCABULARY LEARNING

The Language for Reading (L4R) intervention was iteratively developed over a 3-year time span to support vocabulary development in preschoolers from families living in poverty (Scott et al., under review), which extended procedures developed in a previous 3-year study with similar aims (Dickinson et al., 2019b; Toub et al., 2018). One of the central innovations in the L4R method is the pairing of shared BR and play activities to build both breadth (i.e., knowing a relatively large number of words) and depth of vocabulary knowledge (i.e., knowing individual words "better," or having highly differentiated semantic networks for words [Hadley et al., 2016; Hadley, Dickinson, Hirsh-Pasek et al., 2019]). Prior research suggests that teaching new vocabulary words in the context of shared BR is effective, and yet may not be powerful enough to support robust vocabulary learning (Marulis & Neuman, 2010; Noble et al., 2019). We sought to take advantage of play as a developmentally appropriate and engaging context for word learning that could complement BR. Our aim in pairing BR and play was to first establish fast-mapped representations of words in initial BR sessions through explicit instruction and discussion (Dickinson et al., 2019b) and then use that learning to springboard into the play setting, where children could actively process nuances of word meaning (e.g., McKeown & Beck, 2014; Toub et al., 2018). Our model built in repeated encounters with new words in varied contexts (moving between BR and play sessions) so that children could gradually accumulate deeper layers of word knowledge (Bolger et al., 2008; Frishkoff et al., 2011).

In previous studies, we found that play guided by a teacher was most effective for word learning, as opposed to unstructured play (Toub et al., 2018), but teachers also told us that they were uncomfortable with the

lack of structure in small-group sociodramatic play activities (Dickinson et al., 2019a). Based on these findings, we not only designed a tool kit of playful learning experiences (PLEs) for the present study that were engaging and flexible, but also provided a more consistent structure.

To balance controlling for threats to internal validity and establishing the external validity of our treatment procedures for the classroom context, we took a “soft-scripted” approach (Blases et al., 2018; Neuman et al., 2021) in which we listed instructional strategies, word definitions, and questions for children on a one-page guidance card, customized for each BR and PLE session (see Supplemental Digital Content Figure S2, available at: <http://links.lww.com/TLD/A93>), but did not restrict teachers to using exact scripted language. To keep the guidance cards succinct, we listed only the key strategies that we believed were essential for implementing the intervention successfully. In this article, we call these *core* strategies. These *core* strategies included both *teacher-focused* strategies in which teachers shared explicit definitions for words along with visual and gestural supports and *child-focused* strategies in which teachers invited children to use new words and answer open-ended questions. We used a gradual release of responsibility model in which teachers took a more didactic approach in early sessions and then invited children to take the lead in sharing information about target words as they learned more.

PRESENT STUDY

The intervention we describe here is drawn from a larger 3-year project. Previous articles from the L4R project have reported on (1) the efficacy of a digital app (one of the PLEs) to support vocabulary learning (Dore et al., 2019), (2) the relationship between child and word characteristics and learning of expressive vocabulary in the BR setting (Hadley et al., 2021), and (3) the main effects and overall efficacy of the intervention

in Year 2 (Scott et al., under review). Using a within-subjects design, the third study found that children demonstrated significantly better learning of target words compared with control words on both receptive (odds ratio = 1.08) and expressive (odds ratio = 26.13) measures of target word knowledge. The present study also uses data from Year 2 of the intervention but represents a secondary analysis designed to explore variations in teacher fidelity and investigate relationships between teacher practices and child vocabulary outcomes.

We address the following research questions: (1) Did teachers’ adherence to a vocabulary intervention vary for *core* versus *bonus* strategies (i.e., additional vocabulary instruction beyond *core* strategies), for BR versus PLE settings, and for *teacher-focused* versus *child-focused* strategies? (2) How did teachers’ implemented instructional strategies relate to children’s expressive and receptive knowledge of words at posttest?

METHOD

Participants

The current study was conducted in 10 preschool classrooms: five in a Head Start program in a northeastern U.S. city and five in a state-funded preschool program in a southern U.S. city. There were 138 child participants (74 girls; $M_{\text{age}} [SD] = 57.88 [5.44]$ months) and their mean (SD) norm-referenced standard score for receptive vocabulary was 94.32 (15.42). According to parental report, 61.59% of children were Black/African American, 15.22% Latinx/Hispanic, 10.14% were of another race or more than one race, 5.80% were White, and 7.25% did not respond. The highest level of maternal education reported for most of our sample was high school (44.20%), with 37.68% completing some college or trade school, 7.97% completing college, and 10.14% not responding. Languages other than English spoken at home included Spanish (13.77%), Arabic (7.25%), and other non-English languages (4.35%). Classroom

teachers administered the intervention. One teacher had an associate's degree, six had bachelor's degrees, and three had master's degrees in early childhood education or related fields.

Intervention description

Teachers taught a total of four units and 80 target words over the course of the school year.

During each 4-week unit, children participated in a large-group BR session and a PLE activity and were taught a set of 20 target vocabulary words across both activities. Words consisted of concrete nouns, abstract nouns, and verbs, and were selected to be uncommon for preschoolers to ensure that participants were unlikely to encounter words outside of the intervention (e.g., *nudge*, *task*, *rodent*; see the study by Scott et al. [under review] for additional information on vocabulary selection). Instruction was repeated with minor adaptations across weeks to encourage children to participate more as they learned the words.

Teacher training and professional development

Teachers participated in a research staff-led orientation that introduced the project goals and the vocabulary instructional procedures. Trained coaches supported teachers throughout the intervention. Coaches met with teachers before each unit to review materials and procedures. Coaches recorded observations of teachers implementing the intervention and gave teachers feedback about what went well and what could be improved. Coaches provided more support at the beginning of the intervention, coaching six sessions in Unit 1, gradually fading support to four sessions in Unit 2 and three sessions in Units 3 and 4.

