

Understanding, Measuring, and Fostering  
Preschool Children's Acquisition of Vocabulary Depth

By

Elizabeth Burke Hadley

Dissertation

Submitted to the Faculty of the  
Graduate School of Vanderbilt University  
in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Learning, Teaching, and Diversity

May, 2017

Nashville, Tennessee

Approved:

David K. Dickinson, Ph.D.  
Deborah Wells Rowe, Ph.D.  
Amanda Goodwin, Ph.D.  
Kathy Hirsh-Pasek, Ph.D.

To John and William, my heart and my home.

## ACKNOWLEDGMENTS

I would like to thank my advisor, **David Dickinson**, for his expert mentorship, for always pushing me a bit beyond what I believed I was capable of, for helping me to develop as a writer and researcher, for never accepting less than my best, and for (partially) breaking me of my semi-colon habit. David has given me so many opportunities, and I am grateful to him for taking a chance on me back in 2011. Thanks to my dissertation committee, whose generous engagement with my work has shaped and refined my thinking. I would like to thank **Debbie Rowe** for her kindness and generosity, for expanding my understanding of early literacy, and whose own work inspires me to be a better writer. Thank you also to **Amanda Goodwin** for the thoughtful feedback on my work, willingness to share professional advice, and for modeling for me the path of an early career researcher and teacher. I would like to thank **Kathy Hirsh-Pasek** for her incisive, thoughtful critiques, infectious enthusiasm for research, and endless good cheer and support (I particularly love receiving a BRAVA!!! email from Kathy). Thanks to **Molly Collins** for her friendship and mentorship, for working with me on research problems large and small, and for leaving me a hot meal or bunch of wildflowers on my front porch just when I needed them. Thank you also to **Roberta Golinkoff** and the rest of the RPL/L4R team. Roberta insisted on sending me a marked-up manuscript in the mail addressed to “Dr. Hadley” when I was barely halfway through my program; thank you for knowing that this day would come.

I would like to thank the children and teachers in the Nashville schools who opened their classrooms to me, and to my own former students and colleagues at Elizabeth Seton High School who inspire my current work.

I would like to thank my academic sisters: **Jin-Sil Mock**, **Katherine Newman**, **Mary Miller**, and **Sascha Mowrey**. From the moment Jin-Sil picked me up from the airport at recruitment weekend to the final dissertation writing session with Sascha, you all walked with me through this program, sharing your love, compassion, and empathy along the way.

Thank you to my beloved sister **Ruth Sweeney** and her husband **Conor**, for their support in a million tangible and intangible ways, but most notably, lending me the computer that allowed me to finish this dissertation. I would like to thank my parents, **Kate and Bill Burke**, for their unconditional love and support, for the excellent educational opportunities they provided for me, and for teaching me the values of hard work, generosity of spirit, and old-fashioned Nebraskan tenacity. I hope I always make you proud. **John and William**, I love you more than words can say, and I’m grateful for the many ways in which you made this work possible.

## TABLE OF CONTENTS

	Page
DEDICATION .....	ii
ACKNOWLEDGMENTS .....	iii
LIST OF TABLES .....	vi
LIST OF FIGURES .....	vii
Chapter	
I. INTRODUCTION .....	1
References.....	7
II. EXAMINING THE ACQUISITION OF VOCABULARY DEPTH AMONG PRESCHOOL STUDENTS.....	10
Methods.....	21
Results.....	28
Discussion.....	36
References.....	43
III. MEASURING YOUNG CHILDREN’S WORD KNOWLEDGE: A CONCEPTUAL REVIEW.....	52
Vocabulary Assessment in Early Childhood .....	53
Dimensions of Vocabulary Knowledge: Breadth & Depth .....	55
Review of Selected Vocabulary Measures .....	67
Conclusions and Recommendations .....	81
References.....	87
IV. BUILDING SEMANTIC NETWORKS: THE IMPACT OF A VOCABULARY INTERVENTION ON PRESCHOOLERS’ DEPTH OF WORD KNOWLEDGE .....	99
Theoretical Framework.....	100
Factors that Support Depth of Knowledge.....	102
Methods.....	115
Results.....	126
Discussion.....	136
References.....	146
V. Appendix	

