

# Long-Term Effects of Kindergarten Vocabulary Instruction and Intervention on Target Vocabulary Knowledge Through Second Grade

*Michael D. Coyne, D. Betsy McCoach,  
Susan M. Loftus-Rattan, Doris L. Baker,  
and Sharon M. Ware*

We evaluated the long-term effects of a supplemental, small-group kindergarten vocabulary intervention in fall and spring of first grade and winter of second grade. Participants included students from two studies, an initial efficacy study and a subsequent replication study, identified as at risk for language and learning difficulties, who were randomly assigned in clusters to either a control group that received only classroom vocabulary instruction or a treatment group that received the classroom instruction plus small-group supplemental intervention. We also identified a group of not-at-risk students who received classroom vocabulary instruction as a reference group. Analyses using multilevel modeling indicated that students in the treatment group continued to outperform students in the control group on measures of expressive and receptive knowledge of words taught during the intervention through the winter of second grade, almost 2 years after the end of the intervention. Although long-term effects of the intervention on target word learning in second grade were substantial and meaningful, there was some deterioration of intervention effects across time. There were no effects of the kindergarten intervention on general vocabulary knowledge at posttest or any follow-up points. Findings suggest that the effects of direct and extended vocabulary instruction and intervention are generally durable and sustain over time for taught words.

**Key words:** *follow-up study, kindergarten, long-term effects, multitiered systems of support (MTSS), vocabulary, vocabulary intervention*

---

**Author Affiliations:** *Department of Educational Psychology, Neag School of Education, University of Connecticut, Storrs (Drs Coyne and McCoach); Acadience Learning Inc, Eugene, Oregon (Dr Loftus-Rattan); Department of Special Education, Sanchez School of Education, The University of Texas at Austin (Dr Baker); and Department of Special Education, University of St Joseph, West Hartford, Connecticut (Dr Ware).*

*The research reported here was supported, in whole or in part, by the National Center on Special Education Research, Institute of Education Sciences, U.S. Department of Education, through Grant R324A110135 to the University of Connecticut. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.*

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

**D**IRECT AND EXTENDED vocabulary instruction teaches students the meanings of individual words and provides them with opportunities to engage with those words in meaningful and interactive contexts. The purpose of direct and extended vocabulary instruction is to enable students to develop deep, flexible knowledge of important

---

**Corresponding Author:** *Michael D. Coyne, PhD, Department of Educational Psychology, Neag School of Education, University of Connecticut, 249 Glenbrook Rd, Unit 3064, Storrs, CT 06269 ([mike.coyne@uconn.edu](mailto:mike.coyne@uconn.edu)).*

DOI: 10.1097/TLD.0000000000000296

academic vocabulary and to support comprehension of sentences and passages that include those words. Because this approach to vocabulary instruction is time intensive and focuses on directly teaching relatively few words, an implicit goal is durable word knowledge that sustains over time. The purpose of this study, therefore, was to investigate the long-term effects of kindergarten vocabulary instruction and intervention on students' target word learning and overall vocabulary knowledge through first and second grades.

### **DIRECT AND EXTENDED VOCABULARY INSTRUCTION AND INTERVENTION**

A growing body of research supports the efficacy of direct and extended vocabulary instruction with young students, particularly for learning words targeted for instruction. In a meta-analysis of studies evaluating the effects of vocabulary instruction and intervention in prekindergarten and kindergarten on vocabulary learning, Marulis and Neuman (2010) reported an overall effect size of 0.88 and found that instructional approaches with direct intentional teaching produced larger effects than approaches that relied on incidental teaching. They also found that effects were larger on experimenter-developed measures of target word learning than on standardized measures. Recent intervention studies continue to find positive effects of directly teaching vocabulary to students in preschool and the early grades (e.g., Apthorp et al., 2012; Dickinson et al., 2019; Goldstein et al., 2016; Lonigan & Phillips, 2016; Neuman, et al., 2021; Silverman et al., 2013; Wasik & Hindman, 2020).

In our own work, we have evaluated the effects of direct and extended vocabulary instruction on the learning of students in kindergarten at risk for language and literacy difficulties (e.g., Coyne et al., 2009; Coyne et al., 2010; Coyne et al., 2022). Work by Beck et al. (2013) has informed our approach to vocabulary instruction, including (a) focusing on directly teaching Tier 2 academic

vocabulary that reflects the decontextualized sophisticated language common in books and across content areas (Foorman et al., 2016), (b) giving students clear and understandable definitions of words targeted for instruction, (c) ensuring that students encounter words multiple times in different meaningful contexts, and (d) providing extended opportunities for students to discuss words in interactive activities that promote deep processing.

Our recent research investigated the effects of small-group vocabulary intervention that supplemented whole-class vocabulary instruction across two studies, an initial efficacy study and a subsequent replication study (Coyne et al., 2019; Coyne et al., 2022). In the initial efficacy study, we randomly assigned kindergarten students identified as at risk for experiencing language and vocabulary difficulties to either classroom vocabulary instruction plus supplemental vocabulary intervention (treatment) or classroom vocabulary instruction only (control). We also identified a group of not-at-risk students who received only classroom vocabulary instruction as a comparison group. Kindergarten teachers provided whole-class vocabulary instruction daily for 15–20 min. School-based interventionists provided the supplemental vocabulary intervention for 30 min, 4 days per week, between November and May.

We found that at-risk students who received the treatment condition experienced greater vocabulary learning of words targeted for instruction than students in the control condition. Large effect sizes were evident for experimenter-developed measures of expressive ( $g = 1.05$ ) and receptive vocabulary learning ( $g = 1.14$ ). We also found that at-risk students who received supplemental vocabulary intervention outperformed students in the control group on a measure of listening comprehension measuring students' ability to answer questions about passages read aloud that contained target vocabulary ( $g = 0.47$ ). We did not find any differences on standardized measures of general vocabulary knowledge.

We also found that at-risk students who received supplemental vocabulary intervention experienced comparable, or greater, target word learning than non-at-risk students in the comparison group who received only classroom vocabulary instruction. Not-at-risk students in the comparison group outperformed at-risk students who received supplemental intervention on our listening comprehension measure, but there was a smaller difference between the comparison group students and treatment group students than between the comparison group students and at-risk control students who received only classroom vocabulary instruction.

We then conducted a closely aligned replication study with a different cohort of kindergarten students in the same schools the following year using the same methods and design. Results were replicated again, indicating that direct and extended vocabulary instruction is effective in accelerating students' learning of words targeted for instruction and their ability to comprehend passages that include those target words. Moreover, small-group vocabulary intervention that supplements classroom vocabulary instruction appears to decrease learning differences between at-risk students and their not-at-risk peers.

