Contemporary Educational Psychology 52 (2018) 15–24

Contents lists available at ScienceDirect

Contemporary Educational Psychology

journal homepage: www.elsevier.com/locate/cedpsych

Developing low-income children's vocabulary and content knowledge through a shared book reading program

Susan B. Neuman[a,*], Tanya Kaefer[b]

[a] New York University, USA
[b] Lakehead University, Canada

A B S T R A C T

This study examines the effects of a shared book intervention designed to improve low-income children's oral language vocabulary and content knowledge in science. Classrooms (preK-through grade 1) from 12 elementary schools in a large metropolitan area were randomly selected into treatment (N = 36) and control groups (N = 34). The year-long intervention involved children in read aloud books about science topics, using cross-cutting concepts and vocabulary within taxonomic categories to build knowledge networks. Pre- and post-tests examined child outcomes in vocabulary, science concepts, language, and knowledge of the information genre. Results indicated that pre-K and kindergartners' learned significantly more words and science concepts than controls. Growth for ELL students exceeded that of native English speakers. Standardized scores in language, however, remained largely flat.

1. Introduction

Children need both word and world knowledge to learn how to read (Hirsch, 2003; Neuman, Pinkham, & Kaefer, 2016; Willingham, 2006). Vocabulary knowledge is essential and correlates strongly with oral language and reading performance (Coyne et al., 2013; Storch & Whitehurst, 2002). Nevertheless, to make constructive use of vocabulary, the reader (or listener) will need a threshold of knowledge about the topic (Ouellette, 2006; Perfetti, 2007). Language is full of semantic breaks in which knowledge is assumed and dependent on making correct inferences. Domain knowledge enables the reader or listener to fill in those gaps, leading to a more coherent understanding of text, and a greater potential for retention and inference generation (Stahl, 2003).

Yet despite a growing interest in vocabulary instruction in the early years, few of these programs have focused on its connection to content knowledge (Williams, Stafford, Lauer, Hall, & Pollini, 2009). Although there is an impressive new corpus of research that has embraced a more synergistic relationship with content domains (Cervetti, Barber, Dorph, Pearson, & Goldschmidt, 2012; Gonzalez et al., 2011), the goal of these programs has been to promote vocabulary and has not measured gains in content knowledge. Similarly, programs integrating math and science with literacy skills (Clements & Sarama, 2008; Gelman & Brenneman, 2004) have largely limited their analyses to gains in content and inquiry process skills, not literacy development. To date, there is a dearth of programs that reflect the dual goals for both vocabulary and content knowledge. With the exception of the work of Wright and Gotwals (2017) who developed an integrated approach to literacy and science for kindergartners, there is little evidence for the effectiveness of such instruction on content, language and comprehension.

The problem with a lack of attention in the early prekindergarten and kindergarten years is that acquiring word and domain knowledge is a gradual and cumulative process (Hirsch, 2003). It needs to start early. A full and flexible knowledge of a word includes not only a basic understanding of its meaning but how its meaning may change in different contexts (Stahl & Nagy, 2006). Getting a *hit* in a baseball game has a very different connotation than a *hit* in the boxing ring. Words are embedded in knowledge domains and are sometimes so closely tied that they cannot be defined outside them. Children gain better understanding of words when they are heard frequently, over time, and in multiple content domains (Stahl, 2003).

Starting early is especially crucial given the striking differentials in readiness skills at the 'starting gate.' Poverty represents an enormous factor in school readiness (Neuman & Celano, 2012). Even before entering kindergarten, the average cognitive score of children in the highest socioeconomic group is 60% above the score of the lowest SES group (Lee & Burkam, 2002). These problems are further compounded for children who are non-native English speakers, representing nearly one in five students in American classrooms, and substantially higher in urban school districts (Sparks, 2016). These students will need to develop not only their oral and written language proficiency, but the

* Corresponding author.
E-mail address: sbneuman@nyu.edu (S.B. Neuman).

https://doi.org/10.1016/j.cedpsych.2017.12.001

Available online 09 December 2017

academic language and domain-specific vocabulary in order to benefit from content area instruction (Llosa et al., 2016). Consequently, previous estimates of the ‘word gap’ (Hart & Risley, 1995), the stark differences in the accumulated number of words children are exposed to, may actually underestimate the problems associated with vocabulary differentials and school learning.

Words are more than mere labels however. Rather, in the early years, they are the primary means to convey content knowledge (Pinkham, Kaefer, & Neuman, 2012). Before children can read substantive texts on their own, content is conveyed through oral discourse (Dickinson, McCabe, & Essex, 2006). Therefore, although students will need differential means of support and practice, there are clearly some activities that may strongly enhance language growth and content knowledge. Studies have shown that one of the most important vehicles for rich vocabulary and content is the read-aloud experience (Coyne et al., 2013), in which texts are selected for their quality, substance, and vocabulary, followed by discussion that builds on their understanding of the ideas, topics, and words in the story. For emergent bilingual students, in particular, the read-aloud experience can provide a meaningful context for language learning as well as an authentic purpose for communication (August & Shanahan, 2006).

The aim of the present study was to design a theoretically-based instructional program that could accommodate children from very diverse language backgrounds and previous experiences in and out of school to improve both vocabulary and content knowledge. Using the read-aloud experience, the intervention is based on the premise that bundling together different types of texts on high interest topics along with teachers’ scaffolding support can help to build children’s knowledge networks across important concepts in the life, biological and physical sciences.

2. Book genre as a variable in the read-aloud experience

Programs that combine both vocabulary development and content instruction are likely to require informational text, because the genre supports knowledge about topics like science and social studies (Williams et al., 2005). However, many researchers have reported the dearth of informational text reading in the primary grades. For example, in a survey of over 1000 teachers, prekindergarten to grade 3, Yopp and Yopp (2012) found that informational texts represented only a very small proportion of books read aloud at all grade levels, ranging from 5% in preK to 9% by second grade teachers. In her observational study of first grade classrooms, Duke (2000) found little attention to informational materials on the bulletin boards, classroom libraries, read-aloud and writing activities.

One reason why there might be so little attention to information text, despite the call for its greater use in the Common Core, is that it is considered more difficult than narrative text (Yopp & Yopp, 2006). For example, narrative texts tend to support conversations about characters and their traits, whereas information or expository texts encourages descriptions of the attributes of a topic (Price, van Kleeck, & Huberty, 2009). In addition, narrative texts are more likely to use words that reflect mental states (e.g., *think*, *know*) and temporal connections (e.g., *before*, *after*), while information books tend to use more academic language (e.g., *predict*, *observe*) and conceptually dense words (e.g., *investigation*) (Pappas, 1991). Furthermore, unlike storybooks, informational text often contains diagrams and embedded print within the illustrations and diagrams and other text features like indices, and glossaries. Teachers may not be familiar or particularly comfortable with reading information books aloud compared to traditional storybooks, which have a more predictable story structure (Price, Bradley, & Smith, 2012).

Although recognizing that different genres bring different instructional opportunities, rarely have we considered the possibility of how they might work together to enhance vocabulary and content knowledge. Predictable books, for example, have repetition, rhyme, and cumulative or known sequences that encourage chiming and memory of new words (Dickinson & Smith, 1994). They tend to be highly engaging, and often support co-participation in read-alouds. Other types of books, especially for young children, often blur the lines between genres (Price et al., 2009). Alphabet, counting, and shape books, in many cases, serve dual purposes that include basic skills and scientific concepts. In addition, some information books have narrative structures, with goals, events, and final outcomes. While they are clearly storybooks (e.g. *Fraser Bear* (De Vries, 2010), the story of a young cub), they may include informational concepts and present opportunities for teachers to make these ideas more meaningful to young children. Pappas (1991) described this hybrid genre as narrative nonfiction, mixing the features of the information and narrative genres.