Shared BR and PLE materials

A team of researchers and teachers designed intervention materials collaboratively, which were piloted in classrooms in a previous phase of the study. Units were designed

around four children's books: *Pumpkin Soup* by Helen Cooper, *Farmer Duck* by Martin Waddell and Helen Oxenbury, *Pearl's New Skates* by Holly Keller, and *Raccoon on His Own* by Jim Arnosky.

One of the original aims of the study was to develop a tool kit of playful vocabulary learning experiences (i.e., PLEs). Teachers were assigned randomly to different PLEs in each unit (see Supplemental Digital Content Table S1, available at: <http://links.lww.com/TLD/A92>), all of which included guessing and review of the target words. The PLEs included (1) large-group games including adaptations of classic children's games such as Hot Potato, played as a whole class; (2) small-group games (with three to five children) including adaptations of board games and bingo; (3) sociodramatic play (with three to five children) including reenacting the book stories with popsicle stick characters or magnets; and (4) music PLEs in which children sang along with classic children's songs, rewritten to incorporate target words. Individual teachers used the *same* PLE for all activities in each unit; when they started a new unit, they were asked to use a different PLE according to the assignment schedule.

Vocabulary instruction during shared BR and PLEs

Each target word was taught three times during each BR and PLE session in brief teaching episodes (i.e., "learning moments"). The exact nature of this instruction varied on the basis of the activity and the week of instruction. For example, the first time teachers introduced target words in a unit, we asked them to say the word, define the word, and make a gesture for the word. By the third and fourth BR sessions, we asked teachers to gradually release responsibility to children by asking children to guess the target words and to give definitions.

BR instruction

Before reading, the teachers showed children a picture representing each word, encouraged children to guess the word, and

reviewed each definition. During the BR session, after reading the target word, the teachers provided a definition and gesture (e.g., for *rung*, the definition was “the part of the ladder that you step on,” and the gesture was pretending to climb a ladder) and encouraged the children to repeat after them. The third learning moment occurred after BR when the teachers asked the children scripted questions, including a yes/no and an open-ended question for each target word (see Supplemental Digital Content Figure S2, available at: <http://links.lww.com/TLD/A93>).

PLE instruction

The first learning moment, similar to BR, entailed a picture card review prior to the activity. The second and third learning moments both occurred during the PLEs and varied slightly based on the activity but included saying the word, giving the definition, and sharing a picture or gesture. For example, in a large group fly-swatter game, the teacher began by propping up three picture cards at the children’s eye level. One at a time, the students were asked to swat the picture that matched the target word said by the teacher. All children were prompted to repeat the target word, gesture, and definition. In the second round, the teacher prompted the children with the definition of the word. After the child swatted the correct picture, the teacher encouraged all students to say and define the word.

Guidance cards

To assist teachers in completing these instructional strategies, guidance cards were provided for each BR and PLE session. These cards listed the target words, definitions, and which strategies teachers should use for each of the three learning moments.

Measures

Children were pre- and posttested individually on the target vocabulary words in each unit using receptive and expressive measures of vocabulary knowledge (for a total of eight waves of vocabulary assessments).

To measure receptive word knowledge, the children heard an audio recording of each word and were told to select the matching image on a touch screen tablet. Each trial contained the correct image, a conceptual foil, a thematic foil, and a phonological foil. Sample-based internal consistency (Cronbach’s α) for all receptive vocabulary items was 0.89.

To measure expressive word knowledge, the New Word Definition Test-Modified (Hadley et al., 2016) asked the children to say and show (via gesture) everything they knew about a word. Participants’ responses were coded for information units, or pieces of semantic knowledge, about the word, including synonyms, antonyms, gestures, functional or perceptual features, part-whole relationships, meaningful context, or basic context. After a training period, the research assistants reached at least 90% agreement with a doctoral student anchor coder. The research assistant and the anchor coder double-coded every fifth set of child responses. When 90% agreement was not attained, the anchor coder’s scoring was used and discrepancies were discussed until agreement was reached. Sample-based internal consistency for all expressive vocabulary items was 0.96.

Children’s general receptive vocabulary knowledge was assessed using the Peabody Picture Vocabulary Test (PPVT-IV; Dunn & Dunn, 2007), which was used as a covariate in the present study. Sample-based internal consistency for this measure was 0.94.

Teacher fidelity to the intervention

Using the guidance cards and the training materials, we generated a list of nine *core* strategies that we saw as the active ingredients of the intervention (see Table 1). These also were the strategies that were listed on the guidance card for teachers to complete (see Supplemental Digital Content Figure S2, available at: <http://links.lww.com/TLD/A93>). Teachers were not expected to implement all nine *core* strategies for every target word and every learning moment, as we built in variation across learning moments and instructional sessions. We did not refer to these

Table 1. Average percentage of core and bonus strategies (top panel) and core strategies by PLE types (bottom panel) implemented by teachers (*n* = 10) during BR and PLEs

Strategy Type	BR		PLEs	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Core and bonus strategies</i>				
Core strategies (all)	69.0	18.0	56.0	17.0
Teacher-focused strategies	82.0	14.0	70.0	16.0
Shows picture of word	95.0	8.0	70.0	14.0
Says target word	94.0	6.0	83.0	13.0
Gives definition	83.0	15.0	72.0	20.0
Makes gesture for word	70.0	26.0	54.0	37.0
Child-focused strategies	59.0	21.0	27.0	17.0
Asks children to repeat word	65.0	27.0	10.0	30.0
Asks children to define word	56.0	32.0	2.0	2.0
Asks children to gesture	36.0	26.0	48.0	30.0
Asks children to guess word	86.0	23.0		
Asks closed-ended question	79.0	26.0		
Asks open-ended question	51.0	32.0		
Asks children to show picture			50.0	30.0
Bonus strategies (all)	15.0	7.0	14.0	13.0
<i>Core strategies by PLE type</i>				
Small-group games (<i>n</i> = 5)			70.0	14.0
Music (<i>n</i> = 5)			66.0	14.0
Sociodramatic play (<i>n</i> = 4)			58.0	15.0
Large-group games (<i>n</i> = 5)			49.0	16.0

Note. Mean percentage scores represent how often teachers implemented a strategy, divided by the number of times teachers were instructed to use that strategy by guidance cards. Blank cells indicate that a particular strategy was not used in that setting. Bonus strategies = “extra” instructional strategies that teachers provided on their own; BR = book reading; Child-focused strategies = core strategies in which a teacher asks a child to do something (e.g., encourages child to gesture); Core strategies = strategies that were essential to the intervention; PLEs = playful learning experiences; Teacher-focused strategies = core strategies implemented solely by teachers (e.g., giving a definition).

as *core* strategies with teachers (as this is a post hoc analysis), but we did explicitly list these strategies on each guidance card.