#### **LONG-TERM EFFECTS OF DIRECT AND EXTENDED VOCABULARY INSTRUCTION AND INTERVENTION**

The goal of direct and extended vocabulary instruction is to enable students to develop deep and flexible knowledge of academic vocabulary by directly teaching word meanings and providing scaffolded opportunities for students to process words across a range of different contexts. Deep and flexible knowledge of individual word meanings, in turn, facilitates the development of strong lexical representations (Perfetti, 2007) that support comprehension of sentences and passages that include those words as well as subsequent vocabulary learning (Ouelette, 2006; Spencer et al., 2017; Stahl, 1991).

Despite the positive effects of direct and extended vocabulary instruction, this approach is time intensive. Compared with embedded vocabulary instruction that briefly introduces new words and provides a quick definition, extended instruction dedicates substantial time to additional activities that promote deep processing (Coyne et al., 2007; Coyne et al., 2009). Because of the instructional investment in each word targeted for instruction, an important question is whether knowledge of words gained through direct and extended instruction is durable and sustains over time. Durable word knowledge would continue to support comprehension and additional vocabulary learning. Evidence that direct and extended vocabulary instruction produces word knowledge that sustains over time would provide support to justify the instructional investment in initial instruction.

In multiple studies (Coyne et al., 2007; Coyne et al., 2009), we found that statistically significant effects favoring extended instruction compared with embedded instruction, incidental exposure, and business as usual instruction were still evident at delayed posttests 6–8 months after intervention in the absence of subsequent instruction or planned review. However, overall means as well as effect sizes decreased, indicating that there was some deterioration of word learning for students who received direct and extended instruction. For example, in the study by Coyne et al. (2009), effects sizes at posttest for expressive target word (ETW) and receptive target word (RTW) knowledge were 2.57 and 0.97, whereas those at an 8-week delayed posttest were 1.18 and 0.84, respectively. These findings are consistent with other studies of follow-up effects of vocabulary instruction with young students (e.g., Bowyer-Crane et al., 2008; Neuman et al., 2011; Vadasy et al., 2013). For example, Vadasy et al. (2015) investigated follow-up effects of a kindergarten vocabulary intervention for English learners in winter of first grade and found a sustained impact on an experimenter-developed measure of

target word learning, although effect sizes were smaller than at posttest. Effect sizes for a reading vocabulary measure of target word knowledge were 0.64 at end-of-kindergarten posttest and 0.29 at follow-up in the winter of first grade.

Among the studies of maintenance effects of vocabulary instruction with young students, there are few studies that examine impacts a year or more after the end of intervention. The purpose of this study was to extend research on the effects of kindergarten vocabulary intervention by investigating the long-term effects of kindergarten vocabulary instruction and intervention on students' target word learning and overall vocabulary knowledge through first and second grades, in the absence of planned review or continued intervention.

## METHOD

To investigate follow-up effects of direct and extended vocabulary instruction, we used data from our initial efficacy study and our replication study. The efficacy study was conducted with students from kindergarten classrooms in 48 elementary schools located in a mix of urban, suburban, and rural districts in the northeastern and northwestern United States. The replication study took place the following year in the same classrooms and schools but with a new cohort of kindergarten students. Methods and procedures were the same across both the efficacy and replication study (see the study by Coyne et al., 2022).

### Participants

The Peabody Picture Vocabulary Test (PPVT) was administered to all kindergarten students early in the school year. Students with a standard score of 92 (30th percentile) or below were considered at risk for language and learning difficulties and eligible to participate. Each participating classroom included approximately six to eight at-risk students. We created clusters of three to four students in each classroom matched on initial

**Table 1.** Race/ethnicity and gender disaggregated by treatment group

	Treatment	Control	Reference
Black	165	143	135
Latinx	314	289	201
White	139	144	230
Asian	31	27	18
Native	3	4	4
Other	6	6	3
Multi	59	48	60
<i>n</i>	717	661	651
Female	338	302	346

PPVT scores (i.e., each cluster represented a similar mean and range of PPVT scores) and randomly assigned clusters to either the treatment group or the control group. We also identified a not-at-risk reference group of three to four students from each classroom with initial PPVT standard scores of 95–105 (37th to 67th percentiles). Student demographic information for each of these groups across studies is provided in Table 1, and the number of participating schools, classes, and intervention groups, disaggregated by study, is provided in Table 2. There were no statistically significant differences between treatment and control groups on any of the demographic variables or pretest assessments.

Before implementing the classroom vocabulary instruction, all teachers received a day of professional development focused on implementing effective instruction with the classroom vocabulary materials. Kindergarten classroom teachers provided 20 min of whole-class vocabulary instruction per day to

**Table 2.** Number of participating schools, classes, and intervention groups disaggregated by study

	Efficacy <i>N</i> Groups	Replication <i>N</i> Groups
School	47	36
Class	159	119
Cluster	127	101

all their students. School-based interventionists delivered the supplemental small-group vocabulary intervention to at-risk students in the treatment group. Interventionists included paraprofessionals, certified teachers, reading teachers, and other professionals. Interventionists received a day of professional development focused on implementing the supplemental vocabulary intervention with fidelity and were provided with additional coaching during the study.

### **Classroom vocabulary instruction**

To standardize whole-class vocabulary instruction, teachers implemented the Elements of Reading-Vocabulary program (EOR-V; Beck & McKeown, 2004), a commercially available curriculum with evidence of efficacy (e.g., Apthorp et al., 2012). The EOR-V program includes lessons and materials for 24 weeks of daily instruction. The EOR-V focuses on teaching high-utility academic vocabulary that occurs across content areas (e.g., *transform*, *observe*, *memorable*, *alert*, *option*). During classroom vocabulary lessons, teachers introduced all students in the classroom to challenging target vocabulary words that were drawn from a story that was read aloud by the teacher. Teachers introduced student-friendly definitions and read anchor sentences that provided meaningful contexts for new vocabulary. Students had multiple opportunities to use the academic vocabulary and definitions in interactive oral, listening, and workbook-based activities.

### **Supplemental vocabulary intervention**

Students in the treatment group received supplemental, Tier 2 vocabulary intervention. School-based interventionists implemented a researcher-developed Early Vocabulary Intervention (EVI) to small groups of three to four students for 30 min per day, 4 days per week for approximately 22 weeks. Early Vocabulary Intervention was developed to align with the classroom vocabulary lessons and to incorporate features of effective vocabulary instruction (Beck et al., 2013; Coyne et al., 2009). Interventionists provided students with di-

rect and extended vocabulary instruction on three to four words per week introduced during classroom instruction. Interventionists provided student-friendly definitions of academic vocabulary, modeled using new words in meaningful sentences, and scaffolded opportunities for students to use new words in their own sentences. Interventionists provided immediate and specific feedback and error correction to support student learning. Target vocabulary also was reviewed systematically across lessons through specially written stories and interactive activities. Words introduced during the first 8 weeks were reviewed three times over the course of the intervention. Words introduced in the second 8 weeks were reviewed twice, and words introduced in the final 8 weeks were reviewed once.

