

2009

# Direct Vocabulary Instruction in Kindergarten: Teaching for Breadth versus Depth

Michael D. Coyne

D. Betsy McCoach

*See next page for additional authors*

Follow this and additional works at: [https://digitalcommons.uri.edu/psy\\_facpubs](https://digitalcommons.uri.edu/psy_facpubs)

Terms of Use

All rights reserved under copyright.

---

## Citation/Publisher Attribution

Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Jr. & Kapp, S. (2009). Direct Vocabulary Instruction in Kindergarten: Teaching for Breadth, Versus Depth. *Elementary School Journal*, 110(1),1-48.

Available at <http://www.jstor.org/stable/10.1086/598840?origin=JSTOR-pdf>

This Article is brought to you for free and open access by the Psychology at DigitalCommons@URI. It has been accepted for inclusion in Psychology Faculty Publications by an authorized administrator of DigitalCommons@URI. For more information, please contact [digitalcommons@etal.uri.edu](mailto:digitalcommons@etal.uri.edu).

---

**Authors**

Michael D. Coyne, D. Betsy McCoach, Susan Loftus, Richard Zipoli Jr., and Sharon Kapp



---

Direct Vocabulary Instruction in Kindergarten: Teaching for Breadth versus Depth

Author(s): Michael D. Coyne, D. Betsy McCoach, Susan Loftus, Richard Zipoli Jr., and Sharon Kapp

Source: *The Elementary School Journal*, Vol. 110, No. 1 (September 2009), pp. 1-18

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/598840>

Accessed: 04/04/2013 14:18

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).



*The University of Chicago Press* is collaborating with JSTOR to digitize, preserve and extend access to *The Elementary School Journal*.

<http://www.jstor.org>

# Direct Vocabulary Instruction in Kindergarten: Teaching for Breadth versus Depth

Michael D. Coyne  
D. Betsy McCoach  
*University of Connecticut*

Susan Loftus  
*University of Rhode Island*

Richard Zipoli Jr.  
*West Hartford Public Schools, CT*

Sharon Kapp  
*Marlborough Public Schools, MA*

*The Elementary School Journal*  
Volume 110, Number 1  
© 2009 by The University of Chicago. All rights reserved.  
0013-5984/2009/11001-0001\$10.00

## Abstract

The purpose of this study was to compare 2 methods for directly teaching word meanings to kindergarten students within storybook read-alouds that varied in instructional time and depth of instruction along with a control condition that provided students with incidental exposure to target words. Embedded instruction introduces target word meanings during storybook readings in a time-efficient manner. Extended instruction is more time intensive but provides multiple opportunities to interact with target words outside the context of the story. Participants included 42 kindergarten students who were taught 9 target words, 3 with each method. Target words were counterbalanced in a within-subjects design. Findings indicated that extended instruction resulted in more full and refined word knowledge, while embedded instruction resulted in partial knowledge of target vocabulary. Implications are discussed in relation to the strengths and limitations of different approaches to direct vocabulary instruction in kindergarten and the trade-offs between instruction that focuses on teaching for breadth versus depth.

Children enter school with significant differences in vocabulary knowledge (Hart & Risley, 1995) and these differences grow larger in the early grades (Biemiller & Slonim, 2001). Converging evidence has suggested that instruction in code-based skills is insufficient to meet the needs of students who are at risk for experiencing reading problems because of language and vocabulary difficulties (Catts, Hogan, & Adolf, 2005; Storch & Whitehurst, 2002; Vellutino, Tunmer, Jaccard, & Chen, 2007). Therefore, there is growing recognition of the importance of accelerating vocabulary development in young school-age children

through targeted and teacher-supported instruction and other intervention efforts (Biemiller, 2001; Catts et al., 2005).

Researchers and practitioners interested in accelerating academic achievement face a difficult challenge concerning how best to leverage scarce instructional time. In vocabulary intervention research, discussions about leveraging instructional time often revolve around the trade-offs between teaching for breadth or depth. The purpose of this study was to evaluate two approaches for supporting vocabulary learning with kindergarten students in the context of storybook read-alouds. Specifically, we compared two methods for directly teaching word meanings that varied in instructional time and depth of instruction. We also compared these methods with a control condition that provided students with incidental exposure to target words.

### **Teaching for Breadth: Embedded Vocabulary Instruction**

At the beginning of kindergarten, typically developing students know thousands more word meanings than their peers at risk of language and learning difficulties (Hart & Risley, 1995). This vocabulary gap only grows larger in the primary grades. For example, Biemiller and Slonim (2001) estimated that by second grade, children with large vocabularies know approximately 4,000 more root word meanings than children with delays in vocabulary development. Although it is impossible to teach directly all the words necessary to close this gap (Anderson & Nagy, 1992), one goal of direct vocabulary instruction is to introduce students to as many new words as possible.

Instructional approaches that allow the introduction of many word meanings most often provide students with brief definitions of target words within the context of oral language experiences such as story read-alouds (e.g., Elley, 1989; Penno, Wilkinson, & Moore, 2002). This approach characterizes

the research of Biemiller and his colleagues. Biemiller and Boote (2006) conducted two studies in which vocabulary learning among primary students improved when brief explanations of word meanings were provided during repeated storybook readings. Knowledge of word meanings was tested by rating verbal explanations of words presented in context sentences. Children demonstrated a 22% gain in instructed words compared to a 12% gain for noninstructed words, indicating that explicit explanations resulted in a statistically significant increase in word learning. In a second study, two reviews of each word meaning were provided, including an opportunity to review word meanings in new context sentences. Children in this study showed an average gain of 41%. Biemiller and Boote concluded that teaching many word meanings without in-depth discussion appeared to be an effective approach to direct vocabulary instruction.

Embedded vocabulary instruction has a number of benefits. First, it is time efficient. Providing brief definitions of target words within the context of a story read-aloud takes very little time, perhaps 1 minute per word (Coyne, McCoach, & Kapp, 2007). This allows the introduction of many word meanings during instructional time. Embedded vocabulary instruction can also be incorporated into story readings with minimal disruption to the story flow. Finally, embedded instruction provides students with definitions within a meaningful and supportive context (i.e., a story), two critical features of effective vocabulary instruction (National Reading Panel [NRP], 2000; Stahl & Fairbanks, 1986).