Consequently, organized deliberately, text genre could be used as an instructional scaffold for word and world knowledge. For example, words in rich contexts could be introduced first through predictable text, with its memorable stances serving as mnemonic devices. Narrative nonfiction might introduce similar words and concepts next in an engaging manner for children to understand and recall. Finally, these words and concepts, now more familiar through frequent exposure in multiple contexts, could be introduced through a more traditional expository or informational text, written with the purpose of providing scientific information about the topic. The density of vocabulary and the conceptual load, therefore, might be lessened through a more deliberate stair-step approach to increasing the difficulty in vocabulary, concepts and comprehension of text.

3. Building knowledge networks

When these books are arranged within topics that are coherently structured to develop rich content knowledge both within and across grades, these texts can act as a support to build vocabulary and comprehension (Pollard-Durodola et al., 2012). Such an approach may work especially well with science-related topics, where knowledge building is more hierarchically structured than some other domains such social studies. Cervetti, Wright, and Hwang (2016), for example, examined the effects of having fourth graders read a set of conceptually coherent texts on birds compared to those who read books on a range of topics unrelated to one another. Results of their analysis showed significant gains in vocabulary and knowledge of the concepts in the texts for those who read a set of texts that cohered around those concepts. Similarly, Gonzalez and his colleagues in a study of a content-focused shared book-reading program with low-income preschoolers Gonzalez et al. (2014) reported statistically and practically significant effects (Cohen’s d ranging from 0.93 to 1.41) on standardized measures of vocabulary. Preschoolers benefited from instruction that organized information into meaningful knowledge networks with opportunities for repeated practice and extended learning.

Science is a well-structured domain. Many of the concepts developed in scientific domains (e.g. physical, life, early and space science) are based on taxonomic categorization with shared superordinate categories or functions (Gelman & Kalish, 2006). For example, whales, dolphins, and manatees are taxonomically-related because they belong to the class of marine mammals; pigeons, penguins, and flamingos are taxonomically-related because they belong to the same basic class of birds. By extension, both are a sub-class of wild animals. These categories are typically universal (e.g., birds are animals in North America as well as Europe), conventional (e.g., most people agree that birds are animals), and constant (e.g., birds remain animals even if their feathers are plucked) (Wellman & Gelman, 1998).

When these taxonomic categories are made explicit, they can help children develop a more coherent conceptual framework through which knowledge is encoded, organized, and remembered (Gelman & Kalish, 2006). They may also provide more fluent access to information, and help to link prior knowledge with new information (Booth, 2009). This, in turn, may facilitate the successful construction of meaning, a crucial step for word learning and oral language comprehension (Gernsbacher, 1990; Markman, 1989). In one of our extension tasks, for example, we identified the class (e.g. wild animals) of an unknown word (e.g. sloth) to preschoolers, and asked them to describe some the characteristics of the animal (e.g. that it is not tame; lives out doors away from people; needs food and water to survive). Children who had been taught new words in a taxonomic framework significantly outperformed those taught through business-as-usual techniques (Neuman, Newman, & Dwyer, 2011).

In addition to providing a meaningful conceptual framework, taxonomies may also support children's learning by scaffolding their ability to draw inferences about texts. Such inference-making is an especially critical skill for comprehension. During book-reading, for example, children must be able to go beyond the information directly provided in the text in order to fill in information necessary for comprehension. Invoking taxonomic category membership may help scaffold children's ability to make inferences about content that has only been provided implicitly by the text. This powerful ability to facilitate children's inference-making is referred to as the "inductive potential" of taxonomic categories (Rehder & Hastie, 2004). If children know that animals need food to survive, for example, they may be able to make inferences about category members at increasingly specific levels: birds are animals, therefore birds need food to survive; penguins are birds, birds are animals; therefore, penguins need food to survive and so forth.

Furthermore, when these words and concepts are embedded in sequences of topics that build on children's learning experiences, the end result can be a more solid body of content knowledge (Neuman et al., 2016). With a clearly-defined set of cross-cutting scientific concepts, children can begin to identify big ideas across the physical, life, earth and biological sciences that serve as the foundation for their progress in science education (Corcoran, Mosher, & Rogat, 2009).

4. Questioning strategies in read-alouds

Considerable evidence has shown that certain types of interactional routines and questioning techniques can further facilitate children's active engagement in learning from text (Hamre, Pianta, Hatfield, & Jamil, 2014). Intervention studies have demonstrated that talk focused on vocabulary-eliciting questions has beneficial effects on expressive vocabulary (National Early Literacy Panel, 2008; Whitehurst & Lonigan, 1998). In addition, discussion while books are being read, and before and after they are read can also lead to enhanced understanding of text (Gonzalez et al., 2014).

The cognitive demand of these conversation-eliciting and questioning strategies has been the topic of research in a number of studies of book reading (Dickinson & Porche, 2011; Neuman & Dwyer, 2011). For example, talk with lower cognitive demand might include the labeling of objects or action, skill routines which occurring the reading, and chiming of familiar lines or phrases (e.g. choral reading). On the other hand, talk with greater cognitive demand might include analyzing events or making comparisons across text, or evaluating the authenticity of particular events in the text (Carlisle, Kelcey, & Berebitsky, 2013). Although there is some research that suggests that all questions, whether they either high or low cognitive demand may contribute to children's learning, a recent study by Carlisle and her colleagues with older students found that the quality of discourse actions mattered. When teachers engaged in actions of greater cognitive demand (e.g. such as discussion), there was greater vocabulary growth among third graders.

However, none of these studies have accounted for the interplay between genre of the text and questioning techniques. For example, asking recall questions for a book designed to play with rhythmic words (e.g. Chicka Chicka Boom Boom) would be highly inappropriate; similarly, summarizing extensive chunks of text, a high cognitive demand task may be inappropriate with a narrative nonfiction book. In a study of teacher read-aloud practices in 25 classrooms, Dickinson and Smith (1994) found such differences, reporting that the type of classroom talk varied with different genre of text. Teachers were more likely to use cognitively challenging talk (e.g. analysis of characters; predictions of coming events) with narratives or storybooks. In contrast, they often book-focused utterances (e.g. skill routines) and chiming (e.g. choral responses) with alphabet books, and predictable text. Information books, on the other hand, seemed to elicit more immediate recall as well as vocabulary-related questions (Smolkin & Donovan, 2000). Therefore, the types of questioning strategies are related the genre of the text (Price et al., 2009). Recognizing the affordances and constraints of various genres of text and adjusting our questioning strategies appropriately might enable educators to use both text and talk to greater advantage.

4.1. Research questions and hypotheses

To summarize, accelerating instruction in word and world knowledge early on is critical if we are to improve low-income children's success in reading comprehension and performance. However, merely adopting the use of more complex informational texts is not likely to be sufficient to bring about improvement in vocabulary and content knowledge. Rather, given the extent of the differences in readiness skills at the starting gate, we need to adopt a more systematic approach, helping children to build knowledge through text in ways in which they can be most successful.