Early Vocabulary Intervention provides students with interactive activities designed to give students multiple opportunities to use new vocabulary in varied and supportive contexts as well as encourage extended language use. Early Vocabulary Intervention taught students to (a) discriminate between examples and nonexamples of pictures representing target words, (b) use target words to describe pictures and to discuss personal experiences, (c) discuss connections between target vocabulary and other words and concepts, and (d) participate in meaningful conversations with peers about the target words.

### **Training**

Research team members provided teachers and interventionists with a full day of professional development at the beginning of the school year. The professional development began with information about vocabulary development and effective practices for teaching vocabulary. Classroom teachers then focused on implementing the EOR-V curriculum and interventionists focused on implementing the EVI materials. All teachers received materials and implementation guides and received guided practice on delivering vocabulary instruction or intervention. Research team members demonstrated and

modeled lessons and activities and teachers and interventionists were provided opportunities to practice with feedback. Early Vocabulary Intervention is a scripted intervention that provided clear and consistent instructional language and modeling that the interventionists practiced to develop a fluent delivery. Interventionists also learned how to provide effective scaffolding of student language and how to provide effective feedback.

Classroom teachers and interventionists were observed at least three times over the course of the school year and received feedback and coaching about their teaching after these observations. In addition, the research team provided frequent check-ins, support, and consultation to interventionists as needed.

### **Fidelity of implementation**

We observed kindergarten classroom teachers three times across the year using a fidelity checklist that captured the key elements of each component of the EOR-V classroom vocabulary program. Teachers' mean fidelity was above 80% across both studies. In addition to ongoing coaching and support, members of the research team observed each interventionist providing supplemental small-group intervention three times during the study using an observation form that documented adherence to each component of the EVI intervention. For example, in each EVI lesson, adherence items included the following for each of the three activities: (a) interventionist uses appropriate materials, (b) interventionist explains/models the activity, (c) interventionist provides students with opportunities to practice (group and/or individual turns), (d) interventionist provides feedback to students, and (e) interventionist completes activity for all words. Each item was coded as observed or not observed. The fidelity observation form also included items to assess the quality of the instructional delivery (e.g., modeling, feedback, opportunities to respond). Interventionists' mean adherence to the EVI intervention was above 85% across both stud-

ies. The mean score for quality of instruction was above 0.90 on a scale of 0–1.0.

Fidelity observations were conducted by research team members who had in-depth knowledge and experience with both the classroom program and the EVI intervention. Observers received training and were required to demonstrate 90% reliability for administering and scoring the fidelity measures and interobserver agreement in the field was above 90%.

### **Measures**

Screening, pretest, and posttest assessments were administered to students individually by trained research staff who were required to demonstrate 90% reliability for administration and scoring of each measure. Screening and pretest measures were administered in the fall before the start of the intervention and posttest measures were administered in the spring within 2–3 weeks of the end of the intervention. Follow-up measures were administered in the fall and the spring of first grade and in the winter (February/March) of second grade.

#### ***Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)***

The PPVT-4 is a norm-referenced, individually administered measure of receptive vocabulary (Dunn & Dunn, 2007). Students were presented with a set of four color pictures on each page presented on an easel. Students were asked to point to the picture that best represented the word spoken by the examiner. Split-half reliabilities range from .89 to .97; test–retest correlations range from .92 to .96.

#### ***Expressive measure of target words (ETW)***

This researcher-developed measure is an individually administered assessment of students' knowledge of target word definitions. We selected 26 of 66 target words taught in the intervention to assess. We purposefully sampled words to include different word

classes (i.e., verbs, adjectives, nouns) and to represent vocabulary introduced across the intervention (i.e., nine words from the first third of the intervention lessons, nine from the second third, and eight from the last third). The examiner asked the student, “What does \_\_\_\_\_ mean?” Two points were awarded for complete and accurate responses, one point for partial and related responses, and zero points for an unrelated response or no response. Cronbach’s  $\alpha$  value for the ETW was 0.88.

**Receptive measure of target words (RTW)**

This researcher-developed measure is an individually administered assessment of students’ receptive knowledge of target vocabulary. We selected 16 of the 26 words from the ETW to assess in this format based on a pilot examining item-level statistics, factor analyses, and reliability analyses. The examiner presented the student with 16 sets of

four color pictures, drawn from pictures representing target vocabulary used in both the classroom and intervention curriculum materials. Each set of pictures contained one that corresponded to the target vocabulary word and three semantically or graphically similar distractors. The examiner said the target vocabulary word and students were asked to point to the correct picture. Cronbach’s  $\alpha$  value for the RTW was 0.83.

**Analysis**

To determine the long-term effects of supplemental kindergarten vocabulary intervention, we combined data from both our efficacy and replications studies. Table 3 includes descriptive statistics for PPVT, ETW, and RTW measures across all five time points, disaggregated by treatment group.

Theoretically, we expected students in all three groups to show improvement between fall and spring kindergarten, given that the intervention occurred during this time

**Table 3.** Descriptive statistics for PPVT, expressive target word, and receptive target word measures across all five time points disaggregated by treatment group

	TRT			Control			Reference		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
<b>PPVT</b>									
Fall K	84.19	5.09	825	84.36	5.18	778	100.91	2.55	740
Spring K	91.55	9.00	740	91.31	9.47	699	103.17	8.43	677
Fall 1	91.60	8.57	426	92.06	9.25	378	104.05	8.72	368
Spring 1	92.26	8.95	283	91.72	9.10	273	102.32	8.30	252
Winter 2	92.76	10.88	225	92.71	10.33	227	102.38	9.05	207
<b>ETW</b>									
Fall K	0.83	1.36	795	0.84	1.54	751	2.26	2.41	703
Spring K	18.69	11.82	738	8.27	7.06	692	13.78	8.42	669
Fall 1	13.12	8.41	400	7.44	6.03	351	12.34	6.91	334
Spring 1	13.17	7.80	398	8.62	6.12	367	13.99	7.23	344
Winter 2	14.23	7.54	224	10.35	6.43	220	17.00	7.25	202
<b>RTW</b>									
Fall K	4.76	1.83	453	4.71	2.02	412	5.73	2.11	409
Spring K	13.57	2.89	732	9.74	3.52	688	12.04	2.98	666
Fall 1	12.62	2.94	396	9.79	3.26	347	11.74	2.76	341
Spring 1	12.05	2.88	281	9.79	3.19	268	11.93	2.72	249
Winter 2	12.67	2.51	227	10.04	3.01	222	12.70	2.33	204

Note. ETW = expressive target word; PPVT = Peabody Picture Vocabulary Test; RTW = receptive target word; TRT = treatment.

frame. We expected students' scores to remain constant or decline between spring of kindergarten and fall of first grade, consistent with the literature on summer learning loss. From fall of first grade through second grade, we were unsure how scores would change. The further students were from the intervention, the more learning loss was likely to occur. However, as students mature and have more exposure to vocabulary in both academic and nonacademic settings, their scores are likely to improve. Before fitting any models, we graphically and descriptively examined the expressive and receptive vocabulary scores across the five time points. We did not expect the growth to follow a linear trend, nor did we expect it to follow any polynomial or other obvious nonlinear trend. Therefore, to allow for the greatest flexibility of the growth trajectory and to capture the potentially jagged, nonlinear nature of the change, we included indicator variables (coded 0/1) for the second through fifth time points (spring kindergarten, fall first grade, spring first grade, and winter/spring second grade).