There are also limitations to embedded vocabulary instruction, primarily related to the minimal instructional time allocated for teaching each word. Research has indicated that multiple and repeated exposure to word meanings is a critical feature of effective vocabulary instruction (NRP, 2000; Stahl & Fairbanks, 1986). Limited time restricts the number of exposures students have to target words in embedded vocab-

ulary instruction. These exposures are also constrained by the story context. In other words, because students are introduced to new words within a fixed narrative, they do not receive experiences with or exposure to words in varied contexts. Finally, with embedded instruction, students do not actively engage in learning tasks that require them to discriminate, manipulate, and interact with word meanings.

### **Teaching for Depth: Extended Vocabulary Instruction**

Although breadth (i.e., the number of word meanings in a student's lexicon) is one important dimension of vocabulary knowledge, depth (i.e., how well the student knows those word meanings) is another significant dimension. Knowledge of each word meaning exists on a continuum from no knowledge to varying levels of partial knowledge to more complete and full knowledge (Nagy & Scott, 2000; Schwanenflugel, Stahl, & McFalls, 1997). Depth of word knowledge has important implications for listening or reading comprehension. How well, or deeply, a word is known determines whether or not it can be discriminated from other words and understood in novel contexts or in different morphosyntactic forms. Therefore, another critical goal of direct vocabulary instruction is to help students develop sufficient depth of word knowledge to support comprehension.

Instructional approaches that focus on developing depth of vocabulary knowledge most often provide students with extended opportunities to discuss and interact with words outside story readings. This approach characterizes the work of Beck, McKeown, and their colleagues (Beck, McKeown, & Kucan, 2002). In the first of two studies, Beck and McKeown (2007) reported that oral vocabularies of kindergarten and first-grade students were enhanced by providing rich instruction following storybook readings. Rich instruction included explanations of word meanings and the

presentation of selected words in the context of the original story and new contexts. Students were also asked to discriminate among potential exemplars of word meanings and to offer their own examples of word meanings. Instructed children were compared to children who had not received vocabulary instruction on a picture-recognition task that involved making interpretations in novel contexts. Children who received rich instruction demonstrated significantly higher gains in target words. In a second study, instructed students learned twice as many words when provided with multiple interactions with words across several days.

A number of benefits are associated with extended vocabulary instruction. Because extended instruction allocates more instructional time per word, students receive more encounters with and exposure to target vocabulary. Extended vocabulary instruction also provides students with opportunities to interact with words outside the narrative constraints of the story. This allows teachers to give students examples of how target words can be used in multiple and novel contexts. Activities associated with extended instruction also focus on enabling students to engage in rich dialogic interactions around words and word meanings. The deep and refined knowledge of word meanings that is the goal of extended instruction may better support comprehension across varied contexts (NRP, 2000; Stahl & Fairbanks, 1986). Finally, there is an implicit hypothesis that extended instruction may lead to greater word consciousness or metalinguistic awareness (Nagy, 2007). Because this instruction encourages deep processing of word meanings and challenges students to move beyond memorizing simple dictionary definitions to understanding words at a richer, more complex level, students may become more attuned to novel words they encounter and better able to infer word meanings incidentally (Nagy, 2007; Nagy & Scott, 2000).

The primary limitation of extended vocabulary instruction is the amount of time required to teach each target word. Ideally,

teachers would have unlimited instructional time to spend teaching vocabulary. However, in reality, teachers have limited time. Therefore, an extended approach to vocabulary instruction compels teachers to focus on teaching fewer words. Finally, because evidence to guide teachers in choosing which words to target for instruction is limited, extended instruction forces teachers to make difficult decisions about what words to teach (Coyne, Simmons, Kame'enui, & Stoolmiller, 2004).

### **Evaluating the Effects of Different Approaches to Vocabulary Instruction**

Studies have evaluated the effects of either embedded or extended vocabulary instructional approaches separately on young school-age children's vocabulary learning (e.g., Beck & McKeown, 2007; Biemiller & Boote, 2006; Coyne et al., 2004). Few studies, however, have directly compared these approaches with each other. We recently conducted a study in which we evaluated the effects of both types of instruction on the word learning of 32 kindergarten students from an urban school in the Northeast (Coyne et al., 2007). Children listened to three readings of a storybook in small groups of four for a week. All students received embedded instruction on three words from the storybook and extended instruction on three additional words in a within-subjects experimental design. Embedded instruction included providing simple definitions of target words within the context of the story. Extended instruction included providing simple definitions as well as extended activities after the story reading. Target vocabulary words receiving embedded and extended instruction were counterbalanced across groups to control for word effects. Measures administered at posttest and a 6-week delayed posttest included experimenter-designed assessments of expressive and receptive knowledge of target word definitions and a

receptive measure of target words used in novel contexts. Results indicated that extended instruction resulted in greater word learning than embedded instruction in five of six comparisons and that these differences were maintained on delayed posttests.

Although these findings support the efficacy of extended instruction, a number of important considerations must be addressed when directly comparing the relative effects of embedded and extended vocabulary instruction. In intervention studies, researchers most often use inferential statistics to determine if instruction produces statistically significant differences between treatment and comparison groups. In this case, one would be interested in whether the additional investment in time and intensity associated with extended instruction would produce greater effects on vocabulary outcome measures (Coyne et al., 2007). However, because extended and embedded instruction have different goals (i.e., depth vs. breadth of vocabulary knowledge), statistical tests tell only part of the story. To get a full picture of the relative benefits of the two types of instruction, researchers need to describe the amount and quality of word learning that occur within each approach.

One indicator of word learning is the number of target words learned. For example, if 10 target words were introduced, how many of those words did students learn with each approach? Although this appears to be a straightforward question, the answer is highly dependent on how researchers define what it means to know a word. For example, one might say that a student knows a word when she can recognize its definition. Alternately, one might say that a student knows a word only when she can produce its definition. Therefore, a second indicator of word learning is quality or depth of word learning.

Nagy and Scott (2000) suggested that vocabulary learning is incremental. According to this theoretical model, a student's knowledge of a word's meaning de-

velops incrementally from no knowledge, through varying levels of partial knowledge, to more full and complete knowledge. In other words, "knowing a word is not an all-or-nothing proposition" (Beck et al., 2002, p. 9). Beck, McKeown, and Omanon (1987) developed the following continuum of levels of word knowledge: (a) no knowledge; (b) general sense; (c) narrow, context-bound knowledge; (d) generalized receptive knowledge; and (e) rich, decontextualized knowledge of a word's meaning and its relation to other words. Similar levels of word knowledge have been proposed by Calfee and Drum (1986) and Dale (1965). It is important to note, however, that these are only descriptive examples of what differing levels of word knowledge might look like. In reality, connectionist models of word learning suggest that the development of word knowledge is more likely to be truly incremental, with knowledge of a word's meaning becoming gradually more refined with every new exposure to that word (Landauer & Dumais, 1997).