In this study, we examine an intervention designed to help children structure knowledge networks by immersing them in the vocabulary and content of science-related topics. Because these skills are best served by spending extended time on reading and listening to texts on the same topic, we focus on an approach in which multiple text genres are organized into coherent text sets, with accompanying questioning strategies that engage children in cognitively challenge talk appropriate to the text genre. Designed as a year-long supplemental program, the intervention involves children in science topics, using cross-cutting concepts and vocabulary within taxonomic categories to build knowledge networks to enhance young children's academic vocabulary, conceptual knowledge and content knowledge. In this study, we examine the potential benefits of the intervention on child outcomes, focusing on three essential research questions and hypotheses:

1. What is the impact of the intervention on science-related vocabulary? We hypothesize that treatment children will retain science-related vocabulary to a greater degree when words are taught within taxonomic categories than comparison children, and that the repeated exposure to these words in different text genre will allow children to engage in conversations that build knowledge around these specialized words.
2. To what extent does the intervention improve children's knowledge of science-related concepts? Related to our first question and hypothesis, we hypothesize that the deliberate grouping of words into taxonomic categories will enhance the development of science-related concepts. Studies (Waxman & Namy, 1997) suggest that young children use a variety of different types of category relationships to organize information depending on context and background knowledge. However, taxonomic categories are different than other category structures (e.g. such as themes) because they are based on shared properties and are hierarchical. With taxonomies, principles of class inclusion (and exclusion) can be applied between lower and higher level categories, making this form of categorization ideal for concept-building (Gelman & Kalish, 2006). Given that science is a well-structure domain, we hypothesize that children will make significant gains in developing science-related concepts.
3. Does the intervention improve more global receptive and expressive vocabulary? We hypothesize that learning words in taxonomic categories can improve children's gains in overall receptive and expressive language. We base our hypothesis on the research

demonstrating an association between conceptual organization and vocabulary development (Gopnik & Meltzoff, 1987). For example, computational modeling studies have demonstrated that the more words learned by a computer model, the more likely the concepts associated with those words will be organized into a taxonomic hierarchy (Borovsky & Elman, 2006). At the same time, previous studies (Kaefer & Neuman, 2011; Pinkham, Kaefer, & Neuman, 2014) suggest that the reverse may also be true. Taxonomies may improve word learning. Although the causality of this relationship has not been clearly established (e.g. whether conceptual knowledge influences the acquisition of novel words or the other way around, or some combination of the two) (Gopnik, 2012), we hypothesize that gains in overall receptive and expressive language may result from using taxonomies as a powerful scaffold for new word learning.

5. Method

5.1. Participants and sites

Children in prekindergarten, and kindergarten classrooms from 45 classrooms in 12 elementary schools from a large metropolitan area participated in the study. Schools ranged from 71% to 100% free and reduced lunch. In nine of the schools, classrooms at the designated grade levels (preK; K) were randomly assigned to the treatment or the control group; in three of the schools, however, there was only one preK classroom which was assigned to treatment with no additional class to assign as a control. The total sample, therefore, included 24 treatment classrooms and 21 control classrooms.

Teachers in both conditions were primarily of minority status (33% Latina, 23% African American, or 7% Multiracial); 34% were white and 3% were Asian. All were female with the exception of one male teacher, and all held Masters' degrees. Over 87% had six years or more of teaching experience. For 75% of the teachers, English was their primary language, whereas for 25%, Spanish was their primary language.

The families of all children in the enrolled classrooms received recruitment packets and informed consent agreements that requested permission to conduct ongoing assessments over the year. Although all children in the treatment classrooms participated in the program, six (3 boys; 3 girls) were randomly selected from each classroom to participate in the study. Table 1 describes the child-level characteristics by study condition.

Over two-thirds of the children were of minority status, largely Latino from different South American countries or African-American. About 15% of the children had identified disabilities. In both groups, there was a large emergent bilingual population. Children's scores on the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007) (described below) were approximately one standard deviation below the norm.

Table 1
Demographic Characteristics of Treatment and Control Students.

	Treatment (N = 148)	Control (N = 117)
Gender		
Male	38%	41%
Female	62%	59%
Ethnicity		
Hispanic or Latino	36%	41%
African American	21%	27%
West Indian or Caribbean	10%	8%
Multiracial	10%	4%
Asian	4%	5%
White	1%	2%
Native American or Alaska native	0%	1%
SES (% with free lunch)	78%	80%
SPED	14%	15%
ELL Status	13%	14%
PPVT	91.13	91.46

Note: SES: Socioeconomic Status; SPED: Special Education; ELL status: English Language Learners; PPVT: Peabody Picture Vocabulary Test.
No significant differences between groups.

5.2. Intervention

The instructional intervention consisted of an adaptation of the World of Words (WOW) (Neuman, 2016), a shared book reading designed to promote vocabulary and content knowledge in science. Structurally, the curriculum is organized by topics that represent animate taxonomies in life, earth, and physical sciences aligned with the Next Generation Science standards. Within the curriculum, words are selected that represent content-rich vocabulary within the category structure of the topic (e.g. plants; stem, leaf, roots). Recognizing that words are conveyors of content knowledge, this vocabulary and its use will be encountered in multiple contexts, building a foundation for learning the concepts in the domain.

There are ten topic-related text sets per grade using multiple genre. Five read-aloud books are clustered in each text set. Each text set begins with one or more predictable books, using the book structure as a mnemonic device to support and help children retain words they are learning. The next genre in the series is one or two narrative nonfiction books, and the last, an expository text, with the assumption that the higher density of vocabulary is now more familiar due to the repeated exposures in the other texts. Within each text set, lessons are organized to prime children's background knowledge and strategically integrate concepts with previously learned materials. Picture cards of several exemplars of the category (e.g. camel; scorpion, animals in the desert) are included for each topic to emphasize the important properties of the category, and to enhance the perceptual accessibility of the words for the English language learners (Hadley, Dickinson, Hirsh-Pasek, Golinkoff, & Nesbitt, 2015).

Each text set includes teacher lesson plans, highlighting the target words to be taught, the concepts to be developed and before, during, and after reading activities. Explicit teaching techniques are described throughout the manual with a rationale, and background information for enacting the program. At the end of the program at each grade level children will have learned 100 topical words; 30 challenge words (e.g. chrysalis; words that are designed to accelerate development and problem-solve using evidence) and 100 supporting words (e.g. predict; summarize–academic vocabulary that supports children's ability to talk about the topic).

WOW is enacted during the Morning Meeting, when children gather together to begin their daily activities. Teachers introduce words explicitly, read- and re-read the text, stopping only at key points to clarify the meaning of words or concepts. Following the reading, she/he asks challenge questions that focus on new words within and outside the category, helping children to build knowledge of the properties of certain concepts (e.g. insects have six legs, and three body parts). Challenge items are designed to encourage students to apply the concepts they acquire to think critically about what might or might not constitute category membership (Wellman & Gelman, 1998). Each text set is to be used over a two-week period, for a 20-week intervention.

The instructional sequence was designed to help teachers systematically scaffold students' learning of words and concepts. In the beginning, for example, the teacher's lesson plan focuses on explicit instruction, helping children 'get set' and give meaning to the topic. For example, the teacher might introduce the topic of marine mammals by using picture cards, explaining that while they live in the ocean, they have lungs and breathe air, and have backbones, and that one type is a dolphin. As the instructional sequences progresses, the teacher begins to build bridges to what student have already learned, engaging them in establishing inter-textual linkages with the other books in the text set. She might ask children compare/contrast questions, such as how is a dolphin like a starfish, and how might it be different. Here the teacher is encouraged to release control, stepping back to support more openended responses. Throughout the sequence children uses familiar words to talk about the topic, at the same time, they are encouraged to use content-specific words. Text sets at each grade level are designed to build on one another.

6. Procedures

The study began in the fall, 2014 and continued throughout the year, ending in May 2015. Before classes began, teachers in the treatment group (prek and kindergarten) participated in a day-long professional development training. In this workshop, we discussed the theoretical foundation of the supplemental curriculum, the pedagogical strategies used throughout the program (e.g. differences in genre features, and teacher interactions) and reviewed text sets. Teachers worked within grade and across grade level groups to examine the alignment of the materials with their standards, and the ways in which topics built on one another to establish big ideas in science. In addition, teachers in each school were assigned a coach who would be responsible for supporting the intervention throughout the year.

6.1. Student assessments

We administered a battery of standardized and author-created assessments prior to and following the intervention. Our purpose was to understand how the curriculum might influence science-related word knowledge and content knowledge, as well as more global aspects of receptive and expressive language.

Standardized measures. Standardized assessments of receptive and expressive vocabulary were individually administered to the randomly selected students in each grade.

General receptive vocabulary. Children's general receptive vocabulary knowledge was measured with the Peabody Picture Vocabulary Test-IV, a receptive vocabulary test which yields both raw scores and standard equivalent scores related to national norms (Dunn & Dunn, 2007). On the PPVT, the assessor provides a verbal prompt and the child is requested to point to one of four pictures on a panel that represents the object or the action. Internal consistency ranges from 0.86 to 0.98 for both Forms A and B.