We fit a four-level multilevel model in which the observations across time were nested within students, nested within clusters, and nested within schools. Treatment group was included as a cluster-level variable (at Level 3). Model 1 is a baseline mode that includes only the indicator variables for time. The intercept represents the expected vocabulary score in fall of kindergarten (prior to the start of the intervention). In Model 1, the spring kindergarten coefficient ( $\gamma_{1,000}$ ) represents the change in the outcome score between the fall of kindergarten and the spring of kindergarten. The fall first-grade coefficient ( $\gamma_{2,000}$ ) represents the change in the outcome score between the spring of kindergarten and the fall of first grade. The spring first-grade coefficient ( $\gamma_{3000}$ ) represents the change in the outcome score between the fall of first grade and the spring of first grade. Finally, the second-grade coefficient ( $\gamma_{4000}$ ) represents the change in

the outcome score between the spring of first grade and the winter/spring of second grade.

Model 2 includes the indicator variables from Model 1 and adds a main effect for treatment (at Level 3) and interaction between treatment and the four time variables. The main effect for treatment tests whether the treatment and control groups are equivalent in the fall of kindergarten, prior to the start of the intervention. The effect of treatment on the intercept was to demonstrate that there were no differences between the two groups prior to intervention, which establishes the equivalence of the two groups.

The treatment by time interactions capture whether the changes between time points differ as a function of treatment group (i.e., treatment vs. control). Positive coefficients indicate that the treatment group exhibited more positive change than the control group between the two consecutive time points.

Model 3 includes the treatment and treatment by time interaction variables from Model 2 and also includes fall kindergarten PPVT score at the student level (centered around the grand mean for the treatment and control groups) and PPVT by time as well as PPVT by treatment by time interactions.

Model 4 is a trimmed model that eliminates PPVT by time and PPVT by treatment by time interactions that are not statistically significant in Model 3. It also eliminates the effect of treatment on the intercept (fall kindergarten score), given that this effect is not statistically significant, nor should it be, given that the treatment and control groups were matched on PPVT scores and then randomized into treatment and control groups. Eliminating these effects does not worsen the fit of the model, and the Akaike information criterion and the Bayesian information criterion are highest for Model 4.

Tables 4 and 5 contain the results for the multilevel models, and Figures 1 and 2 graph the model-predicted vocabulary scores across all five time points for the treatment,



**Table 4.** Multilevel results for the expressive target word measure

	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Level 1				
Spring K				
Coefficient	12.97**	7.66**	7.56**	7.56**
Standard error	0.54	0.56	0.52	0.51
Fall 1				
Coefficient	- 4.14**	- 1.45**	- 1.44**	- 1.44**
Standard error	0.19	0.26	0.26	0.26
Spring 1				
Coefficient	0.45	0.99**	0.99**	0.99**
Standard error	0.19	0.27	0.26	0.26
Grade 2				
Coefficient	1.82**	2.30**	2.35**	2.32**
Standard error	0.23	0.31	0.31	0.31
Spring K by TRT				
Coefficient		10.28**	10.38**	10.38**
Standard error		0.43	0.41	0.39
Fall 1 by TRT				
Coefficient		- 5.11**	- 5.14**	- 5.14**
Standard error		0.35	0.35	0.35
Spring 1 by TRT				
Coefficient		- 1.07*	- 1.07*	- 1.07*
Standard error		0.37	0.36	0.36
Grade 2 by TRT				
Coefficient		- 0.99	- 1.00	- 0.98
Standard error		0.44	0.43	0.43
PPVT × Spring K by TRT				
Coefficient			0.31**	0.30**
Standard error			0.07	0.07
PPVT × Fall 1 by TRT				
Coefficient			- 0.17*	- 0.18**
Standard error			0.07	0.04
PPVT × Spring 1 by TRT				
Coefficient			- 0.00	
Standard error			0.07	
PPVT × Grade 2 by TRT				
Coefficient			- 0.13	
Standard error			0.08	
PPVT × Spring K				
Coefficient			0.26**	0.26**
Standard error			0.06	0.05
PPVT × Fall 1				
Coefficient			0.01	
Standard error			0.05	
PPVT × Spring 1				
Coefficient			- 0.02	
Standard error			0.05	

(continues)

**Table 4.** Multilevel results for the expressive target word measure (*Continued*)

	Model 1	Model 2	Model 3	Model 4
PPVT × Grade 2				
Coefficient			0.09	
Standard error			0.06	
Intercept				
Coefficient	0.84**	0.84**	0.84**	0.84**
Standard error	0.09	0.13	0.13	0.09
Level 2				
PPVT				
Coefficient			0.05*	0.05*
Standard error			0.02	0.02
Level 3				
Treatment				
Coefficient		− 0.01	− 0.00	
Standard error		0.18	0.18	
Random effects				
Level 1				
Sigma-square				
Coefficient	13.67	12.10	12.00	12.01
Standard error	0.33	0.29	0.29	0.29
Level 2				
var(p1)				
Coefficient	55.92	40.03	36.12	36.10
Standard error	2.99	2.26	2.09	2.09
Tau00				
Coefficient	0.12	0.15	0.12	0.12
Standard error	0.07	0.07	0.06	0.06
cov(p1, cons)				
Coefficient	2.56**	2.44**	2.07**	2.08**
Standard error	0.69	0.55	0.52	0.52
Level 3				
var(p1)				
Coefficient	5.27	7.03	7.25	7.25
Standard error	1.79	1.72	1.65	1.65
Level 4				
var(p1)				
Coefficient	8.62	7.99	6.02	6.02
Standard error	2.87	2.68	2.13	2.13
Model fit				
N	4928	4928	4928	4928
Log likelihood	− 15173.23	− 14780.92	− 14700.03	− 14702
No. parameters	11	16	25	19
AIC	30368.46	29593.84	29450.06	29441.99
BIC	30439.99	29697.88	29612.63	29565.54

*Note.* AIC = Akaike information criterion; BIC = Bayesian information criterion; cov(p1,cons) = covariance between the piece 1 slopes and the intercepts; PPVT = Peabody Picture Vocabulary Test; Tau00 = variance in the randomly varying intercepts; TRT = treatment; var(p1) = variance in the piece 1 slope.

\* $p < .05$ , \*\* $p < .01$ .