In the current study, we were interested in replicating and extending our previous findings by attempting to better characterize the amount and quality of word learning that resulted from embedded and extended vocabulary instruction. We developed measurement tasks that would capture different levels of partial word knowledge. In this way, we believed that we would be better able to describe the breadth and depth of vocabulary learning.

### **Responsiveness to Vocabulary Intervention**

Research on shared storybook reading has suggested that students who are at risk for language and literacy difficulties and have smaller initial vocabularies are less likely than their peers with larger vocabularies to learn words incidentally while listening to stories (Coyne et al., 2004; Nicholson & Whyte, 1992; Robbins & Ehri, 1994; Senechal, Thomas, & Monker, 1995). Research on direct

vocabulary instruction with young students has produced inconsistent findings on the relation between individual differences and responsiveness to vocabulary interventions. Although a few studies have showed that all students responded similarly to instruction (Coyne et al., 2004; Elley, 1989), other researchers have found that at-risk students are less responsive to vocabulary instruction (Coyne et al., 2007; Penno et al., 2002).

We were interested in whether students most at risk for language and reading disabilities responded differentially to embedded and extended vocabulary instruction. Overall receptive vocabulary knowledge is a strong predictor of language and comprehension outcomes (Scarborough, 2005; Storch & Whitehurst, 2002). Therefore, we used students' pretest Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997) scores as an indicator of risk.

### **Purpose of the Study**

The purpose of this study was to compare the effects of two interventions that differed in their approach to direct vocabulary instruction for young students. The goal of embedded instruction is to promote breadth of vocabulary knowledge by introducing target word meanings during storybook readings in a time-efficient manner. The goal of extended instruction is to promote depth of vocabulary learning by allocating more instructional time per target word so that students can have multiple opportunities to interact with words outside the context of the story. We compared these approaches both statistically and descriptively to better characterize the amount and quality of word learning that resulted from each. We also compared both approaches with a control condition consisting of incidental exposure to target words within the story reading. Our goal was to provide data so that researchers and practitioners can make more informed decisions about the benefits associated with different approaches to vocabulary instruction, as well as the trade-offs in instructional time and in-



tensity that are required of each. Additionally, we were also interested in whether students maintained vocabulary knowledge gained through instruction at an 8-week delayed posttest. Finally, we examined whether students' responsiveness to vocabulary instruction was moderated by general receptive vocabulary knowledge measured at pretest.

## Method

### Participants

Participants in this study included 42 kindergarten students enrolled in a PreK–8 elementary school in a large city in the Northeast. Approximately 69% of the students in the school were Hispanic, 24% were black, and 6% were white. Approximately 65% were eligible for free or reduced-price lunch. On the 2005 state mastery test, 8% of fourth graders met the state goal in reading compared to 53.8% of students statewide.

Consent forms were sent home to parents of all 58 students in the three kindergarten classrooms. Fifty consent forms were signed and returned to the teachers. Because of space constraints, we randomly selected 45 students from this pool to take part in the study. Three students were absent for two of three intervention sessions and were therefore not included in the analyses. The 42 participants included 23 females and 19 males; 12 students were Hispanic, 8 were African American, 3 were Caucasian, and 1 student was Asian. The average age of participants was 5 years, 4 months and ranged from 4 years, 10 months to 6 years, 1 month.

### Design

We used an experimental design with two within-subjects factors: type of instruction (story words taught with extended instruction, story words taught with embedded instruction, and story words receiving incidental exposure) and time (posttest and

delayed posttest). In a within-subjects design, each student receives each instructional condition and serves as her own control. There are a number of benefits of using this method to investigate vocabulary instruction. First, because each student received all three instructional procedures, we were able control for between-subject variability, thus increasing our power considerably. Second, in vocabulary instruction, directly teaching the meaning of one word does not immediately affect learning the meanings of other words, especially during short, focused intervention studies. Therefore, carryover effects that often make within-subjects designs unfeasible in educational research are eliminated. Third, within-subjects designs control for almost all threats to internal validity, making this method particularly rigorous and trustworthy. A number of research teams have used within-subjects designs successfully in research on vocabulary instruction and intervention (e.g., Beck & McKeown, 2007; Bimiller & Boote, 2006), and we have used this method in our previous research (Coyne et al., 2007).

### Instructional Conditions

We developed the intervention around the storybook *Goldilocks* by James Marshall (1998). We selected nine target words using the following procedures. First, we identified words in the story that appeared only once and were located near pictures that illustrated their meanings. We then replaced these words with low-frequency synonyms that kindergartners would be unlikely to know. Our goal in selecting target words was twofold. First, we wanted target words that were unfamiliar to students but whose meanings they would be able to understand. These are the features of words that Beck et al. (2002) characterized as Tier II words—words they recommend for teaching directly to primary-grade students. Second, to assess the strength of the intervention accurately, we

SEPTEMBER 2009

wanted to ensure that students had no experience with any of the target words. Therefore, we selected rare words of very low frequency. The nine target words included six nouns (*weald*, *duvet*, *domicile*, *parlor*, *lass*, and *shards*) and three adjectives (*torrid*, *fatigued*, and *dismayed*).

It is important to note that, although our words were selected to approximate the kinds of Tier II words that would be targeted for instruction in kindergarten classrooms, we do not suggest the target words we used be taught to kindergarten students, because of their low frequency. Instead, we recommend that teachers select higher-frequency Tier II words that students would be likely to encounter often in academic discourse (Beck et al., 2002).

We developed three instructional conditions, two intervention conditions (i.e., words taught with embedded instruction or extended instruction) and a comparison condition (i.e., words receiving only incidental exposure). We created three versions of the intervention in which target words were counterbalanced across conditions. Each version included three words that were taught using extended instruction, three words taught using embedded instruction, and three words receiving incidental exposure. Target words appeared in the story in a random order to control for order effects. Each group of three words included two nouns and one adjective. Children were assigned randomly to version A, version B, or version C of the intervention, and instructional procedures were identical across the three versions.