A list of *bonus* strategies also was generated to capture any additional vocabulary instruction teachers provided for target words beyond the *core* strategies. For example, teachers sometimes provided additional semantic information (e.g., “usually a *banister* is made out of wood”), connected a word to children’s real-life experiences, gave a physical example of a word (e.g., pointing to the coffee in their cup to illustrate the word “liquid”), or used PLE-specific strategies to support word meaning (e.g., asking children to act out the target words during

sociodramatic play). This instruction was technically not “required” (i.e., not listed on the guidance cards) but was potentially important in understanding the full scope of teachers’ vocabulary instruction during the intervention.

Core and *bonus* strategies were coded as either present or absent for each of the three learning moments and for each individual target word. *Core* strategies were then analyzed in greater detail by dividing them into strategies that were *teacher-focused* strategies that did not require child involvement and *child-focused* strategies in which teachers elicited children’s engagement with the vocabulary words (see Table 1). How frequently teachers

used the *core* strategies for each of the PLE types also was investigated.

Teacher fidelity scoring

The research team coded videos of intervention sessions to evaluate teachers' treatment adherence. One BR video and one PLE video from each unit were randomly selected for each teacher, equaling a total of 80 videos (eight per teacher), distributed roughly evenly across units. A graduate student double-coded 20% of the videos. If the percent agreement with the anchor coder was below 75%, the coders resolved their disagreements through discussion before proceeding.

During each BR and each PLE activity, the teachers were asked to teach 10 target words three times, using around five core strategies in each learning moment. For example, we asked teachers to use the following five core strategies for the word "chaos" during learning moment number 2 in BR: (1) point to the picture showing "chaos," (2) say the word; (3) define the word ("a time when everyone and everything is out of control"); (4) children repeat the word; and (5) teacher and children do a gesture. Thus, the teachers were asked to implement about 150 instances of *core* strategies during an activity, with slight variations across activities. To compare fidelity scores across different activity types, we used a percentage score that accounted for the actual number of *core* strategy opportunities listed on the guidance card for each BR or PLE activity. This ensured that teachers were not penalized for not completing *core* strategies that did not apply to a particular learning moment or activity. For *bonus* strategies, we divided the number of strategies implemented by the number of times teachers could have potentially used a bonus strategy (i.e., the number of word exposures), yielding an overall percentage of how often teachers used *bonus* strategies when they had the opportunity to do so. Fidelity was calculated separately for BR and PLEs so that we could compare strategy use across contexts.

RESULTS

Preanalysis data inspections

The descriptive statistics for the 138 child participants are displayed in Table 2. All variables were normally distributed, with the exception of the expressive vocabulary measure, which showed evidence of floor effects at pretest. All bivariate relations were linear and in the expected direction (for the correlation matrix of key variables, see Supplemental Digital Content Table S3, available at: <http://links.lww.com/TLD/A94>).

Because of the frequency of pretest and posttest assessments (i.e., eight waves), some children were absent for one or more assessment sessions. The amount of missingness for each measure is reported in Table 2. We used multiple imputation to avoid biased estimates and loss of statistical power from listwise deletion of cases with missing data (see Collins et al., 2001; Graham, 2012). Simulation studies indicate that even when 50% of the data in the dependent variable is missing, multiple imputation is a reliable and valid approach for handling missing data within analyses (Collins et al., 2001). To achieve statistical power comparable with using full information methods, 40 data sets were multiple imputed in SPSS (Version 26). In the present study, the missing data in the imputed data sets were based on participant age as well as all variables related to both research questions (i.e., PPVT-4 raw scores, eight pretest and eight posttest measures of receptive and expressive vocabulary, and 16 measures of teacher IF).

Data analytic overview

To investigate variations in fidelity between BR and PLE contexts, for different strategy types, and for different kinds of PLEs, we calculated the percentage of *core* and *bonus* strategies in BR and PLE contexts. To examine the relations between teacher fidelity practices and child outcomes, hierarchical linear models (HLMs; Raudenbush & Bryk, 2002) were estimated to account for

Table 2. Descriptive statistics for the child participants ($n = 138$) at pretest and posttest

	Pretest					Posttest				
	<i>N</i>	% Missing	Mean	<i>SD</i>	Skew	<i>N</i>	% Missing	Mean	<i>SD</i>	Skew
PPVT	115	17	70.06	20.54	-0.70	- ^a		- ^a	- ^a	- ^a
Receptive Unit 1	129	7	4.48	1.98	0.49	126	9	9.17	4.28	0.19
Receptive Unit 2	88	36	4.84	1.75	0.03	89	35	8.76	4.36	0.40
Receptive Unit 3	117	15	5.11	1.92	0.42	112	19	9.04	4.46	0.51
Receptive Unit 4	128	6	4.87	1.90	0.33	127	8	9.55	4.26	0.33
Expressive Unit 1	125	9	0.30	0.70	2.90	121	12	6.69	7.58	1.23
Expressive Unit 2	125	9	0.60	1.11	2.27	126	9	9.03	9.33	0.98
Expressive Unit 3	115	17	0.35	1.05	4.07	110	20	8.11	9.52	1.11
Expressive Unit 4	120	13	0.85	1.54	2.35	123	11	10.49	10.20	0.99

Note. PPVT = Peabody Picture Vocabulary Test.