**Table 5.** Multilevel results for receptive target word measure

	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Level 1				
Spring K				
Coefficient	6.98**	5.04**	5.03**	5.01**
Standard error	0.20	0.22	0.21	0.20
Fall 1				
Coefficient	-0.73**	-0.19	-0.18	-0.18
Standard error	0.08	0.12	0.12	0.12
Spring 1				
Coefficient	0.05	0.38*	0.38*	0.38*
Standard error	0.10	0.14	0.14	0.14
Grade 2				
Coefficient	0.43**	0.38	0.37	0.38
Standard error	0.11	0.15	0.15	0.15
Spring K by TRT				
Coefficient		3.76**	3.80**	3.84**
Standard error		0.19	0.18	0.14
Fall 1 by TRT				
Coefficient		-1.03**	-1.03**	-1.04**
Standard error		0.16	0.16	0.16
Spring 1 by TRT				
Coefficient		-0.64**	-0.65**	-0.64**
Standard error		0.19	0.19	0.19
Grade 2 by TRT				
Coefficient		0.09	0.08	0.08
Standard error		0.21	0.21	0.21
PPVT × Spring K by TRT				
Coefficient			-0.04	
Standard error			0.03	
PPVT × Fall 1 by TRT				
Coefficient			0.05	
Standard error			0.03	
PPVT × Spring 1 by TRT				
Coefficient			-0.02	
Standard error			0.04	
PPVT × Grade 2 by TRT				
Coefficient			0.03	
Standard error			0.04	
PPVT × Spring K				
Coefficient			0.14**	0.12**
Standard error			0.02	0.02
PPVT × Fall 1				
Coefficient			-0.02	
Standard error			0.02	
PPVT × Spring 1				
Coefficient			0.01	
Standard error			0.03	

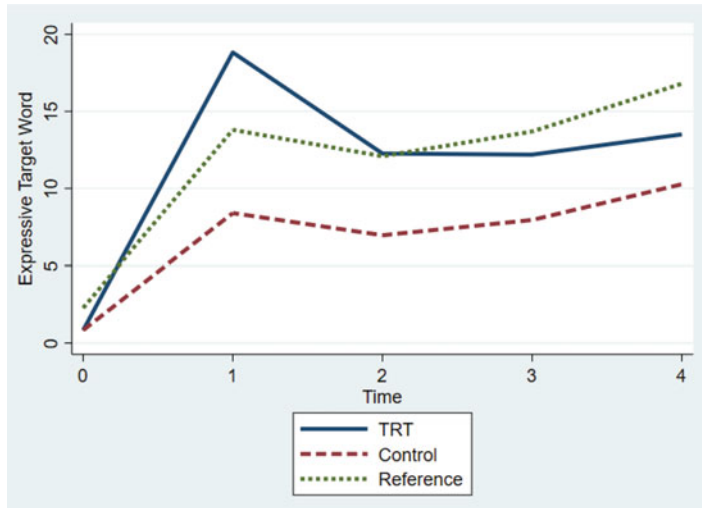
(continues)

**Table 5.** Multilevel results for receptive target word measure (*Continued*)

	Model 1	Model 2	Model 3	Model 4
PPVT × Grade 2				
Coefficient			− 0.01	
Standard error			0.03	
Intercept				
Coefficient	4.74**	4.72**	4.67**	4.70**
Standard error	0.07	0.09	0.09	0.07
Level 2				
PPVT				
Coefficient			0.05**	0.05**
Standard error			0.01	0.01
Level 3				
Treatment				
Coefficient		0.04	0.05	
Standard error		0.13	0.13	
Random effects				
Level 1				
Sigma-square				
Coefficient	2.71	2.55	2.55	2.56
Standard error	0.09	0.09	0.09	0.09
Level 2				
var(p1)				
Coefficient	7.44	4.77	4.41	4.38
Standard error	0.57	0.46	0.44	0.44
Tau00				
Coefficient	0.97	1.13	1.06	1.05
Standard error	0.20	0.20	0.19	0.19
cov(p1, cons)				
Coefficient	− 0.00	− 0.19	− 0.35	− 0.34
Standard error	0.27	0.24	0.24	0.24
Level 3				
var(p1)				
Coefficient	0.75	1.11	1.12	1.12
Standard error	0.25	0.25	0.24	0.24
Level 4				
var(p1)				
Coefficient	1.15	1.03	0.80	0.80
Standard error	0.41	0.38	0.31	0.31
Model fit				
N	4023	4023	4023	4023
Log likelihood	− 9294.921	− 9017.369	− 8937.072	− 8939.443
No. parameters	11	16	25	17
AIC	18611.84	18066.74	17924.14	17912.89
BIC	18681.14	18167.53	18081.64	18019.98

*Note.* AIC = Akaike information criterion; BIC = Bayesian information criterion; cov(p1,cons) = covariance between the piece 1 slopes and the intercepts; PPVT = Peabody Picture Vocabulary Test; Tau00 = variance in the randomly varying intercepts; TRT = treatment; var(p1) = variance in the piece 1 slope.

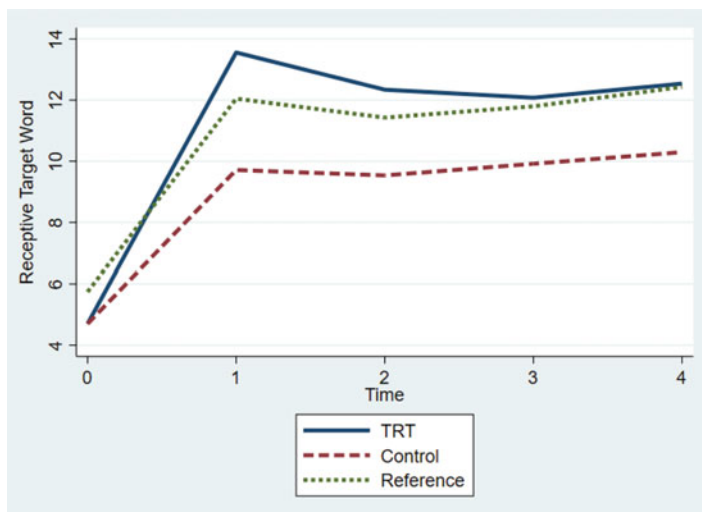
\*p < .05, \*\*p < .01.



**Figure 1.** Performance on the expressive target word measure from fall of kindergarten (time = 0) to second grade (time = 4). TRT = treatment.

control, and reference groups. The reference group consists of not-at-risk students with average PPVT scores who received classroom vocabulary instruction. Although the reference group was not randomized, it serves as another point of comparison to demonstrate how treatment students' scores compare with the scores of average students who received Tier 1 classroom instruction. Because the reference group was not

randomized, it is not included in the multilevel models depicted in Tables 4 and 5. However, the reference group is included in Figures 1 and 2, and we conducted supplemental multilevel analyses to determine whether the differences between the treatment group and the reference group and the treatment group and the control group were statistically significant at each of the five time points.



**Figure 2.** Performance on the receptive target word measure from fall of kindergarten (time = 0) to second grade (time = 4). TRT = treatment.