All students listened to three readings of the story and were exposed to all nine target words during each reading. Students in version A were taught *weald*, *torrid*, and *duvet* through extended instruction; *domicile*, *parlor*, and *fatigued* through embedded instruction; and *lass*, *shards*, and *dismayed* through incidental exposure. Students in version B were taught *domicile*, *parlor*, and *fatigued* through extended instruction; *lass*, *shards*, and *dismayed* through embedded in-

struction; and *weald*, *torrid*, and *duvet* through incidental exposure. Students in version C were taught *lass*, *shards*, and *dismayed* through extended instruction; *weald*, *torrid*, and *duvet* through embedded instruction; and *domicile*, *parlor*, and *fatigued* through incidental exposure.

The intervention was delivered in small groups of three to four students by four graduate students who were members of the research team. Each graduate student taught one group receiving version A of the intervention, one group receiving version B, and one group receiving version C. The intervention was delivered in three sessions over the course of 1 week. The storybook was read during each session. Each session lasted approximately 30 minutes, with 15 minutes dedicated to reading the storybook and introducing target words and 15 minutes dedicated to postreading vocabulary activities.

Prior to the intervention, graduate students were trained to implement it. Training focused on fidelity of implementation and the standardization of administration across interventionists. Procedures for maintaining student attention and eliciting student responses were explained. Trainers modeled implementation of the intervention and provided graduate students opportunities to practice administration with corrective feedback.

**Embedded instruction.** Graduate students introduced words receiving embedded instruction prior to the storybook reading and instructed students to repeat each target word aloud. Students were told to listen for the target words and raise their hands when they heard the words in the story. When a word was encountered in the storybook, the interventionist asked students to identify the target word and then reread the sentence containing the word (e.g., "Oh, good. Some of you raised your hands! What word did you hear? Yes, lass. 'Once there was a lass named Goldilocks.'"). Students were then provided with a simple definition of the word (e.g., "A lass is a little

girl.”). Next, the graduate student reread the sentence and replaced the target word with its definition (e.g., “Now I’ll say the sentence again with the words that mean lass. ‘Once there was a little girl named Goldilocks.’”). The interventionist then showed the students the supportive picture that accompanied the target word (e.g., “In the picture we can see that Goldilocks is a lass, or a little girl.”). Finally, students were prompted to pronounce the target word together once again (e.g., “Everyone say lass.”). These procedures were repeated for each reading of the storybook, so that each target word receiving embedded instruction was introduced and defined a total of three times. Teaching words with embedded instruction took approximately 30 seconds per word over the three readings. Procedures for words taught with embedded instruction were similar to those in other studies that provided this type of instruction (Biemiller & Boote, 2006; Justice, Meier, & Walpole, 2005; Penno et al., 2002).

**Extended instruction.** Words receiving extended instruction were introduced prior to the storybook reading and defined during the storybook reading using the same procedures as those used for words in the embedded instruction condition. In extended instruction, however, graduate students led students in activities following the storybook reading that allowed them to experience the target words in different contexts. Each activity began by reintroducing the target word using the context of the story and the supportive picture. Once the word was reviewed, additional examples of the word’s usage were provided (e.g., “Other things could also be torrid. If a pan on the stove got very hot, the pan would be torrid. A fire is also torrid, or very hot.”).

During the first two intervention sessions, students participated in three activities, one for each word receiving extended instruction. Activities required both group and individual student responses. In one activity, students were shown various pictures and were asked to determine whether or not the pic-

ture was a positive example of the target word. Other activities required students to respond to yes or no and open-ended questions regarding the target words. Examples of yes or no questions include, “Would wood burning in a fire be torrid?” and, “Is the inside of a refrigerator torrid?” Examples of open-ended questions include, “After walking and walking, she finally reached the top of the tall building. Why was she fatigued?” and, “We built a snowman in the parlor. What is silly about that sentence?” Graduate students followed up children’s responses in all activities by either confirming the student’s response (e.g., “Yes! A pizza would be torrid, or hot, if it just came out of the oven.”) or asking the student to justify his or her response (e.g., “Why is that picture an example of a duvet?”).

During the third intervention session, students participated in activities that combined the three target words receiving extended instruction. For example, students were asked open-ended questions containing two target words, such as, “Would you need a duvet if it was torrid in your house?” Student responses were followed up with confirmation or correction from the interventionist. Students were also given prompts to extend or expand their responses when correct in order to encourage them to demonstrate their understanding of both target words. Providing extended instruction took approximately 15 minutes per word over the three readings (5 minutes per word per reading).

**Incidental exposure.** Words in the incidental exposure condition appeared in the story but were not taught directly. Students heard these words in the context of the story once per storybook reading. Interventionists did not directly discuss these words at any time.

### Fidelity of Implementation

Critical components of the interventions were identified and an observation checklist was developed to document and eval-

uate fidelity of implementation (Gersten, Baker, & Lloyd, 2000; Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000). Examples of critical components incorporated into the fidelity checklist included whether or not graduate students delivered each instructional element, modeled procedures appropriately, maximized opportunities to respond, provided error correction, and read storybooks with enthusiasm. The project director observed each interventionist during one of the three instructional sessions. Fidelity of implementation averaged above 90% for each graduate student.

### Measures

The National Reading Panel (2000) concluded that specific vocabulary growth is best assessed through researcher-developed measures because they are more sensitive to gains achieved through instruction than are standardized tests. For this study, we developed measures that assessed children's knowledge of the nine words targeted in the intervention conditions. Our goal was to develop assessment tasks that would capture different levels of word knowledge. We used a standardized instrument, the Peabody Picture Vocabulary Test (PPVT), to assess general vocabulary knowledge. All instruments were administered individually.

*Expressive target word definitions (expressive definitions).* The expressive-definitions measure assessed expressive knowledge of the nine target word definitions. Students were asked to provide a definition for each target word. For example, for the word *domicile*, students were asked, "What does the word *domicile* mean?" Students' responses were recorded verbatim. If a student failed to respond after 5 seconds, or replied "I don't know," the administrator asked a follow-up question, "Can you tell me anything about the word *domicile*?" Children were given 2 points for a complete response (e.g., "A domicile is a house"), 1 point for a partial or related response

(e.g., "A domicile is inside"), and 0 points for an unrelated response, or no response. The maximum score a student could receive for each condition was 6 (i.e., 2 points possible for each of three words).