A. Examples of Child Responses and Codes Assigned .....	156
B. Form Class, Difficulty, and Means for Target, Exposure, and Control Words Tested on NWDT-M Measure.....	157
C. Words Used in Intervention.....	159
D. Selected Transcripts from Play Sessions .....	160
E. Teacher-Child Interaction Codes and Examples.....	161
F. Examples of Student Responses and Codes Assigned.....	162

## LIST OF TABLES

Table	Page
CHAPTER II	
1. Demographic Information for Sample .....	23
2. NWDT-M Descriptive Statistics: Number of Information Units Per Word by Level of Instruction.....	29
3. NWDT-M Descriptives for Target Words by Form Class.....	34
4. Unstandardized Coefficients (Standard Errors) and Effect Sizes of Growth in Knowledge of Target Words by Word Type and Semantic Information Type Used to Teach Words.....	35
CHAPTER III	
1. General and Specific Vocabulary Assessments in Use with Preschool-1 <sup>st</sup> Grade Learners.....	71
CHAPTER IV	
1. Breadth and Depth Measure Unadjusted Means (Standard Deviations) .....	127
2. Parameter Estimates (Standard Errors) for Growth in Word Knowledge (Top Panel) and Effect Sizes (Bottom Panel).....	128
3. Parameter Estimates (Standard Errors) for Teaching Words in One vs. Two Units (Top Panel) and Effect Sizes (Bottom Panel).....	130
4. Parameter Estimates (Standard Errors) for Effect of Taxonomy Words (Top Panel) and Effect Sizes (Bottom Panel).....	131
5. Correlation Matrices for Outcome Measures, Interaction Types, and Word Supports .....	133
6. Parameter Estimates (Standard Errors) for Interaction Types (Top Panel) and Effect Sizes (Bottom Panel).....	135
7. Parameter Estimates (Standard Errors) for Word Supports (Top Panel) and Effect Sizes (Bottom Panel).....	136
8. One child's responses on depth measure for a taxonomy and a theme word at pretest and posttest.....	140

## LIST OF FIGURES

Figure	Page
CHAPTER II	
1. Cohen's <i>d</i> pre-post effect sizes for concrete nouns, verbs, abstract nouns, and adjectives.....	33

## CHAPTER I

### INTRODUCTION

The connection between vocabulary and reading comprehension has been well-established (Cunningham & Stanovich, 1997; Storch & Whitehurst, 2002), but the nature of this relationship is complex and not yet fully clear. In the past, much of this research has studied the impact of vocabulary breadth, or the number of words in a person's lexicon, on reading comprehension. However, a growing body of research suggests that vocabulary depth, or the quality of one's knowledge about words, may play a unique and particularly powerful role in supporting students' ability to understand what they read (Ouellette, 2006; Roth, Speece, & Cooper, 2002; Tannenbaum, Torgesen, & Wagner, 2006). In this proposal, I draw on several theories, Perfetti's (2007) Lexical Quality Hypothesis, Anderson & Freebody's (1981) knowledge hypothesis, and Kintsch's (1988) construction-integration model of text comprehension, that highlight the importance of depth for reading comprehension. Additionally, I draw on language research that employs the tools of network science to model vocabulary knowledge as semantic networks (Steyvers & Tenenbaum, 2005). These theories serve as the framework for my investigation of how to best measure, understand, and foster children's acquisition of depth, the three main topics of the dissertation proposed here.

The first of these theories, the Lexical Quality Hypothesis (LQH; Perfetti, 2007) is chiefly concerned with depth at the word-specific level, or the quality of information a learner might have for individual words (Perfetti, 2007; Perfetti & Hart, 2002). Children's knowledge of individual words can range from low to high quality along a number of dimensions, including knowledge of form and meaning. Higher-quality representations of words can be retrieved quickly and efficiently, therefore facilitating comprehension of text. The form and meaning



components of these representations are closely connected in that both must be well-developed if a word is to be easily retrieved when needed. High-quality word identities therefore are comprised of both a stable phonological representation as well as rich semantic information that allows the learner to distinguish the word in question from other semantically-related words.