General expressive vocabulary. Children's expressive vocabulary was measured with the Expressive One-word Picture Vocabulary Test-IV (EOWPVT) (Martin & Brownell, 2000). The measure assesses the verbal expressive vocabulary of children ages 2 through 80+ by asking the child to name (in English) objects, actions and concepts pictured in illustrations. Internal consistency ranges from 0.93 to 0.97 with a media of 0.95 across age groups.

Author-created assessments. Our author-created assessments were drawn from the instructional materials and were designed to measure science-related vocabulary and content knowledge.

Science-related Vocabulary. We constructed a 50-item expressive task to measure children's vocabulary from text sets for each grade level. Specifically, it was designed to measure whether children learned the words that were taught in the intervention. Five words were identified across each of the ten text sets per grade level that represented content-specific words important to the topic (e.g. Tier 2/3, according to Beck, McKeown, and Kucan (2002). For example, words selected from a preK topic included *whale, eel, dolphin, coral, manatee* and *shark*, highlighting the category of marine animals. These words were randomly placed in a set order. The child was presented with a picture prompt and asked to label the word. Responses were scored either correct or incorrect for a total score. Internal consistency ranged from 0.85 to 0.90 for each assessment.

Knowledge of science-related concepts. Children's content knowledge was assessed through two tasks. Each task examined children's understanding of key content related to the topics taught at each of the grade levels in the WOW curriculum. For both tasks, target words were placed in new contexts that were not covered in the instruction. Distractors, as well, were not mentioned throughout the instructional program. These measures, therefore, were designed to provide some indication of children's ability to transfer these concepts to a new situation.

Yes/No. We designed a 40-item task to assess children's conceptual knowledge of the target words for each topic. Four conceptual properties from each topic per grade level were selected. Assessment questions were devised to include the target word in a sentence that was related to the science concept (e.g. human body; is our heart attached to our body?) or not related to the concept (e.g. are eye glasses attached to our body?). These questions were designed to examine words in a new context, a form of transfer for very young children. Each conceptual property was tested using both in-category and out-of-category items to measure their ability to apply the word to a new context. Children heard an equal number of yes and no questions across the assessment, and the order of these questions was fully randomized. Students responded either 'yes' or 'no' to each question and a total number of correct responses were recorded. Reliability was 0.81, and 0.79 across grade levels respectively.

Categorical property knowledge. To examine children's content knowledge in greater depth, we constructed a 40-item receptive task to identify categories and properties related to the core ideas and concepts in topics. In this task, children were shown three pictures: a target picture (e.g. manatee); thematically-related to the target (e.g. coral) and an out-of-category, but plausible distractor (e.g. sea horse). Children were then asked to identify which item/object belonged to a particular category (e.g., Which is a mammal?) or to identify the item/object that possessed a particular category attribute. A total of two category level questions were assessed (one for each topic) (20), and a total of 20 concept property questions (two for each topic). Concept property questions were selected as most representative of the category. For example, children were assessed on the property of "backbone" as it is a critical and defining property of the category, "mammals." Responses were tallied for accuracy on category and property questions, and for the overall assessment (Total score possible = 40). Assessments varied by topics, and were developed for each grade level. Alpha coefficients based on our sample ranged from 0.74 to 0.85 for pre- and posttests respectively.

Assessments were conducted by 10 trained research assistants, each of which had participated in a day-long training. The project manager reviewed all assessments throughout the assessment period for quality and consistency. Assessors and scorers were blind to the treatment.

6.2. The coaching model

The World of Words was supported by a coach who visited classrooms weekly throughout the year to observe, occasionally model lessons if needed, and informally monitor children's progress and provide feedback to teachers. Coaches were graduate students in education, with a strong teaching background in early education. They were trained in a one-day workshop on the intervention. Following the workshop, the project manager observed them teaching a lesson to children to ensure fidelity to the intervention model.

Our coaching model was designed to support the integrity of the intervention and the instructional design (Bryk, 2016). Coaches scheduled their visits of approximately 20 mins during the WOW lesson. During the first two weeks of the text set topic, she observed and provided brief written comments following the lesson. On the second week, she randomly selected a group of children and informally monitored their progress, using picture cards to ask children to 'What is it? Can you tell me something about it?' These responses were summarized and given to the teacher for feedback. They were also used in monthly debriefing meetings among the research team to determine if additional supports, such as attention to pacing (e.g. keeping the lesson within the recommended time limits) were needed. Based on these debriefings, coaches would provide additional feedback to teachers in preparation

Table 2
Means (in proportions) and Standard Deviations of Children's Outcomes by Condition, Grade, and ELL Status.

	Pre-Kindergarten				Kindergarten			
	Treatment		Control		Treatment		Control	
	ELL	English	ELL	English	ELL	English	ELL	English
Science vocabulary								
Pretest	0.26 (0.12)	0.26 (0.11)	**0.19 (0.14)**	0.26 (0.10)	0.21 (0.09)	0.23 (0.12)	0.23 (0.11)	0.23 (0.09)
Posttest	0.45 (0.12)	0.40 (0.10)	0.23 (0.11)	0.34 (0.07)	0.33 (0.10)	0.31 (0.13)	0.26 (0.09)	0.28 (0.10)
Knowledge of science								
Concepts								
Pretest	0.59 (0.12)	0.60 (0.11)	0.52 (0.19)	0.55 (0.14)	0.60 (0.10)	0.57 (0.15)	0.59 (0.11)	0.61 (0.11)
Posttest	0.76 (0.12)	0.66 (0.13)	0.65 (0.10)	0.64 (0.13)	0.71 (0.11)	0.66 (0.15)	0.63 (0.11)	0.63 (0.11)
PPVT[a]								
Pretest	95.95 (15.55)	90.85 (14.18)	91.76 (12.39)	90.94 (13.16)	91.37 (12.24)	93.29 (13.46)	90.92 (12.68)	94.38 (12.65)
Posttest	101.10 (13.18)	95.00 (12.48)	92.00 (16.32)	93.89 (10.30)	93.05 (10.78)	93.21 (15.16)	92.95 (11.67)	94.56 (13.88)
EOWPVT[b]								
Pretest	89.24 (8.64)	91.12 (11.58)	81.06 (12.97)	88.61 (11.49)	84.79 (10.65)	86.13 (10.01)	84.19 (9.37)	87.31 (7.30)
Posttest	99.90 (14.28)	97.24 (11.56)	79.76 (15.55)	90.67 (11.63)	90.34 (1363)	90.83 (12.65)	87.43 (11.75)	94.04 (13.57)

Note:
[a] PPVT = Peabody Picture Vocabulary Test.
[b] EOWPVT = Expressive One-Word Picture Vocabulary Test.

for the next text set.

6.3. Fidelity to implementation

Three times throughout the course of the intervention period, research assistants conducted fidelity checks. The checklist included items related to the content of the lesson (e.g. whether all aspects of the lesson were covered such as introduction of words; linkage to concepts; using picture cards; reading book; post-reading discussion), and the quality of the enactment (e.g. lesson pacing, engagement, interaction; and responsive teaching). Teachers received 1 point for each component enacted, and conversely, 0 points if the component was not enacted. Points were tallied, then averaged across observations for each teacher, and a percentage was derived to indicate the degree of fidelity for each teacher. Using this procedure, total fidelity ranged from 89% to 100%.

Control group. Throughout the intervention, we conducted three observations of the control group. Classrooms in the wait-list control group continued with their business-as-usual morning meeting. Typically, this included a storybook reading from their classroom library for children in the preK and kindergarten, and a more structured whole group shared book reading lesson from their curriculum program for the first graders. No additional materials or coaching supports were provided to teachers throughout the year.

The final sample included 265 students (Treatment = 148; Control = 117), representing an attrition of 10% who either moved or were absent over the multiple days of testing.