*Target words used in novel contexts/full knowledge (context/full knowledge).* The context/full knowledge measure assessed receptive knowledge of the target words provided in neutral contexts. This test was designed to assess high levels of target word knowledge by requiring children to make finer discriminations about word meanings. Students were asked questions that required a yes or no answer (Beck & McKeown, 2007). Each target word was represented by four questions—two positive examples of the word's usage and two negative examples. For example, for *dismayed*, the positive examples were, "If you lost your toy, would you be dismayed?" and "Would you be dismayed if you got in trouble for something that you did not do?" The two negative examples were, "If your mother were dismayed, would she be smiling?" and "If your mom made your favorite dinner, would you be dismayed?" The four questions for each target word were separated and distributed across the measure. Students received 1 point for each correct answer and 0 points for each incorrect answer. Each student received a separate total score for target words introduced within each instructional condition (i.e., embedded, extended, incidental exposure). The maximum score a student could receive for each condition was 12 (i.e., 4 points possible for each of three words).

*Target words used in novel contexts/partial knowledge (context/partial knowledge).* The partial-knowledge measure assessed receptive knowledge of the target words provided in supportive contexts. This test was designed to detect low target word knowledge. Items were developed so that students could correctly answer questions without full knowledge of the word meanings. Children were asked to respond to questions that required a yes or no answer (Beck & McKeown, 2007). Each target

word was represented by four questions—two that required a positive response and two that required a negative response. For example, for the word *parlor*, questions that required a positive response were, “Have you ever been in a parlor?” and “Could someone have a parlor in their house?” Questions that required a negative response were, “Could you put a parlor in a bag?” and “Could you find a parlor in your kitchen?” The four questions for each of the nine target words were separated and distributed across the measure. Students received 1 point for each correct answer and 0 points for each incorrect answer. The maximum score a student could receive for each condition was 12 (i.e., 4 points possible for each of three words).

**Receptive target word definitions (receptive definitions).** The receptive-definitions measure assessed receptive knowledge of the nine target word definitions. Students were asked questions that required a yes or no answer (Beck & McKeown, 2007). Each target word was represented by two questions—one that corresponded to the correct definition and one that corresponded to an incorrect definition. For example, the two questions for the target word *duvet* were, “Is a duvet a warm blanket?” and “Is a duvet a fast car?” Children received 1 point for each correct answer and 0 points for each incorrect answer. The two questions for each of the nine target words were separated and distributed across the measure. Each student received a separate total score for target words introduced within each instructional condition (i.e., embedded, extended, incidental exposure). The maximum score a student could receive for each condition was 6 (i.e., 2 points possible for each of three words).

**Peabody Picture Vocabulary Test—III (PPVT).** The PPVT is a norm-referenced measure of receptive vocabulary. We used it to characterize children’s overall receptive vocabulary knowledge prior to the intervention. Students are presented with four pictures and are asked to point to the picture that best represents the word given by the

examiner. Standardized scores (mean = 100;  $SD = 15$ ) are computed based on number of items correct and the student’s chronological age. Reliability of the PPVT is satisfactory, with alternate-forms reliability coefficients ranging from .88 to .96 and test-retest reliability coefficients ranging from .91 to .94. Overall, research had suggested high reliability and validity for the PPVT (Salvia & Yseldyke, 1998).

### Data Collection and Scoring

Data collection took place at pretest, posttest, and delayed posttest. Members of the research team administered all assessments individually to students in a quiet location. We collected pretest data approximately 1 week prior to the start of the intervention. Measures administered at pretest included the PPVT and the context/full knowledge and context/partial knowledge measures. We collected posttest data between 1 and 5 days after the third reading of the storybook; these data included the expressive definitions, context/full knowledge, context/partial knowledge, and receptive definitions measures. The delayed posttest was administered 8 weeks after posttest and included the expressive and receptive definitions measures. We administered measures in the order they were described. We collected posttest assessments in one 30-minute session. Pretest and delayed posttest assessments were each collected in one 20-minute session.

Data collectors were required to demonstrate at least 90% reliability for administration. All measures were scored by one member of the research team. The project director randomly selected and independently scored 20% of the assessment protocols to check for scoring reliability. Agreement was 100%.

### Results

We administered both the context/full knowledge and context/partial knowledge assessments at pretest. Students’ to-

TABLE 1. Means and Standard Deviations at Posttest and Delayed Posttest, by Condition

Measure	Extended Instruction		Embedded Instruction		Incidental Exposure	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Expressive definitions: <sup>a</sup>						
Posttest	3.79	2.06	1.47	1.74	.24	.63
Delayed posttest	1.82	1.78	1.11	1.48	.24	.88
Receptive definitions: <sup>a</sup>						
Posttest	3.97*	1.38	3.58*	1.00	3.03	1.06
Delayed posttest	4.06*	1.37	3.61*	1.08	3.08	1.08
Context/full knowledge <sup>b</sup>	7.54*	1.79	6.34	1.62	5.98	1.44
Context/partial knowledge <sup>b</sup>	7.85*	1.75	6.90*	1.62	6.15	1.42

<sup>a</sup>Range = 0 to 6.

<sup>b</sup>Range = 0 to 12.

\*Scores above chance level ( $p < .0033$ ).

tal scores on the preassessments indicated that they were not performing above the chance level on either the context/full knowledge assessment (mean = 17.97 out of 36;  $t = .057$ ,  $p = .96$ ) or the context/partial knowledge subtest (mean = 18.46 out of 36;  $t = 1.51$ ,  $p = .14$ ). Scores at the chance level would indicate that students performed no better than guessing on the yes/no questions. We also compared students' scores on the full and partial knowledge subtests by condition using repeated-measures analysis of variance (ANOVA). These two analyses indicated that the pretest scores on the full knowledge and partial knowledge assessments were similar across conditions ( $p > .16$  for full knowledge;  $p > .32$  for partial knowledge). Scores on the pretest measures indicated that students had no measurable knowledge of target word meanings prior to receiving the intervention and that there were no differences between students' knowledge of words across instructional conditions. Participants' mean PPVT score measured at pretest was 95.24 ( $SD = 9.73$ ).

#### Comparison of Instructional Conditions

To determine whether there were statistically significant differences among the

means of words taught in the three conditions and whether these differences persisted over time, we conducted a series of repeated-measures ANOVAs on the measures we collected at posttest and delayed posttest (receptive and expressive definitions). The two within-subjects variables were time (posttest and delayed posttest) and instructional condition (words taught with extended instruction, words taught with embedded instruction, and words receiving incidental exposure). Descriptive statistics for these analyses are reported in Table 1. Cohen's  $d$  effect sizes for the comparisons between the instructional conditions are reported in Table 2.