Given studies that demonstrate that children continue to refine their understanding of even very common words into their elementary years (Gropen, Pinker, Hollander, & Goldberg, 1991), it seems likely that children can produce a consistent, fully specified phonological representation of a word long before their semantic knowledge of that word is similarly well-developed. Language research with young children supports this hypothesis, showing that 1.5-2 year-olds are able to establish very minimal representations of word forms, consisting only of phonological (Estes, Evans, Alibali, & Saffran, 2007; Swingley, 2007) or syntactic (Yuan & Fisher, 2009) representations of words that can be “filled in” later with semantic information. These initial, shallow representations of words are often quite fragile (Bion, Borovsky, & Fernald, 2013). In one study children initially “learned” words for novel objects, but had forgotten them again after only a five-minute delay (Horst & Samuelson, 2008).

Considered as a whole, this line of research suggests that many of the lexical representations in young children’s vocabularies may be quite shallow and lacking in semantic information. If children are to retain these words long-term, deeper processing of words is needed (Craik, 2002). This deep processing involves active, hands-on experiences analyzing the semantic meaning and connotations of words (McKeown & Beck, 2014), which support the learner’s ability to fit the word in question into a larger semantic framework of related words and concepts (Craik, 2002). Another related conclusion to be drawn from the work on children’s initial representations of words is that the basic phonological and syntactic forms of words may

be established first, with the semantic aspect lagging behind. Support for semantic meaning in particular, then, must be provided for many words, especially those that represent complex concepts (Stahl & Nagy, 2006), if they are to develop into the high-quality lexical representations needed for reading comprehension. Semantic knowledge therefore emerges as a key aspect of the LQH for those wishing to support the acquisition of vocabulary depth in young children.

The LQH provides an explanation of how individual high-quality lexical representations facilitate comprehension. However, this micro-level view lacks a full account of how lexical representations are connected to one another, and how these semantic relationships also contribute to a learner's depth of knowledge. Anderson & Freebody's (1981) knowledge hypothesis takes a broader, macro-level view of depth, considering the quality or richness of the learner's overall vocabulary and the interconnections between individual word identities. The knowledge hypothesis proposes that it is not vocabulary words themselves that directly support comprehension, but the rich underlying networks of conceptual knowledge represented by those words that students draw on as they read. Word labels in this view are just the tip of the iceberg, signaling the interconnected networks of knowledge beneath the surface that are necessary for making meaning from text. These schemas provide a filter for reading text that allows learners to "fill in the gaps" as they read and make inferences beyond what is on the page.

More recent work, most notably Kintsch's (1988, 2005) construction-integration (CI) model, extends and specifies the general account given by Anderson & Freebody. Both Anderson & Freebody (1981) and Kintsch (1988) see knowledge as essential to comprehension. But rather than viewing knowledge as organized into static schema, the CI model employs a more flexible representation of knowledge called associative networks, which includes multiple

meanings of words, contextual information, and related words (or semantic neighbors), parts of which are activated when text is encountered. During the construction phase of the CI model, an initial mental representation of the text is formed, using the words and phrases from the text as well as the learner's associative network of knowledge for those words and phrases. The resulting "text base" can itself be conceptualized as a chaotic, incoherent associative network in which many different meanings of words are represented, whether or not they fit the context at hand. In the integration phase, associations or meanings that are not appropriate to the context are deactivated, and the text base is edited and shaped into a coherent mental representation of the text.

Both stages of this process, construction and integration, rely on learner's deep word knowledge. In the construction phase, all of a learner's knowledge about word meanings and associations between words is activated to construct an initial mental representation of the text (Kintsch & Mangalath, 2011). If the learner does not have sufficient knowledge about words to draw on, comprehension is impaired. Likewise, the integration phase selects only the aspects of word meaning that are relevant to the particular context, a process that depends on a learner's prior experience with words in context to operate correctly. Text comprehension, then, is limited without a deep, flexible knowledge of words, their multiple meanings, and how they might be used in context. The construction-integration model highlights the importance of a learner's high quality knowledge of not only individual words, but of the relationships between words and how they are used in context.