## RESULTS

Table 3 includes descriptive statistics across all five time points, and Figures 1 and 2 display student performance on the ETW and RTW measures, respectively, over time for students in the treatment, control, and reference groups. Table 6 reports Hedge's *g* effect sizes for all comparisons. We interpret the results of Model 4 for both the ETW and RTW measures. However, the parameter estimates for Model 3 (the untrimmed model) are virtually identical.

Consistent with results reported in our efficacy and replications studies (Coyne et al., 2022), students in the treatment group outperformed students in the control group at the end-of-kindergarten posttest on the ETW and the RTW with large effect size differences. Also, consistent with our previous research, we found that PPVT performance at pretest moderated treatment effects on the ETW. In other words, students with higher PPVT scores at pretest responded more strongly to the intervention as measured by the ETW (for more detailed discussion of moderation effects in the efficacy and replications studies, see the study by Coyne et al., 2019). Students in the treatment group also outperformed the reference students on the

ETW and the RTW at posttest. There were no treatment effects on the PPVT.

### Expressive measure of target words

The results of the multilevel analysis for ETW appear in Table 4 and Figure 1. In the completely unconditional model, 6% of the variance was between schools, 4% of the variance was between clusters within schools, 14% of the variance was between students within clusters, and 76% of the variance was within students. For ETW, students in both the treatment and control groups scored on average 0.84 points. In spring of kindergarten, holding PPVT constant at the sample mean, students in the control group scored 7.56 points higher than they did in the fall. In the control group, students with higher PPVT scores did gain slightly more than students with lower PPVT scores. For every 1-point increase in PPVT, the control students' expressive scores increased by 0.26 points. In the treatment group, students of average PPVT outscored students in the control group by 10.38 points. In other words, they scored over 18 points higher in spring of kindergarten than they did in the fall. Again, treatment students with higher PPVT scores made even greater gains than treatment students with lower PPVT scores, and

**Table 6.** Comparisons of treatment, control, and reference students at follow-up points

	Hedge's <i>g</i> Effect Size		
	Treatment/ Control	Treatment/ Reference	Reference/ Control
Expressive target word			
K Spring	1.06	0.47	0.71
First fall	0.77	0.10	0.76
First spring	0.64	− 0.11	0.80
Second winter	0.55	− 0.37	0.97
Receptive target word			
K spring	1.19	0.52	0.70
First fall	0.92	0.31	0.64
First spring	0.74	0.04	0.72
Second winter	0.95	− 0.01	0.98

*Note.* In the treatment/control and treatment/reference comparisons, a positive effect size favors the treatment group; in the reference/control comparisons, a positive effect size favors the reference group.

they made greater gains than comparable control students ( $b = 0.30$ ). Whereas control students scored 0.26 points higher on EWT for every point increase in PPVT; treatment students scored 0.56 ( $0.26 + 0.30$ ) points higher on EWT for every point increase in PPVT. In fall of first grade, control students lost 1.44 points on average, and treatment students lost 5.14 points more than control students did, making the total expected loss of 6.58 points for a treatment student ( $-1.44 + -5.14$ ). The PPVT performance did not appear to moderate ETW loss for control students. However, treatment students with higher PPVT scores lost slightly more ground ( $b = -0.18$ ) than those with lower PPVT scores (but recall that they gained more between the fall and spring of kindergarten). In spring of first grade, control students gained about 1 point ( $b = 0.99$ ) over their fall first-grade scores; treatment students' spring first-grade scores remained similar to their fall first-grade scores: they gained 1.07 points less than control students during this period, which means that their change was close to 0. Between first and second grades, control students gained 2.32 points; treatment students gained about a point less than control students did ( $b = -0.98$ ). Although the gap between the treatment and control students did narrow during this time period, the treatment students still outscored the control students. The predicted score for an average treatment student in second grade was 13.44 (95% confidence interval [CI] = 12.36–14.53). In contrast, the predicted score for a control student in second grade was 10.27 (95% CI = 9.18–11.36).

In summary, treatment students' follow-up trajectory on the ETW was characterized by a decline in scores between spring of kindergarten and fall of first grade and then relatively stable performance through winter/spring of second grade. Control students also experienced a decline in ETW scores between spring of kindergarten and fall of first grade and then a gradual increase in scores through winter/spring of second grade. Effect size differences between treatment and

control groups narrowed between posttest ( $g = 1.06$ ) and second grade ( $g = 0.55$ ); however, this effect continued to be substantial and statistically significant. Reference students experienced a decline in ETW scores between spring of kindergarten and fall of first grade and then an increase in scores through winter/spring of second grade. Effect size differences between treatment and reference groups diminished between posttest ( $g = 0.47$ ) and fall of first grade ( $g = 0.10$ ) and then widened through the winter/spring of second grade, favoring the reference students ( $g = -0.37$ ).

### Receptive measure of target words

In the completely unconditional model, 6% of the variance was between schools, 10% of the variance was between clusters within schools, 12% of the variance was between students within clusters, and 72% of the variance was within students. For the RTW measure, students in both groups had expected scores of 4.7 points in the fall of kindergarten. There are no PPVT by treatment interactions for the RTW measure, making interpretations even easier. In spring of kindergarten, students in the control group scored approximately 5 points higher than they did in the fall ( $b = 5.01$ ). Treatment students outscored students in the control group by 3.84 points. Therefore, they scored 8.85 points higher ( $5.01 + 3.84$ ) in spring of kindergarten than they did in the fall. In the fall of first grade, control students' scores were similar to their scores in the spring of kindergarten ( $b = -0.18$ , not statistically significant). Treatment students' scores decreased by 1.04 points more than control students' scores did. In spring of first grade, control students gained about 0.38 points over their fall of first-grade scores; treatment students gained  $-0.64$ , when compared with the control students, which indicates that their scores dropped slightly ( $0.38 - 0.64 = -0.28$ ) over first grade. Between first and second grades, control students again gained 0.38 points; during this time period, treatment students' gain was on par with that of the control students

( $b = 0.08$ , statistically significant). Although the gap between the treatment and control students did narrow during across first and second grades, the treatment students still outscored the control students by second grade. The predicted score for an average treatment student in second grade was 12.54 (95% CI = 12.12–12.96). In contrast, the predicted score for a control student in second grade was 10.30 (95% CI = 9.87–10.72).

In summary, RTW trajectories of treatment, control, and reference students were all characterized by a decline in scores between spring of kindergarten and fall of first grade and then relatively stable performance through winter/spring of second grade. Effect size differences between treatment and control groups on the RTW narrowed between posttest ( $g = 1.19$ ) and second grade ( $g = 0.95$ ); however, this effect continued to be substantial and statistically significant. Effect size differences between treatment and reference groups diminished between posttest ( $g = 0.52$ ) and second grade ( $g = -0.01$ ), but treatment students continued to score similarly to the reference students in second grade.