^aNot administered at posttest. Data are reported prior to imputation.

the clustered nature of our data (students nested within classrooms). First, after the data were imputed and integrated, sum scores were created for all target receptive and expressive vocabulary measures at pretest and posttest across units. Next, we estimated an unconditional two-level model with random intercepts for both of our outcome measures to compute the intraclass correlation coefficients (ICCs) for Level 1 (students within classrooms) and Level 2 (between classrooms). The ICC estimates indicated that a substantial proportion of the variance in expressive (48.72%) and receptive (39.34%) vocabulary outcomes was at Level 2 (between classrooms). Thus, a considerable amount of the variance in outcomes was accounted for by differences between classrooms, in addition to the differences between children. Given this clustering in the data, we used two-level conditional HLMs to test the relations between teacher vocabulary strategies and child vocabulary. All predictors were transformed into *z* scores to allow for comparison of coefficients across models. For each HLM related to research Question 2, the proportion of variance explained is reported in Table 3. Because there is currently no accepted standard for measures of R^2 in hierarchical linear modeling, a common approach is to report the proportion of the variance explained at Level 1 separately from the pro-

portion of the variance accounted for at Level 2 (Snijders & Bosker, 1999). Therefore, we report the proportion of explained variance for each model, as compared with the unconditional model, at both Level 1 and Level 2.

Variation in teachers' adherence to the intervention

Table 1 presents the percentage of the strategies implemented (based on the guidance cards). Overall, teachers' fidelity for the *core* strategies was higher during BR (69%) than PLEs (56%). As can be seen in Table 1, IF for specific *core* strategies ranged from 95% (showing a picture of the word) to 51% (asking an open-ended question) in BR, and 83% (saying the target word) to 2% (asking children to define the word) in PLEs (see Table 1). Teachers' use of *bonus* strategies, or additional instruction above and beyond the *core* strategies, was relatively infrequent (15% in BR; 14% in PLEs).

Next, we examined teachers' implementation of *teacher-focused* and *child-focused* strategies. Overall, teachers had substantially higher fidelity for *teacher-focused* strategies (82% in BR, 70% in PLEs) than for *child-focused* strategies (59% in BR, 27% in PLEs). We also investigated whether teachers' use of the *core* strategies varied across PLE types. These descriptive results should be interpreted with caution, as not all teachers

Table 3. Parameter estimates and standard errors (within parentheses) for relations between teachers' use of core strategies and children's vocabulary outcomes in BR and PLE settings

Variable	Model 1		Model 2		Model 3		Model 4	
	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary
Fixed effects								
Intercept	-0.103 (0.242)	-0.065 (0.216)	-0.043 (0.152)	-0.022 (0.202)	-0.061 (0.160)	-0.013 (0.150)	-0.046 (0.159)	-0.032 (0.137)
PPVT			0.432** (0.050)	0.642** (0.049)	0.441** (0.053)	0.638** (0.051)	0.435** (0.053)	0.648** (0.051)
Pretest			0.236** (0.053)	-0.078 (0.052)	0.228** (0.053)	-0.078 (0.051)	0.232** (0.053)	-0.089 (0.052)
Core BR			0.369* (0.149)	0.111 (0.194)			0.248 (0.217)	-0.327 (0.186)
Core PLE					0.355 (0.157)	0.391* (0.148)	0.174 (0.222)	0.630* (0.190)
Random effects								
Child variance	0.571** (0.071)	0.643** (0.080)	0.259** (0.033)	0.285** (0.036)	0.259** (0.033)	0.285** (0.036)	0.259** (0.033)	0.285** (0.036)
Teacher variance	0.542 (0.279)	0.417 (0.221)	0.219 (0.119)	0.384 (0.204)	0.235 (0.131)	0.203 (0.114)	0.229 (0.134)	0.158 (0.096)
Explained variance								
Child variance			0.571	0.399	0.556	0.562	0.562	0.602
Teacher variance			0.550	0.487	0.548	0.509	0.548	0.514

Note. * $p < .05$. ** $p < .001$. All continuous variables are z scores. Explained variance for Models 2–4 is the proportion of the variance explained by the predictors compared with Model 1 (unconditional model). BR = book reading; Core = core vocabulary instructional strategies that were essential to intervention; PLE = playful learning experiences; PPVT = Peabody Picture Vocabulary Test.

implemented all PLE types (see Supplemental Digital Content Table S1, available at: <http://links.lww.com/TLD/A92>). Overall, there was variability in teachers' use of *core* strategies across PLE types, with teachers implementing *core* strategies with 70% fidelity during small-group games, 66% in music, 58% in sociodramatic play, and 49% in large-group games.

Relations between teacher strategies and child vocabulary outcomes

We conducted preliminary correlational analyses to determine which covariates should be included in our final models (see Supplemental Digital Content Table S3, available at: <http://links.lww.com/TLD/A94>). Children's pretest scores, age, and PPVT scores were included as covariates in our initial models (not reported). However, age was not included in subsequent models as it was not a consistently significant predictor of children's outcomes. Similarly, *bonus* strategies were only weakly related to posttest vocabulary scores and were not included in subsequent analyses. Finally, we did not use the *PLE-type* variable in further analyses, as only four to five teachers implemented each PLE type.

We first tested whether each individual fidelity variable (*core* BR and *core* PLE strategies) was a predictor of children's expressive and receptive vocabulary. The HLM accounts for children ($n = 138$) nested within classrooms ($n = 10$). We included PPVT and pretest scores as covariates to control for general vocabulary knowledge and children's baseline knowledge of target words. Results indicated that *core* BR strategies were significant positive predictors of expressive ($\gamma_{01} = 0.369$) but not receptive vocabulary (Table 3, Model 2). *Core* PLE strategies were significant predictors of receptive ($\gamma_{02} = 0.391$) but not expressive vocabulary (Table 3, Model 3).