For the expressive-definitions measure, the main effects of word learning condition (Wilks's  $\lambda = .284$ ;  $F = 45.42$ ,  $p < .001$ ), and time (Wilks's  $\lambda = .467$ ;  $F = 42.25$ ,  $p < .001$ ), and the interaction between time and condition (Wilks's  $\lambda = .518$ ;  $F = 16.76$ ,  $p < .001$ ) were significant. The linear trend was significant, indicating that the mean score of words taught with extended instruction was higher than that for embedded instruction, which in turn was higher than that for incidental exposure at both posttest and delayed posttest. However, the interaction revealed that students also lost the greatest amount of knowledge between posttest and delayed posttest for words taught with

TABLE 2. Cohen's *d* Effect Sizes for Comparisons between Instructional Conditions at Posttest and Delayed Posttest

Measure	Extended vs. Embedded	Extended vs. Incidental	Embedded vs. Incidental
Expressive definitions:			
Posttest	1.34**	2.57**	.87**
Delayed posttest	.44*	1.18**	.72**
Receptive definitions:			
Posttest	.70**	.97**	.24*
Delayed posttest	.33	.84**	.52
Context/full knowledge	.38**	.91**	.63**
Context/partial knowledge	.56**	1.07**	.49*

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

extended instruction when compared to the other two conditions.

For the receptive-definition measure, the effect of word learning condition was significant (Wilks's  $\lambda = .695$ ;  $F = 7.478$ ,  $p = .002$ ). The effects of time (Wilks's  $\lambda = .994$ ;  $F = .198$ ,  $p = .66$ ) and the interaction between time and condition (Wilks's  $\lambda = .999$ ,  $F = .022$ ,  $p = .98$ ) were not significant. Therefore, there were differences in vocabulary scores among the word learning conditions, and these differences were maintained over time. The linear trend was significant, indicating that extended instruction produced higher scores than embedded instruction, which led to better performance than incidental exposure.

Next, we conducted a series of one-way repeated-measures ANOVAs on the measures we collected at posttest only (context/full knowledge and context/partial knowledge). At posttest, there were significant differences among the three conditions on the context/full knowledge measure (Wilks's  $\lambda = .683$ ,  $p = .001$ ). The linear trend was statistically significant ( $p < .001$ ), indicating that students' scores were higher for words taught with extended instruction, followed by words taught with embedded instruction, followed by words receiving incidental exposure. There were also significant differences among the three conditions on the context/partial knowledge measure (Wilks's  $\lambda = .66$ ,  $p < .001$ ), and the linear trend was significant ( $p <$

.001), showing the same trend as for the context/full knowledge measure.

Finally, we conducted a series of one-sample *t*-tests to determine whether students scored above the chance level on the dichotomous yes/no measures of word learning (receptive definitions, context/full knowledge, and context/partial knowledge). Scores at the chance level would indicate that students performed no better than guessing on the yes/no questions. We used a Bonferroni adjustment (.05/15) to account for the large number of statistical tests being conducted. Therefore, probability levels below .0033 were considered significant. Scores above the chance level are indicated with an asterisk in Table 1.

For the receptive-definitions measure administered at posttest, students scored above the chance level on words that were taught using extended ( $t = 4.96$ ,  $p < .001$ ) and embedded ( $t = 3.81$ ,  $p < .001$ ) instruction. However, students scored at the chance level on words that received incidental exposure ( $t = -.44$ ,  $p = .66$ ). Results were similar for the receptive measure at delayed posttest. Again, students performed above the chance level on words taught using extended ( $t = 5.02$ ,  $p < .001$ ) and embedded ( $t = 3.6$ ,  $p = .001$ ) instruction. However, they scored at the chance level on words that received incidental exposure ( $t = .453$ ,  $p = .65$ ).

For the context/full knowledge measure, children scored above the chance

SEPTEMBER 2009

level on words taught using extended instruction ( $t = 5.50, p < .001$ ). However, they performed at the chance level on words that received embedded instruction ( $t = 1.35, p = .19$ ) and incidental exposure ( $t = -.11, p = .91$ ). For the context/partial knowledge measure, students scored above the chance level on words that were taught using extended ( $t = 6.77, p < .001$ ) and embedded ( $t = 3.56, p = .001$ ) instruction. However, they performed at the chance level on words that received incidental exposure ( $t = -.66, p = .51$ ).

### Characterizing Word Learning

We were also interested in describing the effects of the instructional conditions on breadth and depth of vocabulary learning. Breadth refers to how many words, on average, students learned in each condition, and depth indicates how well they learned those words.

To characterize breadth of word knowledge, we examined the mean scores on each measure of target word learning and estimated how many words were learned based on the criteria for that assessment. For example, a mean score on the expressive-definitions measure of 6 points out of 6 for words from one of the three approaches would indicate that students had learned all three words taught in that condition (i.e., they produced complete 2-point definitions for each of the three target words). A mean score of 4 would suggest that children learned the equivalent of two of the three words. Finally, a mean of less than 2 would indicate that students, on average, could not produce a complete definition of any word.

For the full and partial knowledge context measures, there were four yes/no questions for each word. Therefore, 12 would be a perfect score and indicate that students learned all three of the words taught in that condition according to the criterion of these measures. A score of 8 would mean that children learned the equivalent of two of the three words (i.e.,

four correct yes/no questions for each of two words). If analyses showed that the mean score was statistically no different from chance, this would suggest that students, on average, did not learn any of the words. On the receptive-definitions measure, a mean score of 6 out of 6 would mean that students correctly answered two yes/no questions about each of the three target words in that condition. In contrast, a score of 3 on these measures would indicate that students performed no better than chance.

To characterize depth of word learning, we examined mean scores across measures of target word learning that assessed levels of partial word knowledge. Table 3 summarizes our characterization of amount and quality of word learning within the three conditions across each measure. To simplify interpretation, we rounded off mean scores on each measure before estimating the number of words learned.