7. Results

We present our results in three parts to answer our research questions. First, we examine the impact of the intervention on science-related vocabulary. Given the sizable number of students who were English language learners, we report descriptives for native English speakers and ELL students. Means and standard deviations are reported in percentages across each grade level. There were no differences by gender for any of outcomes variables (all ps > 0.01). We then turn to an analysis of content knowledge and the effects of the intervention on children's knowledge of science-related concepts. Finally, we report on children's growth on the standardized receptive and expressive language measures.

Due to the nested structure in our data (i.e. 265 children; 45 Classrooms; 12 school sites), we chose multilevel modeling to analyze the data. All measures had significant intraclass correlation coefficients (ICC), ranging from 0.19 to 0.50. Multilevel modeling is preferred over traditional fixed effects models for nested data in which individuals within a group are not independent. Given that children were nested within classrooms and sites in 12 elementary schools, we employed three-level models to examine differences for each child outcome.

We used the following standard equation:Level 1 model:

$$\mathbf{Y_{ikj}} = \boldsymbol{\pi_{0kj}} + \boldsymbol{\pi_{1kjn}}(\textbf{pre}-\textbf{test score})_{\mathbf{ikj}} + \boldsymbol{\pi_{2kj}}(\textbf{ethnicity})_{\mathbf{ikj}} + \boldsymbol{\pi_{3kj}}(\textbf{ELL status})_{\mathbf{ikj}} + \varepsilon ikj$$

Level 2 model:

$$\pi \mathbf{0}kj = \beta \mathbf{00}j + \beta \mathbf{01}jtreatment\ groupj + \beta \mathbf{02}jgradej + \beta \mathbf{03}treatment\ group \times gradej + \beta \mathbf{04}jELL \times gradej + \beta \mathbf{05}jtreatment\ group \times ELLj + r\mathbf{0}kj$$

Level 3 model:

$$\beta \mathbf{00}j = \gamma \mathbf{000} + \mu \mathbf{00}j$$

Models for each of our outcomes were based on the same overall formula structure. The Level 1 equation examined individual level variables. In this case, the outcome variable (vocabulary; knowledge of scientific concepts, etc) was predicted by children's pretest scores, ELL status, and ethnicity. π0kj is a random intercept, π1kj − π3kj are fixed slope parameters, and εtkj is error not otherwise accounted for by the predictive variables. The Level 2 equation examines group level variables. In this case, the intercept in the Level 1 model is predicted by an intercept, γ0j, and the effect of being assigned to the treatment group, grade, and the interactions between treatment group, grade, and ELL status. r0kj is the residual error. In the Level 3 equation, no predictors are added, but an intercept, γ000, and associated error, are included.

Table 3
Model Specifications (estimates with standard error in parentheses) of Multilevel Analyses for Models with Independent Variables Entered Together, Controlling for Pre-test scores and Ethnicity.

	Treatment	ELL	Grade	Treatment X Grade	ELL X Grade	ELL X Treatment
Science Voc.	.09** (0.02)	.03 (0.02)	.08** (0.02)	.06 (0.03)	.003 (0.02)	.06* (0.02)
Know. science concepts	.09* (0.03)	.04 (0.03)	.03 (0.03)	.01 (0.04)	.02 (0.03)	.05 (0.03)
PPVT[a]	.17 (2.73)	1.07 (2.09)	3.80 (2.97)	6.00 (3.94)	1.23 (2.70)	1.47 (2.56)
EOWPVT[b]	3.19* (2.62)	.60 (2.26)	6.09 (2.89)	8.85* (3.64)	2.79 (2.87)	6.07* (2.73)

Note:
[a] PPVT: Peabody Picture Vocabulary Test.
[b] Expressive One-word Picture Vocabulary Test.
* $p < .05$.
** $p < .01$.

7.1. Effects of the intervention on science-related vocabulary

Table 2 presents the pre- and posttest means and standard deviations for the vocabulary assessment for treatment and control groups. As shown in the Table, pretest scores were low, indicating on average that for preschoolers and kindergartners, less than one-quarter of the words were known. Given that these words were related to science content representing Tier 3 words, it is not surprising that they were unfamiliar to students. There were no significant differences between treatment and control groups prior to instruction at either of the grade levels.

Posttest scores, however, revealed substantial growth for both native English and ELL students in the treatment group compared to the control. Children in preK and kindergarten appeared to benefit from the intervention. To analyze these differences, we conducted a three-level mixed-effect model (MLM) that included classroom and site as random intercepts, along with pretest scores as a fixed covariate in the model (see Table 3). Treatment group, grade level, ELL status, ethnicity, ELL by grade interaction, ELL by treatment interaction, and treatment by grade level interaction served as fixed independent variables. Because each of our independent variables were fixed, with no random slopes, we report the fixed effect outcomes. Significance tests for fixed effects in a multi-level model operate similarly to a conventional general linear model (Peugh & Enders, 2005), and are interpreted in the same way as a conventionally structured test (Seltman, 2015).

Our analysis revealed a significant effect of pretest scores, $t(251.22) = 14.08$ $p < .001$, as well as a significant effect of treatment, $t(60.50) = 4.39$, $p < .001$, grade, $t(71.23) = 3.46$, $p = .001$, treatment by ELL interaction, $t(248.86) = 2.91$, $p = .004$. There was no significant main effect of ELL status, $t(250.69) = 1.76$, $p = .080$, ethnicity, $t(244.56) = 1.12$, $p = .265$, or ELL by grade interaction, $t(243.73) = 0.13$, $p = .900$, and no significant effect of treatment by grade interaction, $t(35.83) = 1.97$, $p = .057$. Given its potential to inform educational practice, we conducted additional analyses despite its non-significance to examine the treatment's effects on grade levels. Follow-up tests on the grade by treatment interaction indicated that the intervention made a significant impact on children's scores at both grade levels. However, the effect sizes were higher for preK children $t(89) = 6.75$, $p < .001$, $d = 1.14$ than for kindergarten, $t(41.30) = 2.96$, $p = .005$, $d = 0.48$. These results suggest that the intervention had an effect at both grade levels, however, it had the greatest impact with the youngest learners.

Follow-up tests on the ELL by treatment interaction showed that ELL students in the treatment condition showed more growth than native English speakers, $t(144) = 2.26$, $p = .025$. At the same time, native English speaking students showed more growth in the control condition, $t(114) = 2.34$, $p = .021$. Taken together, these results suggest that the intervention may have been particularly beneficial for students who are learning English as a second language, disrupting the business-as-usual condition.

7.2. Knowledge of science-related concepts

Table 2 presents the pre- and posttest means of children's knowledge of science-related concepts assessment and their developing knowledge of the information genre in science. Once again, for science-related concepts, we see a pattern consistent with our earlier results: greater gains were likely to be made for children in the earlier years than for students in kindergarten.

Our multi-level analysis for science-related concepts shown in Table 3 also reported a significant effect of pretest scores $t(249.00) = 3.175$, $p = .002$ on posttest scores, as well as a significant effect of treatment, $t(64.20) = 2.88$, $p = .005$. However, there was no significant effect of ethnicity, $t(250.28) = 0.98$, $p = .326$, ELL status, $t(251.35) = 1.72$, $p = .088$, treatment by ELL interaction $t(247.42) = 1.48$, $p = .140$, grade $t(74.83) = 0.96$, $p = .342$, ELL by grade interaction, $t(243.16) = 0.51$, $p = .609$, or treatment by grade interaction $t(38.43) = 0.22$, $p = .826$. Follow up tests show that English language learners in both treatment and control conditions improved significantly more than English speakers in preK, $t(94) = 2.18$, $p = .032$, $d = 0.40$ suggesting growth in science-related concepts for younger non-native speaking children. There were, however, no differences between groups in kindergarten, $t(164) = 1.43$, $p = .155$ $d = 0.24$ suggesting that growth slowed down after preK.