Next, we included both *core* BR and *core* PLE strategies in the same model to determine whether either variable contributed to vocabulary learning above and beyond the other (Table 3, Model 4). The HLM revealed that *core* PLE strategies were significantly related to receptive vocabulary posttest scores

($\gamma_{02} = 0.630$), even after controlling for *core* BR strategies. However, Model 4 explains only a slightly larger percentage of the variance in outcome scores as the single-predictor models (Models 2 and 3), suggesting that these predictors have substantial overlap in the variance in child outcomes they explain.

We tested whether *teacher-focused* strategies in BR and PLE and *child-focused* strategies in BR and PLE were predictive of child vocabulary outcomes (Models 2-5, Table 4). As in the previous models, we first tested each of the key teacher fidelity predictors in separate models. Results showed that *teacher-focused* strategies during PLEs, but not BR, were significantly related to receptive vocabulary ($\gamma_{04} = 0.365$). There also was a pattern of positive significant relationships between *child-focused* strategies and child vocabulary outcomes. *Child-focused* strategies during BR were significantly positively related to expressive vocabulary ($\gamma_{04} = 0.383$), and *child-focused* strategies during PLEs were significantly positively related to receptive vocabulary ($\gamma_{05} = 0.383$). Finally, we included all four predictor variables in the same model (Model 6). When including all teacher-level variables, none were significant predictors of expressive or receptive vocabulary. These results indicate that none of the strategy types have a unique predictive relationship with vocabulary when controlling for the other strategy types.

DISCUSSION

A robust body of research indicates both the promise of early vocabulary instruction and the difficulty in scaling up vocabulary interventions for use in classroom settings (Dickinson, Freiberg, et al., 2011). An urgent priority is to investigate which procedures teachers can use successfully in authentic classroom settings. The present study adds to this area of research by examining fine-grained data on teachers' implementation of a vocabulary intervention across book reading and play settings. Implementation fidelity data revealed that teachers showed higher

Table 4. Hierarchical linear modeling results for relations between use of teacher-focused and child-focused strategies and children’s vocabulary outcomes in BR and PLE settings

Variable	Model 1		Model 2		Model 3		Model 4		Model 5		Model 6	
	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary	Exp. Vocabulary	Rec. Vocabulary
Fixed effects												
Intercept	-0.103 (0.242)	-0.065 (0.216)	-0.054 (0.181)	-0.019 (0.201)	-0.042 (0.146)	-0.024 (0.203)	-0.078 (0.180)	-0.029 (0.158)	-0.040 (0.163)	0.010 (0.154)	-0.033 (0.149)	-0.020 (0.164)
PPVT			0.436** (0.053)	0.642** (0.051)	0.431** (0.053)	0.643** (0.051)	0.440** (0.053)	0.642** (0.051)	0.440** (0.053)	0.638** (0.051)	0.430** (0.053)	0.646** (0.051)
Pretest			0.233** (0.053)	-0.078 (0.051)	0.238** (0.053)	-0.078 (0.052)	0.231** (0.053)	-0.081 (0.051)	0.230** (0.053)	-0.077 (0.051)	0.239** (0.053)	-0.087 (0.052)
TF-BR			0.283 (0.184)	0.144 (0.203)							0.015 (0.217)	-0.157 (0.322)
CF-BR					0.383* (0.136)	0.091 (0.189)					0.255 (0.199)	-0.104 (0.270)
TF-PLE							0.273 (0.176)	0.365* (0.155)			0.015 (0.217)	0.329 (0.238)
CF-PLE									0.351 (0.161)	0.383* (0.153)	0.255 (0.199)	0.301 (0.219)
Random effects												
Child variance	0.571** (0.071)	0.643** (0.080)	0.259** (0.033)	0.285** (0.035)	0.259** (0.033)	0.285** (0.036)	0.259** (0.033)	0.285** (0.036)	0.239** (0.033)	0.285** (0.036)	0.259** (0.033)	0.285** (0.036)
Teacher variance	0.542 (0.279)	0.417 (0.221)	0.306 (0.165)	0.376 (0.200)	0.192 (0.106)	0.388 (0.207)	0.302 (0.164)	0.229 (0.126)	0.243 (0.134)	0.213 (0.119)	0.196 (0.137)	0.239 (0.165)
Explained variance												
Child variance	0.492	0.406	0.492	0.406	0.595	0.395	0.496	0.538	0.567	0.553	0.591	0.529
Teacher variance	0.539	0.488	0.539	0.488	0.553	0.487	0.540	0.506	0.580	0.508	0.552	0.505

Note. * $p < .05$. ** $p < .001$. All continuous variables are z scores. Explained variance for Models 2–6 is the proportion of the variance explained by the predictors compared with Model 1 (unconditional model). BR = book-reading; CF-BR = child-focused book-reading strategies; CF-PLE = child-focused playful learning experience strategies; PLE = playful learning experiences; PPVT = Peabody Picture Vocabulary Test; TF-BR = teacher-focused book-reading strategies; TF-PLE = teacher-focused playful learning experience strategies.

adherence to practices that were teacher-focused rather than child-focused, and that teachers often omitted open-ended questions and other practices that conferred conversational control to children. Teachers also achieved better adherence in BR than PLE contexts. In investigating the efficacy of specific intervention practices, we found that the key active ingredients in the present intervention were the use of *core* intervention practices during PLEs, the use of *teacher-focused* strategies during PLEs, and the use of *child-focused* strategies in both settings. We unpack these findings in more detail in the following sections.