These theories, considered together, provide conceptual support for the position that at both the word-specific and the general level, vocabulary depth plays a key role in reading comprehension. These theoretical perspectives, which emphasize the importance of depth, help

to shed light on an important quandary in the field: although there is a great deal of correlational evidence (Cunningham & Stanovich, 1997; Ouellette, 2006; Storch & Whitehurst, 2002) for the link between reading comprehension and vocabulary, vocabulary interventions have reported relatively modest effects on reading comprehension outcomes (Elleman, Lindo, Morphy, & Compton, 2009). If vocabulary knowledge is indeed central to reading comprehension, we would expect that increased support for vocabulary would lead to more substantial improvements in reading comprehension.

One possible explanation for the somewhat weak impact of vocabulary instruction on reading comprehension is that many of the studies reported in the meta-analysis by Elleman et al. (2009), and indeed, many of the vocabulary interventions conducted with early childhood learners (Hoffman, Teale, & Paciga, 2014), use measures that tap word knowledge only at the shallowest end of the spectrum. Elleman et al. (2009) specifically invoke “poorly conceptualized, unreliable measures” (p. 35) as a key reason why, in their meta-analysis, the impact of vocabulary instruction on comprehension was relatively weak. Pearson, Hiebert, & Kamil (2007) similarly criticize the current state of vocabulary assessment as “grossly undernourished, both in its theoretical and practical aspects,” and call for measures that “are as conceptually rich as the phenomenon (vocabulary knowledge) they are intended to measure” (2007, pp. 282-283). Measures that consider not just size, but also quality, of learners’ vocabularies are necessary for a more multifaceted view of vocabulary and its relationship with reading comprehension.

A second possible explanation is that the instructional methods used in the interventions may not have adequately supported the building of the high quality lexical representations that are important for reading comprehension. As has been demonstrated in early language acquisition research (Carey, 1978; Dollaghan, 1985), it is quite possible for learners to acquire

only a shallow representation of a word and still be able to identify its referent. However, given the low quality of these representations, they are unlikely to be retained long-term (Bion et al., 2013) or be helpful for reading comprehension (Perfetti, 2007) without further development. More research is needed on the kinds of encounters with words that help to promote depth and retention.

Both of these issues, measurement and instruction, relate to the need for a renewed focus on quality of word knowledge. In the dissertation proposed here, I argue that vocabulary depth may be an overlooked dimension in the field's conceptualization of vocabulary that can help to enrich our understanding of its relationship with reading comprehension. Broadly, this dissertation study is framed in response to the research question, "How can we better measure, understand, and foster children's acquisition of high quality lexical representations?" Paper One responds to the need for more information about vocabulary depth in young children, examining preschoolers' acquisition of depth through a detailed look at the kind of semantic information learned by word type. Paper Two focuses on measurement issues, reviewing commonly used early childhood vocabulary assessments through the lens of lexical quality and providing guidelines for designing conceptually richer measures. Paper Three reports on an intervention designed to investigate how depth can be fostered through instruction, with a particular focus on the impact of teaching words in categories and teaching words across multiple contexts. I also examine associations between growth in word knowledge and 1) adult-child interactions, and 2) supports for word meaning.

Together, these three papers seek to shift the focus in vocabulary interventions and measurement to include a greater attention to quality of word knowledge, as well as to inform theory about depth of vocabulary acquisition.

## References

- Anderson, R. C., & Freebody, P. (1981). *Vocabulary knowledge*. Newark: International Reading Association.
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39-53.  
doi:<http://dx.doi.org/10.1016/j.cognition.2012.08.008>
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.
- Craik, F. I. M. (2002). Levels of processing: Past, present... and future? *Memory*, *10*(5-6), 305-318.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934-945.
- Dollaghan, C. (1985). Child meets word: "Fast mapping" in preschool children. *Journal of Speech and Hearing Research*, *28*(3), 449-454.  
doi:<http://dx.doi.org/10.1044/jshr.2803.454>
- Elleman, A. M., Lindo, E. J., Morphy, P., & Compton, D. L. (2009). The impact of vocabulary instruction on passage-level comprehension of school-age children: A meta-analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.  
doi:10.1080/19345740802539200
- Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254-260. doi:10.1111/j.1467-9280.2007.01885.x

- Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, *18*(1), 115-151.  
doi:<http://dx.doi.org/10.1017/S0305000900013325>
- Hoffman, J. L., Teale, W. H., & Paciga, K. A. (2014). Assessing vocabulary learning in early childhood. *Journal of Early Childhood Literacy*, *14*(4), 482-509.  
doi:10.1177/1468798413501184
- Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128-157. doi:10.1080/15250000701795598
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kintsch, W., & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science*, *3*(2), 346-370.
- McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly*, *29*(4), 520-530.
- Ouellette, G. P. (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.  
doi:<http://dx.doi.org/10.1037/0022-0663.98.3.554>
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, *42*(2), 282-296.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*(4), 357-383. doi:<http://dx.doi.org/10.1080/10888430701530730>
- Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P.

- Rietsma (Eds.), *Precursors of functional literacy* (pp. 67-86). Amsterdam: John Benjamins.
- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95*(5), 259-272. doi:10.1080/00220670209596600
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*: Lawrence Erlbaum Associates Publishers.
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78. doi:10.1207/s15516709cog2901\_3
- 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. doi:http://dx.doi.org/10.1037/0012-1649.38.6.934
- Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology, 43*(2), 454-464. doi:10.1037/0012-1649.43.2.454
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398. doi:10.1207/s1532799xssr1004\_3
- Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-Year-Olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*(5), 619-626. doi:10.1111/j.1467-9280.2009.02341.x



## CHAPTER II

### EXAMINING THE ACQUISITION OF VOCABULARY DEPTH AMONG PRESCHOOL STUDENTS<sup>1</sup>

A number of studies have established the important connection between reading comprehension and vocabulary knowledge (Cain, Oakhill, & Bryant, 2004; Quinn, Wagner, Petscher, & Lopez, 2015; Vellutino, Tunmer, Jaccard, & Chen, 2007). However, there is a growing awareness that vocabulary knowledge is a complex construct that cannot be understood solely in terms of breadth, or number of words known (Christ, 2011; Schoonen & Verhallen, 2008). Vocabulary breadth is a descriptor of the overall number of entries in a learner's lexicon, each of which may be known to a greater or lesser extent. Vocabulary depth, a related but distinct aspect of word knowledge, is a descriptor of how well the individual entries in one's lexicon are known (Anderson & Freebody, 1981). In other words, depth can be defined as a learner's richness of knowledge about individual words and has also been shown to contribute to students' ability to understand what they read (Ouellette, 2006). However, depth has been less frequently explored than breadth in the literature, with many vocabulary intervention studies focusing on number of words learned, without asking how much and what kind of knowledge students have gained about individual words or whether this knowledge is of sufficiently high quality to impact reading comprehension.

Although the concept of depth itself has several different dimensions, this article focuses on two key aspects: richness of semantic representation of words and knowledge of use in typical contexts (Nation & Snowling, 1998; Ordóñez, Carlo, Snow, & McLaughlin, 2002). We respond

---

<sup>1</sup> This chapter was published in *Reading Research Quarterly* in the April/May/June issue, 2016. The citation can be found in the references (Hadley, Dickinson, Hirsh-Pasek, Golinkoff, & Nesbitt, 2016).

to the need for more detailed information about what kinds of instruction can help foster depth. We present a nuanced portrait of preschoolers' acquisition of deep word learning, examining which kinds of semantic information children learned from instruction for words ranging in perceptual accessibility. Our data are drawn from a vocabulary intervention designed to evaluate children's word learning from shared book reading paired with play sessions with varying levels of adult support (Dickinson et al., 2013). This multiphase intervention has been shown to have strong effects on children's depth and breadth of word knowledge (Dickinson et al., 2013). Analyses reported in the present study focus on children's depth of word learning, looking at results by both word type and semantic information.

### **Acquiring Deep Word Knowledge**

The initial process of learning a new word has often been described as fast mapping (Carey, 1978): the quick learning of a few aspects of a word after only a few incidental exposures. Fast-mapped information includes the association between an object and a word label, limited information about the context in which the new word is encountered, and the ability to produce some of the phonemes in the word label (Dollaghan, 1985). Children with a fast-mapped, limited understanding of a word may identify a picture of the word on a receptive vocabulary measure but lack a deeper understanding of the word's nuances and uses in multiple contexts. Therefore, although children may, in a superficial sense, know the word, their semantic knowledge may not be extensive enough for them to use the word in real-world settings or draw on it when comprehending text.