Taken together, these results suggest that the intervention significantly enhanced children's developing knowledge of science-related concepts. Of particular notice, ELL children were as successful in developing these concepts as native English speakers, with growth most noted in preK than kindergarten.

7.3. Global receptive and expressive language

In our final analysis, we examined the effects of the intervention on more global measures of language development. Table 2 also presents the pre- and posttests for receptive and expressive vocabulary. Pretest standardized scores for the PPVT were slightly higher than the EOWPVT. Nevertheless, both pretests indicated that students were below average for their norm group. There were no significant differences between treatment and control group. Posttests of the PPVT, however, revealed a pattern noted in previous research: there was a striking stability in scores rather than change. This was true for native English speakers and ELL students as well.

Conducting an multi-level analysis for receptive vocabulary shown in Table 3, we found a significant effect of pre-test scores $t(247.53) = 12.87$, $p < .001$, but no significant effect of treatment, $t(55.43) = 0.06$, $p = .951$, or grade level, $t(64.35) = 1.27$, $p = .511$. Independent factors including ethnicity, $t(246.22) = 1.21$, $p = .225$, ELL status, $t(247.82) = 0.51$, $p = .610$, ELL by grade interaction, $t(252.50) = 0.457$, $p = .648$, ELL by treatment interaction, $t(252.94) = 0.58$, $p = .566$ or treatment by grade interaction $t(35.56) = 1.52$, $p = .137$ were all statistically insignificant. These results suggest that the intervention did not benefit growth in receptive language.

In contrast, expressive language scores rose for both groups over the year with the exception of the ELL preK students in the control group, which declined at posttest. For expressive vocabulary (EOWPVT) we found a significant effect of pre-test scores, $t(252.93) = 11.50$, $p < .001$, as well as grade level $t(70.60) = 2.11$, $p = .039$, a significant treatment by grade interaction, $t(34.88) = 2.43$, $p = .020$, and a significant treatment by ELL interaction, $t(248.10) = 2.23$, $p = .027$. There were, however, no significant effects of ethnicity, $t(251.25) = 0.99$, $p = .323$, main effects of treatment, $t(59.72) = 1.22$, $p = .229$, Ell, $t(252.59) = 0.27$, $p = .791$, or ELL by grade interaction, $t(241.59) = 0.97$, $p = .331$.

Follow-up tests on the significant treatment by grade interaction indicated a significant effect of treatment for preK students, $t(13.06) = 2.77$, $p = .016$, $d = 0.94$ but not kindergarteners, $t(25.36) = 0.051$, $p = .959$, $d = 0.03$. These results suggest that the intervention significantly enhanced prek children's expressive language skills, closely approximating average proficiency for their norm group, compared to the control group who remained one or two standard deviations below norm. However, the intervention did not have a similar impact for kindergartners in the sample who remained statistically on par with the control group.

Follow up tests on the significant treatment by ELL interaction reported a significant effect of treatment for ELL students, $t(26.55) = 2.43$, $p = .022$, $d = 0.63$, but not native English speakers, $t(24.08) = 0.08$, $p = .932$, $d = 0.06$. This is consistent with our previous findings, indicating that the intervention helped develop language skills for ELL students.

In summary, our results indicated that the preK and kindergarten treatment groups scored significantly higher on science-related vocabulary and knowledge of science concepts than the controls. In addition, it showed a treatment by ELL interaction which, as indicated in the follow-up tests, revealed that ELLs in the treatment group responded to the intervention more than native English speakers. These results were in contrast to the control condition in which the non-ELL performed at higher levels.

8. Discussion

Children need word and world knowledge to successfully learn to read and comprehend complex texts in later grades (Stanovich, West, & Harrison, 1995). However, word meanings do not exist in isolation; rather, they are learned by encountering them frequently in multiple and varied contexts. As children acquire deeper word knowledge through these contexts, they begin to build networks linking words and other related terms to concepts. Comprehension is dependent on the ability to retrieve these networks of concepts. Authors (2006) have argued that once children begin to gain conceptually rich knowledge, they become able to acquire more knowledge at faster rates.

Therefore, the purpose of this project was to integrate word knowledge within the context of read-aloud books to develop and progressively build knowledge networks. Because young children need a fairly extensive network of words and concepts, we designed a read-aloud program in which words were explicitly taught in the context of science concepts and organized around high priority topics, varied text genres, and explicit interactive discussions of content-rich words and big ideas. Our goal was to accelerate children's learning by understanding the relationship between new words and their connected concepts. In this respect, the intervention was designed to make deliberate the connections between words to concepts to knowledge about science.

The study was conducted as a quasi-experimental field experiment in a large urban district selected on the basis of high poverty criteria, and not on the language status or other characteristics of the students. As such, our sample represents what might often be reported in urban districts: schools included almost exclusively children of minority status representing many cultures and countries, with a sizeable number of children of nonnative English speakers as well as a substantial percentage of special education students in these inclusive classrooms. Therefore, the intervention took place in a context that clearly mirrors many of the opportunities and challenges reported in urban school districts today (National Assessment of Educational Progress., 2013).

Within this diverse setting, the results of our study found that preK and kindergarten children learned science-related vocabulary and science concepts, the building blocks of understanding important big ideas in the domain. Perhaps most striking was the growth for ELLs, who in many cases surpassed the gains of native English speakers. These results indicate that a knowledge-rich intervention that prioritized explicit understanding of words and concepts, along with many opportunities to engage in discussion significantly improved pre-K and kindergarten children's word and world knowledge about science. In this respect, it both replicates and extends the research by Gonzalez et al. (2011) who found that a content focused shared book reading program significantly enhanced preschoolers' science vocabulary. In addition to vocabulary, it highlights that such an integrated program can improve both their vocabulary *and* content knowledge.

We can only speculate on the individual program components that promoted vocabulary and content knowledge in science. However, there is evidence from previous studies that topic immersion in which teachers read a sequence of books organized by topic instead of 'stand-aloud' texts and materials can support vocabulary development (Pollard-Durodola et al., 2012), providing opportunities for multiple exposures to words and concepts shown in different contexts and with different genres. Picture supports, in addition, have shown to be especially helpful to ELL students, giving them visual representations of key vocabulary words (Mayer, 2001).

Nevertheless, what distinguished this program from many other read-aloud interventions was that words were taught within taxonomic categories, an instructional design feature that supported the connections between words to concepts from the very outset of the program. Previous research has shown that categorization is an integral part of conceptual knowledge, and may allow children to organize and process incoming information more efficiently by structuring word knowledge hierarchically (Gelman & Kalish, 2006). Treatment children in our study were able to use the inductive potential of categories to develop inferences about these words in new contexts. Through the interplay between words and concepts, children were able to deepen their understanding of the word meaning.

In fact, there is evidence of a reciprocal relationship between words and conceptual development (Booth, 2009). In a study with 36 3-year-olds, Booth found that novel words applied to referents with known conceptual properties were more readily recalled several days later than those words applied to referents for which conceptual properties were unspecified. In addition, in a study of 104 4-year-olds from Head Start, we found that children who learned words in taxonomic categories were better able to sort taxonomically than children in the control group, and that this ability subsequently predicted the number of novel words learned outside of the training program. Although much more research is needed, it does suggest a bi-directional relationship between conceptual organization and word learning. If categorization can enhance young children's abilities to retain words and their conceptual properties and enable children to infer meaning in novel words, it might act as a bootstrap for self-learning in the future.

At the same time, our results on the global receptive and expressive vocabulary gains contrast with Gonzalez et al. (2011). Their findings showed significant gains in receptive but not expressive vocabulary for preschoolers. In contrast, our results showed virtually no substantial growth in receptive vocabulary for both preK and kindergartners, but impressive gains in expressive language for the ELL students particularly in the treatment group in pre-K. Although we can only speculate at this point, such gains might reflect the emphasis on discussion throughout the intervention. Each topic, for example, asked children to reflect on the similarities and differences in the properties related to categories, and to justify these decisions with evidence. Therefore, the role of discussion might be especially important in promoting vocabulary for young children as well as older students (Dwyer, Kelcey, Berebitsky, & Carlisle, 2016; Gonzalez et al., 2014). As a case in point, Lawrence and his colleagues in a large randomized trial (Lawrence, Crosson, Pare-Blagoev, & Snow, 2015) reported that a word learning program organized around engaging and discussable dilemmas produced small, but significant gains on taught vocabulary, and that the quality of classroom discussion mediated 14% of the treatment effects on vocabulary outcomes.