Teachers' adherence varies across settings and by strategy type

Teachers' overall adherence to *core* intervention strategies (69% in BR, 56% in PLEs) was lower than that reported in the systematic review of IF in early literacy interventions (93%) by Capin et al. (2018) but not out of line with other early language and literacy studies that used fine-grained measures of IF (e.g., Mendive et al., 2016). Indeed, considering that we measured teachers' use of more than 150 instances of instructional strategies per session, 100% adherence is likely neither expected nor the ultimate goal, as teachers may have skipped certain strategies for pedagogically sound reasons. For example, children might chime in with a target word before the teacher asks, making the *core* strategy of asking children to guess the target word superfluous.

Examining variations in adherence across individual practices revealed several illuminating patterns. First, teacher-focused strategies, such as showing a picture of the target word, were implemented with nearly perfect adherence in BR and high adherence in PLEs. Adherence dropped below 80% for child-focused practices, where teachers asked children to do something with words (e.g., repeat or define words). Within child-focused practices, teachers had higher adherence for questions or requests that required a one-word response from children (e.g., guessing

target words; 86% in BR) than those that asked children to share longer responses (e.g., open-ended questions; 51% in BR) or asked children to define words (56% in BR, 2% in PLE). Open-ended questions also had the largest standard deviation of all practices in BR, indicating that this was highly variable among the teachers in our sample.

Why, then, did teachers omit open-ended questions and requests for word definitions more than other practices? Prior research suggests that teachers are often more comfortable with, and more effective at, supplying scripted information when implementing an intervention than scaffolding and engaging in discussion, which requires deep knowledge of intervention principles and skill at impromptu language interactions (Zucker et al., 2021). It also is possible that teachers worried that asking open-ended questions might derail the perceived goal of the sessions: delivering a robust "dose" of word meaning information. Hindman et al. (2019) found that Head Start teachers rarely asked open-ended questions during BR and suggested that teachers might be uncomfortable with ceding control of the conversation to potentially off-topic contributions from children. Previous research has highlighted the importance of skillful classroom management as a factor in the success of early childhood language interventions (Cabell et al., 2019), suggesting that interventions are optimized in classrooms where teachers can manage behavior concerns and potential disruptions from children. Thus, teachers' omission of some of the more open-ended, conversation-starting strategies may relate to teachers' relative discomfort with extended conversations in these activity settings, either due to classroom management or a need for additional support in effectively scaffolding conversations.

Teachers' adherence also varied across contexts, with overall adherence higher during BR than in PLEs. This likely signals the relative unfamiliarity of embedding teaching strategies in the PLE activities, whereas BR is a typical part of early childhood practice. In

an early literacy and language intervention, Mendive et al. (2016) found that teachers spent the most time implementing activities that were already part of their daily routine. Book reading implementation may have been particularly straightforward for teachers as it involved what was a relatively small iteration of teachers' already existing BR practices. Although we designed PLEs to be more structured and teacher-friendly in response to teacher feedback from a prior study (Dickinson et al., 2019a), the PLEs still involved spontaneous, playful practices such as turn-taking, singing, and acting, adding an element of unpredictability, which may have been challenging.

Core practices in PLEs and child-focused strategies as active ingredients in vocabulary learning

Teachers' use of *core* BR practices were significantly and positively related to expressive vocabulary learning, and their use of *core* PLE practices were significantly and positively related to receptive vocabulary learning. These findings align with previous studies that found that higher adherence to *core* intervention practices was associated with higher vocabulary outcomes (Mendive et al., 2016; Wasik & Hindman, 2011). These results also confirm the importance of measuring adherence by identifying and tracking the *core* components of an intervention hypothesized to be essential to achieving effects, rather than more general "checklist" items such as having necessary materials, generally following a script, or completing an activity log for the day (Nelson et al., 2012). The finding that *core* BR practices related to expressive vocabulary, whereas *core* PLE practices related to receptive vocabulary, was somewhat unexpected, as the PLE sessions typically gave children more opportunities to use words than BR sessions. Further research is needed to fully explore how different instructional strategies relate to expressive versus receptive vocabulary learning.

When *core* PLE and *core* BR strategies were examined in the same model, *core* PLE strate-

gies emerged as a significant positive predictor of receptive vocabulary above and beyond that accounted for by *core* BR strategies. This finding suggests that giving children the chance to experiment with new target words in a hands-on, playful context, once they have been introduced during BR, may help refine and further differentiate children's semantic networks for new words. In other words, this finding points to the use of the *core* PLE strategies as an active ingredient essential to the overall success of the intervention.

An additional active ingredient was revealed by our analysis of *teacher-* and *child-focused core* strategies. Analyses revealed that *teacher-focused* strategies during PLEs were a positive predictor of children's receptive vocabulary. *Child-focused* strategies in both BR and PLEs also were positive predictors of children's vocabulary outcomes. Considered together, these findings suggest that the *teacher-focused* instruction, especially during PLEs, may have provided a knowledge base that in turn helped children successfully participate in *child-focused* interactions about words with their teachers in later BR and PLE sessions. The importance of *child-focused* strategies for vocabulary learning in both settings may signal that children learned the most when their teachers were comfortable asking questions and actively engaging children in discussions about word meaning. Our finding supports and extends our prior research that found that responsive interactions during BR and sociodramatic play activities were associated with growth in vocabulary knowledge (Hadley & Dickinson, 2019). Similarly, other early childhood interventions have identified teachers' ability to engage children's participation and extend their conversational offerings as a key factor in the success of their intervention (Hamre et al., 2010; Neuman et al., 2021).

Interestingly, results from our first research question revealed that these *child-focused* strategies, although potent, were among the most difficult for teachers to consistently implement. Taken together, these two analyses indicate that in scaling up our intervention,

providing better support for teachers' use of *child-focused* strategies is an important leverage point. Previous interventions have successfully supported teachers' ability to engage in responsive conversations with children (Cabell et al., 2015), but because we saw our *child-focused* strategies as fairly straightforward, we likely did not provide adequate coaching support for using these strategies and managing child responses in large groups. Future studies also may investigate the effect of reducing expectations for using *teacher-focused* strategies to diminish demands teachers may have felt to "get through" the guidance cards.