A deep understanding of word meanings is acquired gradually over time (Bion, Borovsky, & Fernald, 2013; Yu & Smith, 2012). Bloom (2000) described the rate of children's word learning not, as is often cited, as learning 10 new words a day but as "learning one-

hundredth of each of a thousand different words” (p. 25). He pointed to research showing that common verbs such as *pour* and *fill* are not fully understood even by children as old as 7 (Gropen, Pinker, Hollander, & Goldberg, 1991). Research by Medina, Snedeker, Trueswell, and Gleitman (2011), however, suggested that if the context in which a word is learned is highly informative, such as in a book-reading session in which the story and pictures illustrate the word’s meaning, children may be able to gain a great deal of knowledge by encountering the word just once. Encountering a word in multiple, varied contexts, such as in a book-reading session and then during a guided play activity, may also facilitate deeper word knowledge than learning a word in a single context (Bolger, Balass, Landen, & Perfetti, 2008).

Henriksen (1999) described the process of gaining deep word knowledge as network building: discovering links between the word in question and other related terms. Henriksen gave an example of network building for the word *hot*: A child might learn antonyms for this word (e.g., *cold*), synonyms (e.g., *warm*), words that vary in gradation (e.g., *scalding*, *tepid*), and multiple typical contexts for the use of the word.

### **Deep Word Knowledge and Reading Comprehension**

Networks of knowledge associated with words have been identified as a key factor in the relationship between vocabulary knowledge and reading comprehension. Anderson and Freebody (1981) posited that vocabulary and reading comprehension are related because vocabulary serves as an indication of conceptual knowledge. According to this theory, a person can understand what is read not only because he or she knows individual word meanings but also because he or she has built extensive networks of conceptual knowledge from which to draw on, of which vocabulary is the tip of the iceberg. To improve reading comprehension, therefore, vocabulary instruction must build deep, conceptually rich knowledge. Neuman and Celano

(2006) suggested that once children begin to acquire this conceptually rich knowledge, they become able to acquire more knowledge at faster rates and that, conversely, children lacking in conceptually rich knowledge fall further behind their peers.

Perfetti's (2007) lexical quality hypothesis, a theoretical model describing the reading process, similarly highlights the centrality of networks of word knowledge in reading comprehension. The hypothesis views comprehension as dependent on the ability to retrieve word identities, which in turn relies on the lexical quality of a word, or how much knowledge a reader has about the form and meaning of a particular word, as well as how closely these elements are connected to one another. As a learner has more experience with a word in a variety of contexts (Bolger et al., 2008), its phonological representation becomes more stable, more grammatical classes and inflections of the word are learned, and the meaning becomes incrementally more precise and less bound to context. High-quality representations, or semantic networks in which elements of form and meaning are tightly connected to one another, can be quickly retrieved, whereas low-quality representations threaten a reader's retrieval speed and ability to comprehend a passage.

The lexical quality hypothesis provides a theoretical framework for understanding the relationship between high-quality word knowledge and reading comprehension, which has been demonstrated in several studies. A study that measured depth by evaluating fourth-grade students' ability to verbally define words showed that depth predicted reading comprehension beyond the association explained by measures of breadth (Ouellette, 2006). Proctor, Silverman, Haring, and Montecillo (2012) corroborated these findings: Depth was a significant predictor of reading comprehension for students in grades 2–4, even after controlling for decoding and vocabulary breadth. In a longitudinal study, Roth, Speece, and Cooper (2002) reported a strong

correlation ( $r = .70$ ) between kindergartners' abilities to give oral definitions (a measure of depth) and their second-grade reading comprehension. Finally, the National Early Literacy Panel (2008) found that definitional vocabulary (a depth of knowledge measure) was more predictive of later decoding and reading comprehension than breadth.