## DISCUSSION

The purpose of this study was to evaluate the long-term effects of a small-group vocabulary intervention that supplemented classroom vocabulary instruction for kindergarten students identified as at risk for language and learning difficulties. Results from our initial efficacy and replication studies indicated that at-risk treatment students who received intervention outperformed at-risk control students who received only classroom vocabulary instruction on measures of ETW and RTW learning.

### Long-term effects

Findings indicated that at-risk treatment students who received the kindergarten vocabulary intervention continued to significantly outperform at-risk control students on expressive and receptive measures of

taught vocabulary through the winter of second grade, almost 2 years after the end of the intervention. These maintenance effects of the kindergarten vocabulary intervention through the winter of second grade were characterized by moderate to large effects of 0.56 for expressive word learning and 0.66 for receptive word learning. These findings suggest that intervention that includes direct and extended vocabulary instruction in kindergarten produces substantial and meaningful word learning that is durable and that sustains over time, even in the absence of planned review or continued intervention. Our findings are consistent with results from the limited number of follow-up studies of direct vocabulary instruction (e.g., Bowyer-Crane et al., 2008; Neuman et al., 2011; Vadasy et al., 2013) and extend the literature by examining long-term effects almost 2 years after the end of the intervention.

Direct and extended vocabulary instruction involves a significant investment in time, providing multiple opportunities for students to interact with target words in different meaningful contexts across lessons. The implicit assumption is that this investment in instructional time enables students to develop deep and lasting word knowledge that will continue to support comprehension of text that includes target vocabulary over time as well as subsequent vocabulary learning (Ouelette, 2006; Spencer et al., 2017). Findings of this study that at-risk students who received kindergarten intervention continue to demonstrate higher levels of ETW and RTW knowledge almost 2 years after intervention provide evidence for the durability of word learning gained through direct and extended instruction and support the initial investment in instructional time.

### Deterioration of effects

Although robust treatment effects of our kindergarten vocabulary intervention were still evident in second grade, findings also revealed some deterioration of these effects over time, which is consistent with the few studies that examined follow-up effects of



kindergarten vocabulary instruction and intervention (e.g., Coyne et al., 2009; Vadasy et al., 2015). Between the end of kindergarten and the beginning of first grade, both at-risk treatment and control students experienced a decline in their ETW and RTW scores. The decline for treatment students who received intervention and made the most vocabulary gains over the course of kindergarten was more pronounced than the decline for the control students who made significantly less gains in kindergarten, especially for expressive knowledge of target words. Between the beginning of first grade and the winter of second grade, expressive and receptive knowledge of target words decreased slightly for treatment students and increased slightly for control students. Therefore, the overall effect of the intervention was characterized by differences between treatment and control group means that narrowed between the end of the intervention and each additional follow-up point. These findings are evident in Table 6 and Figures 1 and 2.

To help interpret these findings, we calculated the approximate mean percentage of target words that students in the treatment and control groups learned on measures of target word learning administered at posttest and second grade (we divided the mean scores for each group at posttest and at second-grade follow-up by the total score possible). At posttest, at-risk control students who received classroom vocabulary instruction demonstrated expressive knowledge of approximately 32% and receptive knowledge of 60% of target words. In comparison, at-risk treatment students who received classroom vocabulary instruction plus supplemental intervention demonstrated expressive knowledge of approximately 72% and receptive knowledge of 85% of target words. In winter of second grade, control students demonstrated expressive knowledge of approximately 40% and receptive knowledge of 63% of target words and treatment students demonstrated expressive knowledge of approximately 55% and receptive knowledge of 79% of target words.

These findings suggest that word knowledge gained through direct and extended vocabulary instruction may be subject to some deterioration over time. The greater decline in the effects on the expressive measure, which captures depth of vocabulary knowledge, also suggests that higher and more complete levels of word knowledge may be more susceptible to deterioration.

Direct and extended vocabulary intervention involves intensive and explicit instruction that includes multiple exposures to academic vocabulary across different and supportive contexts and systematic review of target words across lessons. Evidence of deterioration of intervention effects over time suggests that sustaining and deepening vocabulary knowledge gained through direct and extended instruction may require ongoing review or at least intentional, planned continued encounters with target academic vocabulary in supportive texts over time.

### **Vocabulary differences among students**

We also were able to collect follow-up data on a group of not-at-risk kindergarten students who received classroom vocabulary instruction only. Results from our efficacy and replication studies indicated that at-risk students who received supplemental vocabulary intervention as well as classroom vocabulary instruction outperformed this not-at-risk reference group on end-of-kindergarten posttest target word measures. These findings suggest that supplemental vocabulary intervention was able to substantially decrease or eliminate vocabulary learning differences between at-risk and not-at-risk students.

Findings from our follow-up analyses revealed that differences between the treatment students and reference students narrowed. By the winter of second grade, reference students demonstrated similar scores on the RTW measure to treatment students and higher scores on the ETW measure. Although these analyses were descriptive and not experimental, these findings provide additional evidence suggesting that effects of

direct and extended vocabulary instruction may deteriorate over time, especially expressive knowledge of taught vocabulary. Despite some deterioration, however, at-risk students' receptive knowledge of vocabulary taught during the kindergarten intervention was still comparable with that of the not-at risk students who received classroom instruction. This finding suggests that the narrowing of vocabulary learning differences between at-risk students and their typically achieving peers that students experienced during the kindergarten intervention was still apparent 2 years later, at least on the receptive vocabulary target word measure.

### SUMMARY AND IMPLICATIONS

Converging evidence supports the efficacy of direct and extended vocabulary instruction and intervention for accelerating student learning of important academic vocabulary. There are few studies, however, that have investigated the long-term effects of these interventions. Results from this study suggest that the effects of a supplemental kindergarten vocabulary intervention characterized by direct and extended instruction were durable and sustained for almost 2 years after the end of the intervention. These findings suggest that the initial benefits of vocabulary intervention on word learning are for the most part lasting and provide support for the instructional investment in a direct and extended approach to vocabulary instruction.

Although the long-term impacts of the intervention on target word learning in second grade were still substantial and meaningful, there was some deterioration of intervention effects across time, especially on our measure of expressive vocabulary that captured depth of word knowledge. Moreover, there were no effects of the kindergarten vocabulary intervention on general vocabulary knowledge measured by the PPVT at posttest or at follow-up.

These results on the PPVT are consistent with other experimental studies of vocabulary instruction and intervention that find

moderate to large effects on proximal and near transfer measures of target word learning and small to no effects on distal standardized measures of vocabulary (Apthorp et al., 2012; Elleman et al., 2009). It may be that it is unrealistic to expect that vocabulary interventions that directly teach a defined set of vocabulary words would lead to changes in overall vocabulary knowledge. Another challenge with evaluating the distal effects of vocabulary intervention is that common standardized measures (e.g., PPVT) may be less sensitive to the effects of targeted interventions (McKeown et al., 2017; Pearson et al., 2007).