### Responsiveness to Instruction

The PPVT scores predicted responsiveness to instruction on receptive and expressive target word measures for both the extended and embedded instructional conditions. Students with higher initial PPVT scores demonstrated greater word learning across the two instructional conditions and the two instructional measures. The relation between PPVT scores and students' scores on the expressive-definitions measure was moderately strong for both words taught with extended instruction ( $r = .49$ ) and words taught with embedded instruction ( $r = .41$ ). We tested the difference between the two dependent correlation coefficients for significance using the formula Glass and Hopkins (1996) recommended. This difference was not significant ( $t = .57, p > .10$ ), suggesting that the relation between PPVT and expressive-definitions scores did not differ across the two instructional conditions. The relation between PPVT and scores on the receptive-definition measure was also moderately



TABLE 3. Breadth of Word Learning across Measures, by Condition

Condition/Time per Word	Expressive Definitions	Context/Full Knowledge	Context/Partial Knowledge	Receptive Definitions
Incidental (30 seconds):				
Words learned <sup>a</sup>	0	0	0	0
Mean score	.24	5.98*	6.15*	3.03*
Perfect score	6.00	12.00	12.00	6.00
Embedded (3 minutes):				
Words learned <sup>a</sup>	0	0	2	2
Mean score	1.47	6.34*	6.90	3.58
Perfect score	6.00	12.00	12.00	6.00
Extended (15 minutes):				
Words learned <sup>a</sup>	2	2	2	2
Mean score	3.79	7.54	7.84	3.97
Perfect score	6.00	12.00	12.00	6.00

<sup>a</sup>Students could potentially learn three words for each measure. Number of words learned was estimated.

\*Scores below chance level.

strong for both the extended ( $r = .45$ ) and embedded ( $r = .49$ ) conditions. This difference was not significant ( $t = .35, p > .10$ ), indicating that the relation between PPVT and receptive-definition scores also did not differ across the two instructional conditions.

## Discussion

### Comparison of Instructional Conditions

Our results indicated that there were statistically significant differences at posttest favoring words taught with extended and embedded instruction over words receiving only incidental exposure during story reading on all measures. These findings are consistent with a growing body of research documenting the efficacy of directly teaching word meanings to young students within oral language activities such as storybook readings (Elley, 1989; Justice et al., 2005; Penno et al., 2002; Walsh & Blewitt, 2006). Moderate to large effect sizes for these comparisons indicate that direct instruction of vocabulary results in reliably greater word learning in kindergarten students than does incidental exposure by itself.

Even with the short duration of the study, we found significant differences across all measures at posttest between

words taught with extended and embedded instruction. Although few studies have compared these types of instruction, our results converge with other findings suggesting that increasing instructional time and providing more exposure to target vocabulary in varied contexts lead to enhanced word learning (Beck & McKeown, 2007; Biemiller & Boote, 2006; Coyne et al., 2007).

### Breadth and Depth of Word Learning

Although the results of statistical tests are important, we were particularly interested in describing the effects of embedded and extended instruction. We wanted to examine the relative benefits related to the amount and quality of word learning that occurred within each approach. We characterized breadth by examining mean scores on each measure and estimating, on average, the number of words students in each of the three conditions learned. We described depth by looking at students' mean scores across measures that required varying levels of word knowledge to answer assessment items correctly.

Results suggested that, with embedded instruction, students demonstrated measurable word learning on approximately two-thirds of the words introduced. How-

SEPTEMBER 2009

ever, word learning was evident on only two of our four measures, suggesting that students possessed relatively little knowledge of those words. For example, students recognized correct and incorrect definitions of target words and answered yes/no questions that required low levels of word knowledge. However, they were unable to produce definitions of target words or respond to yes/no questions that required additional word knowledge. For words taught with extended instruction, students also learned approximately two out of three words. In contrast to embedded instruction, however, this finding was consistent across all measures, suggesting that students developed more complete knowledge of these words.

These descriptive findings help to illuminate both the strengths and limitations of different approaches to direct vocabulary instruction. Embedded instruction (i.e., providing brief definitions of words within the context of a story) is time efficient (i.e., 30–60 seconds per word per story reading) and allows the introduction of many word meanings within limited instructional time. However, word learning appears to be limited to partial knowledge of word meanings. Extended instruction (i.e., providing opportunities to discuss and interact with words outside of story reading) requires more instructional time (i.e., 5 minutes per word per story reading) but provides multiple exposures to target words in varied contexts that result in more full and refined knowledge of these words.

#### Durability of Word Learning

We administered a delayed posttest 8 weeks after the conclusion of the intervention to examine the durability of word learning without planned review or reinforcement. We selected low-frequency vocabulary words that kindergarten students would not know and that were unlikely to be encountered incidentally during school. We also confirmed with classroom teachers

that target words were not discussed in class between the posttest and the delayed posttest.

Results indicated that there was no effect of time on the receptive-definitions measures, suggesting that students' ability to recognize the meanings of target words did not decrease between posttest and delayed posttest. On the expressive-definitions measures, however, there was a significant interaction between time and instructional condition indicating that students' ability to produce definitions of target words taught with extended instruction decreased differentially compared to words taught with embedded instruction and incidental exposure. Scores on the expressive-definitions measure for words taught using embedded instruction did not decrease between posttest and delayed posttest. These findings are consistent with our previous research and suggest that word learning is robust over time but that higher and more complete levels of word knowledge may be more susceptible to deterioration, at least initially (Coyne et al., 2007). In our study, students did not receive formal review or even encounter target words incidentally between posttest and delayed posttest. Word knowledge may not become permanent or fully established unless students receive continued encounters with target words over time or are provided with thoughtful and systematic review.

#### Responsiveness to Vocabulary Instruction

In our previous research, we found that individual differences in initial receptive vocabulary knowledge strongly predicted responsiveness to extended vocabulary instruction (Coyne et al., 2007). We replicated this finding in the current study where students' pretest PPVT scores were highly correlated with target vocabulary outcome measures. We found a similar relation between initial vocabulary knowledge and response to embedded instruction. These findings suggest that students most at risk of language and

literacy difficulties are less responsive to direct vocabulary instruction than their typically developing peers (Penno et al., 2002).

### Implications for Practice

Our findings underscore the trade-offs between different approaches to direct vocabulary instruction and reinforce the importance of carefully considering learning goals when making decisions about vocabulary instruction for young students. For example, embedded instruction may help students establish initial lexical representations of a large number of new vocabulary words. However, our results suggest that the limited exposure to target words associated with this approach is unlikely to result in deeper word knowledge. Therefore, if the goal is to support students in developing fuller and more complete word knowledge, extended instruction would be more effective, although it requires more instructional time and limits the number of words that can be introduced.