Limitations

As a pilot study with a small number of teacher participants, results should be generalized with caution. In addition, given the small sample of teachers, we did not measure other aspects of IF found to relate to child outcomes (e.g., dosage and quality of delivery; Hamre et al., 2010; Mendive et al., 2016), which may have allowed for a more comprehensive look at IF. Finally, our analyses are correlational. Although threats to internal validity were controlled for by accounting for the nested structure of the data and baseline vocabulary knowledge, the relations identified in this study were not causal. To determine causality, experimental studies that investigate the impact of teacher strategies on child outcomes are needed.

CONCLUSIONS AND FUTURE DIRECTIONS

Our findings suggest the importance of child engagement and of strategies that ask children to actively contribute to the

construction of word meanings. However, our work also highlights the need for additional coaching and support, because the improvisational nature of these interactions resists scripting and relies on teachers' internalization of intervention design principles (Neugebauer et al., 2021). On one hand, giving teachers some latitude to adapt questions and scripted language for their specific context (i.e., "structured adaptations") may be more effective than requiring strict adherence (Bleses et al., 2018; Neuman et al., 2021). However, Quinn and Kim (2017) found that the use of such structured adaptations was successful only after an initial year of implementation that focused on adherence and learning core principles of the intervention.

The present study also has implications for measuring fidelity in early literacy and language studies, indicating the importance of measuring IF across different contexts as well as disaggregating and reporting fidelity scores for specific practices. In future work, the results from fine-grained adherence measures in pilot studies could also be used to develop contextualized, study-specific measures of quality, another dimension of IF (e.g., Darrow, 2013), to employ in scaled up versions of interventions. For example, the quality of teachers' interactions with children while using *child-focused* strategies could be rated on a scale from 1 to 3, allowing for further insights into teachers' skill at managing these interactions. Overall, the present study points to the importance of measuring and analyzing fidelity, not just as a check to ensure that implementers have done what researchers have asked but as a rich information source essential in refining and developing effective interventions.

REFERENCES

-
- Bleses, D., Højen, A., Dale, P. S., Justice, L. M., Dybdal, L., Piasta, S., Markussen-Brown, J., Kjærbaek, L., & Haghish, E. F. (2018). Effective language and literacy instruction: Evaluating the importance of scripting and group size components. *Early Childhood Research Quarterly, 42*, 256–269. <https://doi.org/10.1016/j.ecresq.2017.10.002>
- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning

- approach. *Discourse Processes*, 45(2), 122–159. <https://doi.org/10.1080/01638530701792826>
- Cabell, S. Q., Justice, L. M., McGinty, A. S., DeCoster, J., & Forston, L. D. (2015). Teacher-child conversations in preschool classrooms: Contributions to children's vocabulary development. *Early Childhood Research Quarterly*, 30(Part A), 80–92. <https://doi.org/https://doi.org/10.1016/j.ecresq.2014.09.004>
- Cabell, S. Q., Zucker, T. A., DeCoster, J., Melo, C., Forston, L., & Hamre, B. (2019). Prekindergarten interactive book reading quality and children's language and literacy development: Classroom organization as a moderator. *Early Education and Development*, 30(1), 1–18. <https://doi.org/10.1080/10409289.2018.1514845>
- Capin, P., Walker, M. A., Vaughn, S., & Wanzek, J. (2018). Examining how treatment fidelity is supported, measured, and reported in K-3 reading intervention research. *Educational Psychology Review*, 30(3), 885–919. <https://doi.org/10.1007/s10648-017-9429-z>
- Collins, L. M., Schafer, J. L., & Kam, C.-M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6(4), 330–351. <https://doi.org/10.1037/1082-989X.6.4.330>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1), 23–45. [https://doi.org/10.1016/s0272-7358\(97\)00043-3](https://doi.org/10.1016/s0272-7358(97)00043-3)
- Darrow, C. L. (2013). The Effectiveness and precision of intervention fidelity measures in preschool intervention research. *Early Education and Development*, 24(8), 1137–1160. <https://doi.org/10.1080/10409289.2013.765786>
- 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>
- Dickinson, D. K., Collins, M. F., Nesbitt, K. T., Toub, T. S., Hassinger-Das, B., Hadley, E. B., Hirsh-Pasek, K., & Golinkoff, R. M. (2019a). Effects of teacher-delivered book-reading and play on vocabulary learning and self-regulation among low-income preschool children. *Journal of Cognition and Development*, 20(2), 136–164. <https://doi.org/10.1080/15248372.2018.1483373>
- Dickinson, D. K., Freiberg, J. B., & Barnes, E. M. (2011). Why are so few interventions really effective? A call for fine-grained research methodology. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of early literacy research* (Vol. III, pp. 337–357). Guilford Press.
- Dickinson, D. K., Nesbitt, K. T., Collins, M. F., Hadley, E. B., Newman, K. M., Rivera, B., Ilgaz, H., Nicolopoulou, A., Golinkoff, R. M., & Hirsh-Pasek, K. (2019b). Teaching for breadth and depth of vocabulary knowledge: Learning from explicit and implicit instruction and the storybook texts. *Early Childhood Research Quarterly*, 47, 341–356. <https://doi.org/10.1016/j.ecresq.2018.07.012>
- Dore, R. A., Shirilla, M., Hopkins, E., Collins, M., Scott, M., Schatz, J., Lawson-Adams, J., Valladares, T., Foster, L., Puttre, H., Toub, T. S., Hadley, E., Golinkoff, R. M., Dickinson, D., & Hirsh-Pasek, K. (2019). Education in the app store: Using a mobile game to support U.S. preschoolers' vocabulary learning. *Journal of Children and Media*, 13(4), 452–471. <https://doi.org/10.1080/17482798.2019.1650788>
- Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). American Guidance Service.
- Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2011). Predicting robust vocabulary growth from measures of incremental learning. *Scientific Studies of Reading*, 15(1), 71–91. <https://doi.org/10.1080/10888438.2011.539076>
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, 71(2), 149–164. <https://doi.org/10.1177/001440290507100202>
- Graham, J. W. (2012). *Missing data: Analysis and design*. Springer Science & Business Media.
- Hadley, E. B., Barnes, E. M., Wiernik, B. M., & Raghavan, M. (2022). A meta-analysis of teacher language practices in early childhood classrooms. *Early Childhood Research Quarterly*, 59, 186–202. <https://doi.org/10.1016/j.ecresq.2021.12.002>
- Hadley, E. B., Dedrick, R. F., Dickinson, D. K., Kim, E. S., Hirsh-Pasek, K., & Golinkoff, R. M. (2021). Exploring the relations between child and word characteristics and preschoolers' word-learning. *Journal of Applied Developmental Psychology*, 77, 101332. <https://doi.org/10.1016/j.appdev.2021.101332>
- Hadley, E. B., & Dickinson, D. K. (2019). Cues for word-learning during shared book-reading and guided play in preschool. *Journal of Child Language*, 46(6), 1202–1227. <https://doi.org/10.1017/S0305000919000552>
- Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). 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>
- Hamre, B. K., Justice, L. M., Pianta, R. C., Kilday, C., Sweeney, B., Downer, J. T., & Leach, A. (2010). Implementation fidelity of MyTeachingPartner literacy and language activities: Association