Overall, our findings suggest that there is an important role for targeted direct and extended vocabulary intervention that teaches important academic vocabulary. However, providing supplemental vocabulary intervention in addition to classroom vocabulary instruction requires a significant investment in additional time, resources, and personnel. In our study, students in intervention received approximately double the amount of instructional time dedicated to vocabulary instruction, which could be difficult for some schools to replicate and potentially quite costly. Direct and extended vocabulary intervention as devised in most studies including this one also may not be sufficient to meet the language and vocabulary needs of students at risk for language and learning difficulties. First, direct and extended vocabulary instruction and intervention may need to continue across multiple years. This would maximize the number of carefully chosen academic vocabulary that could be taught directly and ensure that previously learned words would continue to be reviewed and practiced over time.

In addition, researchers and practitioners should consider additional approaches to vocabulary instruction that may have more generative and indirect effects on vocabulary learning and comprehension; for example, those that focus on other language skills such as morphological, syntactic, or metalinguistic awareness (Gottardo et al., 2018; Nagy,

2007; Spencer et al., 2017). A comprehensive approach to supporting vocabulary learning and language development could consist of sustained direct and extended vocabulary in-

struction and intervention over multiple years along with aligned and coordinated support for general language development that supports vocabulary acquisition.

## REFERENCES

- Apthorp, H., Randel, B., Cherasaro, T., Clark, T., McKeown, M., & Beck, I. (2012). Effects of a supplemental vocabulary program on word knowledge and passage comprehension. *Journal of Research on Educational Effectiveness*, 5(2), 160-188. <https://doi.org/10.1080/19345747.2012.660240>
- Beck, I. L., & McKeown, M. G. (2004). *Elements of Reading vocabulary teacher's guide level K*. Steck-Vaughn.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). Guilford Press.
- Bowyer-Crane, C., Snowling, M. J., Duff, F. J., Fieldsend, E., Carroll, J. M., Miles, J., Gotz, K., & Hulme, C. (2008). Improving early language and literacy skills: Differential effects of an oral language versus a phonology with reading intervention. *Journal of Child Psychology and Psychiatry*, 49(4), 422-432. <https://doi.org/10.1111/j.1469-7610.2007.01849.x>
- Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction with embedded instruction and incidental exposure. *Learning Disability Quarterly*, 30, 74-88.
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth vs. depth. *The Elementary School Journal* 110, 1-18.
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Ruby, M., Crevecoeur, Y., & Kapp, S. (2010). Direct & extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness*, 3, 93-120. <https://doi.org/10.1080/19345741003592410>
- Coyne, M. D., McCoach, D. B., Ware, S., Austin, C., Loftus, S., & Baker, D. (2019). Racing against the vocabulary gap: Matthew effects in early vocabulary instruction and intervention. *Exceptional Children*, 85, 163-179.
- Coyne, M. D., McCoach, D. B., Ware, S. M., Loftus-Rattan, S. M., Baker, D. L., Santoro, L. E., & Oldham, A. C. (2022). Supporting vocabulary development within a multitiered system of support: Evaluating the efficacy of supplementary kindergarten vocabulary intervention. *Journal of Educational Psychology*, 114(6), 1225-1241. <https://doi.org/10.1037/edu0000724>
- Dickinson, D. K., Nesbitt, K. T., Collins, M. E., Hadley, E. B., Newman, K., Rivera, B. L., Ilgez, H., Nicolopoulou, A., Gofinkoff, R. M., & Hirsh-Pasek, K. (2019). 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>
- Dunn, L. M., & Dunn, D. M., & Pearson Assessments. (2007). *PPVT-4: Peabody picture vocabulary test* (4th ed.). Pearson Assessments.
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, 2, 1-44. <https://doi.org/10.1080/19345740802539200>
- Foorman, B., Beyler, N., Borradaile, K., Coyne, M., Denton, C. A., Dimino, J., Furgeson, J., Hayes, L., Henke, J., Justice, L., Keating, B., Lewis, W., Sattar, S., Streke, A., Wagner, R., & Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade* (NCEE 2016-4008). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. <http://whatworks.ed.gov>
- Goldstein, H., Kelley, E., 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-L15-0227](https://doi.org/10.1044/2015_JSLHR-L15-0227)
- Gottardo, A., Mirza, A., Koh, P. W., Ferreira, A., & Javier, C. (2018). Unpacking listening comprehension: The role of vocabulary, morphological awareness, and syntactic knowledge in reading comprehension. *Reading and Writing: An Interdisciplinary Journal*, 31(8), 1741-1764. <https://doi.org/10.1007/s11145-017-9736-2>
- Lonigan, C. J., & Phillips, B. M. (2016). Response to instruction in preschool: Results of two randomized studies with children at significant risk of reading difficulties. *Journal of Educational Psychology*, 108(1), 114-129. <https://doi.org/10.1037/edu0000054>
- 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, 300-335. <https://doi.org/10.3102/0034654310377087>
- McKeown, M. G., Deane, P. D., & Lawless, R. R. (2017). *Vocabulary assessment to support instruction:*

- Building rich word-learning experiences.* Guilford Press.
- Nagy, W. E. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52–77). Guilford Press.
- Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly, 46*(3), 249–272.
- 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>
- Ouelette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and word comprehension. *Journal of Educational Psychology, 98*, 554–566. <https://doi.org/10.1037/0022-0663.98.3.554>
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Theory and research into practice: Vocabulary assessment: What we know and what we need to know. *Reading Research Quarterly, 42*(2), 282–296. <https://doi.org/10.1598/RRQ.42.2.4>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- 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 & Development, 24*(2), 98–122. <https://doi.org/10.1080/10409289.2011.649679>
- Spencer, M., Quinn, J. M., & Wagner, R. K. (2017). Vocabulary, morphology, and reading comprehension. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 239–256). John Benjamins.
- Stahl, S. A. (1991). Beyond the instrumentalist hypothesis: Some relationships between word meanings and comprehension. In P. Schwanenflugel (Ed.), *The psychology of word meanings* (pp. 157–178). Lawrence Erlbaum Associates.
- Vadasy, P. E., Nelson, J. R., & Sanders, E. A. (2013). Longer term effects of a tier 2 kindergarten vocabulary intervention for English learners. *Remedial and Special Education, 34*(2), 91–101. <https://doi.org/10.1177/0741932511420739>
- Vadasy, P. E., Sanders, E. A., & Nelson, J. R. (2015). Effectiveness of supplemental kindergarten vocabulary instruction for English learners: A randomized study of immediate and longer-term effects of two approaches. *Journal of Research on Educational Effectiveness, 8*(4), 490–529. <https://doi.org/10.1080/19345747.2015.1033120>
- Wasik, B. A., & Hindman, A. H. (2020). Increasing preschoolers' vocabulary development through a streamlined teacher professional development intervention. *Early Childhood Research Quarterly, 50*, 101–113.