In addition, the selection of topics centered around key cross-cutting themes in science that intrigue young children may have supported discussions (Gelman & Brenneman, 2004). Rather than ask students to activate their background knowledge (Pearson, 2015), the intervention used a stair-step approach to introduce them to words and concepts, providing a common foundation of knowledge from which to have discussions. In this respect, it provides further evidence of the integral ties between background knowledge and comprehension. It offers additional support for the *knowledge hypothesis*, the view that words are part of larger knowledge structures and that it is these knowledge structures, not the words per se that impact children's comprehension (Anderson & Freebody, 1981). Unfortunately, previous research has often failed to examine how background knowledge (or the lack thereof) might contribute to active engagement in discussions and subsequent content learning.

There are a number of limitations in this research. For one, our assessments were not particularly sensitive to second language learners. All were in English, and as a result, might have underestimated children's vocabulary and content knowledge. Second, although we attempted through our randomization efforts to randomly assign classrooms to treatment and control conditions, there was not a sufficient number of preK classrooms in three schools. Rather than eliminate these three preK classrooms, we chose to place them in the treatment group. Third, we did not account for teacher background knowledge in science. Based on our observations, it was clear that some teachers had a greater grasp themselves of science concepts than others. Educative curriculum supports for teachers, highlighting science definitions and concepts to aide in the enactment of the curriculum could potentially enhance teachers' comfort with science content (Davis et al., 2014). Previous research has shown that these educative elements can promote higher level questioning strategies throughout discussion in preK.

In summary, this study demonstrates the important interconnections between word and world knowledge. Because language serves as the primary medium through which early content learning occurs (Boals, Kenyon, Blair, Wilmes, & Wright, 2015), instruction that targets content-specific vocabulary exposes children to ideas associated with a given word. In this study, we provide evidence that these behaviors can occur simultaneously. As early as pre-K, children can profit from systematic and focused instruction that begins the process of building knowledge networks essential for further content learning and comprehension development. In this respect, we hope to disrupt the all-too-frequent pattern of offering systematically weaker content and curriculum to low-income children (Schmidt, Burroughs, Zoido, & Houang, 2015). Rather, we argue that these low-income children can and do benefit from interventions that are both challenging and achievable.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.cedpsych.2017.12.001.

References

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.). *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.

August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the national literacy panel on language-minority children and youth.* Mahwah, NJ: Erlbaum.

Beck, I., McKeown, M., & Kucan, L. (2002). *Bringing words to life.* New York: Guilford.

Boals, T., Kenyon, D., Blair, A., Wilmes, C., & Wright, L. (2015). Transformation in K-12 English language proficiency assessment: Changing contexts, changing constructs. *Review of Research in Education, 39*, 122–164. http://dx.doi.org/10.3102/0091732x14556072.

Booth, A. (2009). Causal supports for early word learning. *Child Development, 80*(4), 1243–1250. http://dx.doi.org/10.1111/j.1467-8624.2009.01328.

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: a relation between cognition and early word learning. *Journal of Child Language, 33*, 759–790. http://dx.doi.org/10.1017/s03050009006007574.

Bryk, A. (2016). *Fidelity of implementation: Is it the right concept?* Stanford, CA: Carnegie Foundation for the Advancement of Teaching.

Carlisle, J., Kelcey, B., & Berebitsky, D. (2013). Teachers' support of students' vocabulary learning during literacy instruction in high poverty elementary schools. *American Educational Research Journal, 50*(6), 1360–1391. http://dx.doi.org/10.3102/0002831213492844.

Cervetti, G., Barber, J., Dorph, R., Pearson, P. D., & Goldschmidt, P. (2012). The impact of an integrated approach to science and literacy in elementary school classrooms. *Journal of Research on Science Teaching, 49*, 631–658. http://dx.doi.org/10.1002/tea.21015.

Cervetti, G., Wright, T., & Hwang, H. (2016). Conceptual coherence, comprehension, and vocabulary acquisition: A knowledge effect? *Reading and Writing*, (Published online 28 January 2016).

Clements, D. H., & Sarama, J. (2008). Experimental evaluation of the effects of a research-based preschool mathematics curriculum. *American Educational Research Journal, 45*, 443–494. http://dx.doi.org/10.3102/0002831207312908.

Corcoran, T., Mosher, F., & Rogat, A. (2009). *Learning progressions in science.* Retrieved fromNew York: Consortium for Policy Research in Education.

Coyne, M., Little, M., Rawlinson, D., Simmons, D., Kwok, O., Kim, M., ... Civetelli, C. (2013). Replicating the impact of a supplemental beginning reading intervention: The role of instructional context. *Journal of Research on Educational Effectiveness, 6*, 1–23. http://dx.doi.org/10.1080/19345747.2012.706694.

Davis, E., Palinscar, A., Arias, A., Bismack, A., Marulis, L., & Iwashyna, S. (2014). Designing educative curriculum materials: A theoretically and empirically driven process. *Harvard Educational Review, 84*(1), 24–52.

De Vries, M. (2010). *Fraser bear: A cub's life.* Vancouver, BC, Canada: Graystone Books.

Dickinson, D., McCabe, A., & Essex, M. (2006). A window of opportunity we must open to all: The case for preschool with high quality support for language and literacy. In D. Dickinson, & S. B. Neuman (Vol. Eds.), *Handbook of early literacy research: Vol. II*, (pp. 11–28). New York: Guilford Press.

Dickinson, D., & Porche, M. (2011). Relation between language experiences in preschool classrooms and children's kindergarten and fourth-grade language and reading abilities. *Child Development, 82*(3), 870–886. http://dx.doi.org/10.1111/j.1467-8624.2011.01576.x.

Dickinson, D., & Smith, M. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29*, 104–122. http://dx.doi.org/10.2307/747807.

Duke, N. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*(2), 202–224. http://dx.doi.org/10.1598/rrq.35.2.1.

Dunn, L., & Dunn, D. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington, MN: Pearson Education Inc.

Dwyer, J., Kelcey, B., Berebitsky, D., & Carlisle, J. (2016). A study of teacher discourse moves that support text-based discussions. *Elementary School Journal, 117*(2), 285–309.

Gelman, R., & Brenneman (2004). Science learning pathways for young children. *Early Childhood Research Quarterly, 19*, 150–158. http://dx.doi.org/10.1016/j.ecresq.2004.01.009.

Gelman, S., & Kalish, C. (2006). Conceptual development. In W. Damon, R. Lerner, D. Kuhn, & R. Siegler (Eds.). *Handbook of child psychology* (pp. 687–733). New York: Wiley.

Gernsbacher, M. (1990). *Language comprehension as structure building.* Hillsdale, NJ: Erlbaum.

Gonzalez, J., Pollard-Durodola, S., Simmons, D., Taylor, A., Davis, M., Fogarty, M., & Simmons, L. (2014). Enhancing preschool children's vocabulary: Effects of teacher talk before, during and after shared reading. *Early Childhood Research Quarterly, 29*, 214–226. http://dx.doi.org/10.1016/j.ecresq.2013.11.001.

Gonzalez, J., Pollard-Durodola, S., Simmons, D., Taylor, A., Davis, M., Kim, M., & Simmons, L. (2011). Developing low-income preschoolers' social studies and science vocabulary knowledge through content-focused shared book reading. *Journal of Research on Educational Effectiveness, 4*, 25–52. http://dx.doi.org/10.1080/19345747.2010.487927.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. *Science, 337*, 1623. http://dx.doi.org/10.1126/science.1223416.

Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58*, 1523–1531. http://dx.doi.org/10.2307/1130692.

Hadley, E., Dickinson, D., Hirsh-Pasek, K., Golinkoff, R., & Nesbitt, K. (2015). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly, 51*(2), 181–198. http://dx.doi.org/10.1002/rrq.130.

Hamre, B., Pianta, R., Hatfield, B., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher-child interactions: Associations with preschool children's development. *Child Development, 85*(3), 1257–1274. http://dx.doi.org/10.1111/cdev.12184.

Hart, B., & Risley, T. (1995). *Meaningful differences.* Baltimore, MD: Brookes.

Hirsch, E. D. (2003). Reading comprehension requires knowledge of words and the world. *American Educator, 27* 10, 12-1316-1322,1328-1329, 1348.

Kaefer, T., & Neuman, S. B. (2011). A bi-directional relationship between conceptual organization and word learning. In *Paper presented at the Literacy Research Association, Jacksonville, FL*.

Lawrence, J., Crosson, A., Pare-Blagoev, E. J., & Snow, C. (2015). Word Generation randomized trial: Discussion mediates the impact of program treatment on academic word learning. *American Educational Research Journal, 52*(4), 750–786. http://dx.doi.org/10.3102/000283125579485.

Lee, V., & Burkam, D. (2002). *Inequality at the starting gate.* Washington, D.C.: Economic Policy Institute.

Llosa, L., Lee, O., Jiang, F., Haas, A., O'Connor, C., Van Booven, C., & Kieffer, M. (2016). Impact of a large-scale science intervention focused on English Language learners. *American Educational Research Journal, 53*(2), 395–424. http://dx.doi.org/10.3102/0002831216637348.

Markman, E. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.

Martin, N., & Brownell, R. (2000). *Expressive one-word picture vocabulary test* (4th ed.). Austin TX: Pro-Ed.

Mayer, R. E. (2001). *Multimedia learning.* New York, NY: Cambridge University Press.

National Assessment of Educational Progress. (2013). *2013 Math and reading trial urban district assessment*. Retrieved from Washington, D.C.

National Early Literacy Panel (2008). *Developing early literacy.* Washington, D.C.: National Institute for Literacy.

Neuman, S. B. (2016). *The World of Words (WOW)-revised.* NY: New York University.

Neuman, S. B., & Celano, D. (2012). *Giving our children a fighting chance: Poverty, literacy, and the development of information capital.* New York: Teachers College Press.

Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literacy Research, 43*, 103–129. http://dx.doi.org/10.1177/1086296x11403089.

Neuman, S. B., Newman, E., & 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. http://dx.doi.org/10.1598/RRQ.46.3.3.

Neuman, S. B., Pinkham, A., & Kaefer, T. (2016). Improving low-income preschoolers' word and world knowledge: The effects of content-rich instruction. *Elementary School Journal, 116*, 652–674. http://dx.doi.org/10.1086/686463.

Ouellette, G. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554–566. http://dx.doi.org/10.1037/0022-0663.98.3.554.

Pappas, C. (1991). Young children's strategies in learning the "book language" of information books. *Discourse Processes, 14*, 203–225. http://dx.doi.org/10.1080/01638539109544782.

Pearson (2015). *Readygen.* New York: Pearson.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. http://dx.doi.org/10.1080/10888430701530730.

Peugh, J. L., & Enders, C. (2005). Using the SPSS mixed procedure to fit cross-sectional and longitudinal multilevel models. *Educational and Psychological Measurement, 65*, 717–741. http://dx.doi.org/10.1177/0013164405278558.

Pinkham, A., Kaefer, T., & Neuman, S. B. (Eds.). (2012). *Knowledge development in early childhood.* New York: Guilford.

Pinkham, A., Kaefer, T., & Neuman, S. B. (2014). Taxonomies support preschoolers' knowledge acquisition from stories. *Child Development Research, 1–10*. http://dx.doi.org/10.1155/2014/386762.

Pollard-Durodola, S., Gonzalez, J., Simmons, D., Davis, M., Simmons, D., & Nava-Walichowski, M. (2012). Using knowledge networks to develop preschoolers' content vocabulary. *The Reading Teacher, 65*(4), 265–274. http://dx.doi.org/10.1002/trtr.01035.

Price, L., Bradley, B., & Smith, M. (2012). A comparison of preschool teachers' talk during storybook and information book read-alouds. *Early Childhood Research Quarterly, 27*(3), 426–440. http://dx.doi.org/10.1016/j.ecresq.2012.02.003.

Price, L., van Kleeck, A., & Huberty, C. (2009). Talk during book sharing between parents and preschool children: A comparison between storybook and expository book conditions. *Reading Research Quarterly, 44*(2), 171–194. http://dx.doi.org/10.1598/rrq.44.2.4.

Rehder, B., & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition, 91*, 113–153. http://dx.doi.org/10.1016/S0010-0277(03)00167-7.

Schmidt, W., Burroughs, N., Zoido, P., & Houang, R. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher, 44*(7), 371–386. http://dx.doi.org/10.3102/0012189x15603982.

Seltman, H. J. (2015). *Experimental design and analysis.* Pittsburgh, PA: Author.

Smolkin, L., & Donovan, C. (2000). *The contexts of comprhension: Information book read-alouds and comprehension acquisition (CIERA #2-009)*. Retrieved from Ann Arbor, MI.

Sparks, S. (May 11, 2016). Teaching English Language Learners: What does the research tell us. *Education Week*.

Stahl, S. (2003). How words are learned incrementally over multiple exposures. *American Educator, Spring, 18–19*, 44.

Stahl, S., & Nagy, W. (2006). *Teaching word meanings.* Mahwah, NJ: Erlbaum.

Stanovich, K., West, R., & Harrison, M. (1995). Knowledge growth and maintenance across the life span: The role of print exposure. *Journal of Educational Psychology, 31*, 811–826. http://dx.doi.org/10.1037/0012-1649.31.5.811.

Storch, S., & Whitehurst, G. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*, 934–947. http://dx.doi.org/10.1037/0012-149.38.6.934.

Waxman, S. R., & Namy, L. (1997). Challenging the notion of thematic preferences in young children. *Developmental Psychology, 33*, 555–567. http://dx.doi.org/10.1037//0012-1649.33.3.555.

Wellman, H., & Gelman, S. (1998). Knowledge acquisition in foundational domains. In W. Damon, D. Kuhn, & R. Siegler (Eds.). *Handbook of child psychology: Cognition, perception, and language* (pp. 523–574). New York: Wiley.

Whitehurst, G., & Lonigan, C. (1998). Child development and emergent literacy. *Child Development, 69*, 848–872. http://dx.doi.org/10.1111/j.1467-8624.1998.tb06247x.

Williams, J., Hall, K., Lauer, K., Stafford, B., DeSisto, L., & deCani, J. (2005). Expository text compreheension in the primary grade classroom. *Journal of Educational Psychology, 97*(4), 538–550. http://dx.doi.org/10.1037/0022-0663.97.4.538.

Williams, J., Stafford, K. B., Lauer, K., Hall, K., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology, 101*(1), 1–20. http://dx.doi.org/10.1037/a0013152.

Willingham, D. (2006). How knowledge helps. *American Educator, 30*, 30–37.

Wright, T., & Gotwals, A. (2017). Supporting kindergartners' science talk in the context of an integrated science and disciplinary literacy curriculum. *Elementary School Journal, 117*, 513–537. http://dx.doi.org/10.1086/690273.

Yopp, R., & Yopp, H. (2006). Informational texts as read-alouds at school and home. *Journal of Literacy Research, 38*, 37–51.

Yopp, H., & Yopp, R. (2012). Young children's limited and narrow exposure to informational text. *Reading Teacher, 65*(7), 480–490. http://dx.doi.org/10.1002/trtr.01072.