We believe there is a place for both embedded and extended vocabulary instruction in primary classrooms. We have described this as a tri-level approach to vocabulary instruction (Coyne et al., 2007). It includes (a) reading storybooks to children that contain varied and complex vocabulary, (b) providing embedded instruction on a subset of targeted words contained in the storybook, and (c) providing extended instruction on a second set of words from the story.

Embedded instruction should target words that students will encounter often in academic discourse but are not immediately critical for comprehending the story. Embedded instruction may help students establish an initial lexical representation of words and provide a foundation that will assist them in refining and consolidating word knowledge when they encounter words subsequently, either incidentally or through additional instruction. Words chosen for extended instruction, in contrast, should be immediately essential for under-

standing important ideas and concepts in the story. Extended instruction would help students develop more complete word knowledge that would support improved listening comprehension of the story. A tri-level approach such as this would leverage instructional time efficiently to target the equally important goals of helping students develop both breadth and depth of vocabulary knowledge.

We found that higher and more complete levels of word knowledge were more difficult to maintain when students did not receive continued encounters with target words. This suggests that teachers who provide direct vocabulary instruction should consider ways to review target words over time and ensure that students continue to be exposed to words after they have been introduced. This may be especially important for more sophisticated or low-frequency target words that students are less likely to encounter incidentally.

Finally, our findings revealed that students with less receptive vocabulary knowledge did not make learning gains as large as their peers with more developed vocabularies. This finding indicates that even directly teaching target words is insufficient to close the vocabulary gap among students (Biemiller, 2001). The growing evidence supporting the strong relation between initial vocabulary knowledge and responsiveness to instruction highlights the need to intensify instruction for students most at risk for language and literacy difficulties. It is likely that these students will require additional intervention above and beyond general classroom instruction to make gains similar to their peers who are not at risk (Loftus, Coyne, McCoach, & Zipoli, 2008).

### Limitations and Future Research

This study was of very short duration, and students were taught the meanings of only three words in each condition. Additionally, although we selected words that would approximate the kinds of Tier II

words that would typically be taught in kindergarten classrooms, our target vocabulary consisted of low-frequency words that may not have been representative of higher-frequency Tier II words.

Although many studies have used storybook readings as a context for teaching vocabulary, there are some possible limitations to this approach as well. Even though providing brief definitions of vocabulary during a story reading takes very little time, these interruptions may disrupt listening comprehension, especially if they occur frequently. The possibility that vocabulary instruction may interfere with comprehension may also limit the number of words that can be introduced during a given story reading.

Finally, we examined only proximal measures of target word learning. It will be important for future research to replicate these findings as well as to investigate the effects of long-term implementation of both embedded and extended instruction on both proximal measures of word learning and more distal measures of language and literacy such as generalized vocabulary knowledge and listening comprehension.

## References

- Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator*, *16*(4), 14–18, 44–47.
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal*, *107*(3), 251–271.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford.
- Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Erlbaum.
- Biemiller, A. (2001). Teaching vocabulary: Early, direct, and sequential. *American Educator*, *25*(1), 24–29.
- Biemiller, A., & Boote, C. (2006). An effective method for building vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44–62.
- Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology*, *93*(3), 498–520.
- Calfee, R. C., & Drum, P. A. (1986). Research on teaching reading. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3d ed., pp. 804–849). New York: Macmillan.
- Catts, H. W., Hogan, T. P., & Adolf, S. M. (2005). Developmental changes in reading and reading disabilities. In H. W. Catts & A. G. Kahmi (Eds.), *The connections between language and reading disabilities* (pp. 25–40). Mahwah, NJ: Erlbaum.
- Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disabilities Quarterly*, *30*, 74–88.
- Coyne, M. D., Simmons, D. C., Kame'enui, E. J., & Stoolmiller, M. (2004). Teaching vocabulary during shared storybook readings: An examination of differential effects. *Exceptionality*, *12*(3), 145–162.
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, *42*, 895–901.
- Dunn, L., & Dunn, L. (1997). *Peabody Picture Vocabulary Test—Revised*. Circle Pines, MN: American Guidance Service.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*(2), 174–187.
- Gersten, R., Baker, S., & Lloyd, J. W. (2000). Designing high quality research in special education: Group experimental design. *Journal of Special Education*, *34*(1), 2–18.
- Glass, G. V., & Hopkins, K. D. (1996). *Statistical methods in education and psychology*. Boston: Allyn & Bacon.
- Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research and Practice*, *15*(4), 198–205.
- Hart, B., & Risley, R. T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Brookes.
- Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners.

- Language, Speech, and Hearing Services in Schools*, **36**, 17–32.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, **104**(2), 211–240.
- Loftus, S., Coyne, M. D., McCoach, D. B., & Zipoli, R. (2008). *Effects of a Tier 2 vocabulary intervention on the word knowledge of kindergarten students at risk for language and literacy difficulties*. Manuscript submitted for publication.
- Marshall, J. (1998). *Goldilocks and the three bears*. New York: Puffin.
- 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). New York: Guilford.
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269–284). Mahwah, NJ: Erlbaum.
- National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups*. Bethesda, MD: National Institute of Child Health and Human Development.
- Nicholson, T., & Whyte, B. (1992). Matthew effects in learning new words while listening to stories. In C. K. Kinzer & D. J. Leu (Eds.), *Literacy research, theory, and practice: Views from many perspectives. Forty-first yearbook of the National Reading Conference* (pp. 499–503). Chicago: National Reading Conference.
- Penno, J. F., Wilkinson, I. A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, **94**(1), 23–33.
- Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology*, **86**(1), 54–64.
- Salvia, J., & Ysseldyke, J. E. (Eds.). (1998). *Assessment* (7th ed.). Boston: Houghton Mifflin.
- Scarborough, H. (2005). Developmental relationships between language and reading: Reconciling a beautiful hypothesis with some ugly facts. In H. W. Catts & A. G. Kamhi (Eds.), *The connections between language and reading disabilities* (pp. 3–24). Mahwah, NJ: Erlbaum.
- Schwanenflugel, P. J., Stahl, S. A., & McFalls, E. L. (1997). Partial word knowledge and vocabulary growth during reading comprehension. *Journal of Literacy Research*, **29**, 531–553.
- Senechal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology*, **87**, 218–229.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, **56**(1), 72–110.
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology*, **38**(6), 934–947.
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, **11**(1), 3–32.
- Walsh, B. A., & Blewitt, P. (2006). The effect of questioning style during storybook reading on novel vocabulary acquisition of preschoolers. *Early Childhood Education Journal*, **33**(4), 273–278.