- with preschoolers' language and literacy growth. *Early Childhood Research Quarterly*, 25, 329-347. <https://doi.org/10.1016/j.ecresq.2009.07.002>
- Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children*, 79(2), 181-193. <https://doi.org/10.1177/001440291307900204>
- Hindman, A. H., Wasik, B. A., & Bradley, D. E. (2019). How classroom conversations unfold: Exploring teacher-child exchanges during shared book reading. *Early Education and Development*, 30(4), 478-495. <https://doi.org/10.1080/10409289.2018.1556009>
- Justice, L. M., Mashburn, A., 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\(2008/072\)](https://doi.org/10.1044/1092-4388(2008/072))
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300-335. <https://doi.org/10.3102/0034654310377087>
- 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>
- Mendive, S., Weiland, C., Yoshikawa, H., & Snow, C. (2016). Opening the black box: Intervention fidelity in a randomized trial of a preschool teacher professional development program. *Journal of Educational Psychology*, 108(1), 130. <https://doi.org/10.1037/edu0000047>
- Nelson, M. C., Cordray, D. S., Hulleman, C. S., Darrow, C. L., & Sommer, E. C. (2012). A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The Journal of Behavioral Health Services & Research*, 39(4), 374-396. <https://doi.org/10.1007/s11414-012-9295-x>
- Neugebauer, S. R., Coyne, M., McCoach, D. B., & Ware, S. (2021). Reframing adherence: Active ingredients and impromptu interactions that support vocabulary implementation effectiveness. *Early Childhood Research Quarterly*, 56, 52-64. <https://doi.org/10.1016/j.ecresq.2021.02.004>
- Neuman, S. B., & Danielson, K. (2021). Enacting content-rich curriculum in early childhood: The role of teacher knowledge and pedagogy. *Early Education and Development*, 32(3), 443-458. <https://doi.org/10.1080/10409289.2020.1753463>
- Neuman, S. B., Samudra, P., & Danielson, K. (2021). Effectiveness of scaling up a vocabulary intervention for low-income children, pre-K through first grade. *The Elementary School Journal*, 121(3), 385-409. <https://doi.org/10.1086/712492>
- Noble, C., Sala, G., Peter, M., Lingwood, J., Rowland, C., Gobet, F., & Pine, J. (2019). The impact of shared book reading on children's language skills: A meta-analysis. *Educational Research Review*, 28, 100290. <https://doi.org/10.1016/j.edurev.2019.100290>
- Pence, K. L., Justice, L. M., & Wiggins, A. K. (2008). Preschool teachers' fidelity in implementing a comprehensive language-rich curriculum. *Language, Speech, and Hearing Services in Schools*, 39(3), 329-341. [https://doi.org/doi:10.1044/0161-1461\(2008/031\)](https://doi.org/doi:10.1044/0161-1461(2008/031))
- Quinn, D. M., & Kim, J. S. (2017). Scaffolding fidelity and adaptation in educational program implementation: Experimental evidence from a literacy intervention. *American Educational Research Journal*, 54(6), 1187-1220. <https://doi.org/10.3102/0002831217717692>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- Scott, M. E., Hopkins, E., Masters, A., McMillan, B., Collins, M. F., Dore, R. A., Lawson-Adams, J., Preston, M. L., Toub, T. S., Schatz, J., Weaver, H., Dickinson, D., Golinkoff, R. M., & Hirsh-Pasek, K. (Under review). Beyond book reading: New pathways to vocabulary development through playful learning.
- Snijders, T. A., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Toub, T. S., Hassinger-Das, B., Nesbitt, K. T., Ilgaz, H., Weisberg, D. S., Hirsh-Pasek, K., Golinkoff, R. M., Nicolopoulou, A., & Dickinson, D. K. (2018). The language of play: Developing preschool vocabulary through play following shared book-reading. *Early Childhood Research Quarterly*, 45, 1-17. <https://doi.org/10.1016/j.ecresq.2018.01.010>
- Wasik, B. A., & Hindman, A. H. (2011). Improving vocabulary and pre-literacy skills of at-risk preschoolers through teacher professional development. *Journal of Educational Psychology*, 103(2), 455-469. <https://doi.org/10.1037/a0023067>
- Zucker, T. A., Jacobos, E., & Cabell, S. Q. (2021). Exploring barriers to early childhood teachers' implementation of a supplemental academic language curriculum. *Early Education and Development*, 32(8), 1194-1219. <https://doi.org/10.1080/10409289.2020.1839288>