Studies with struggling readers at the elementary level have further corroborated the importance of deep word knowledge for reading comprehension. Nation and Snowling (1998) found that elementary school-age students who had difficulties with reading comprehension scored the same as their peers on measures assessing phonological skills but did less well on tests that measured semantic abilities. These results indicate that students who struggled with reading comprehension did so not because of weak phonological skills but because they had limited semantic knowledge, which led to slower semantic processing and poor comprehension.

Landi and Perfetti (2007) suggested that this weakness in semantic knowledge is due to a lack of relevant experience with words, such as repeated exposures to words' phonological and semantic features. Once a pattern of inadequate exposure to words is established, it can have long-reaching reciprocal effects: A paucity of high-quality experiences with words leads to weak semantic representations, which then leads to poor reading comprehension (Landi & Perfetti, 2007). As early as first grade, students who struggle with reading have lower self-concepts about reading and are rated by teachers as more likely to be off task during class reading activities and less likely to read independently on their own (Morgan, Fuchs, Compton, Cordray, & Fuchs, 2008). This lack of sufficient exposure to print then limits students' ability to build high-quality representations of words, which continues the cycle (Landi & Perfetti, 2007). Helping young students build high-quality representations of words through rich instruction and repeated exposures is therefore essential in providing a strong foundation for reading success.

These studies have shown the importance of deep word knowledge in reading comprehension for elementary students. Given the predictiveness of early language to later language (Storch & Whitehurst, 2002) and the fact that early language ability is highly predictive of later language competence (Dickinson & McCabe, 2001), the lexical representations that children build early on are likely to be of key importance once they begin reading. Efforts to foster young children's vocabulary learning, therefore, should focus not only on adding new words to their lexicons but also on building rich, high-quality representations of words. These two instructional goals are likely complementary: As learners add new words to their lexicons, their networks of word knowledge become more refined and precise for distinguishing new entries from old ones (Carey, 1978), thereby increasing depth, and when learners gain rich knowledge about a number of aspects of a word, they likely learn other, related words along the way, thereby increasing breadth.

### **Semantic Networks by Word Type**

Depth of word knowledge can be conceptualized as a rich network of semantic associations around a word that support semantic processing and reading comprehension. The content of these networks, however, is thought to vary by form class (Miller & Fellbaum, 1991), which may also have consequences for how well different words are learned. A number of studies have reported that verbs are more difficult to learn than nouns (see Gentner, 1982; Imai et al., 2008), even for adults (Gillette, Gleitman, Gleitman, & Lederer, 1999). Maguire, Hirsh-Pasek, and Golinkoff (2006) suggested that all words lie on a continuum of concepts and that words are easier or more difficult to learn based not on their form class but on how perceptually accessible they are, as determined by the factors of shape, individuation, concreteness, and imageability.

*Shape* can be understood as the extent to which an object or action has a reliable outline or contour (e.g., a cup has a more consistent shape than a person dancing, which has a more consistent shape than someone thinking; Maguire et al., 2006). *Individuation* refers to the ease with which an item can be distinguished from others in a scene (Gentner & Boroditsky, 2001). *Concreteness* refers to whether something is a tangible object (Paivio, Yuille, & Madigan, 1968), and *imageability* to how readily one can produce a mental image for that word (Maguire et al., 2006). These four elements, considered together, characterize a word's perceptual accessibility.

Because verbs tend to lie on the less perceptually accessible end of this continuum, they are generally more difficult to learn, although verbs such as *walking* that easily produce a mental image and have a reliable, consistent shape are easier to learn than more abstract verbs such as *thinking*. Nouns also fall at various points along the continuum. For example, the noun *justice* may be quite difficult to learn, as it is highly abstract and difficult to form a mental image for, whereas a perceptually accessible, easily individuated object such as *sword* may be quite simple to learn. Although verbs vary in perceptual accessibility, in the present study, eight of the nine verbs tested fall on the more perceptually accessible end of the continuum, so we analyzed them together (see Appendix B for more specific information about each word). The nouns tested vary in their perceptual accessibility, so we divided them into “abstract” and “concrete” categories (the abstract nouns are also less perceptually accessible in terms of shape, individuation, and imageability, but we use the terms *concrete nouns* and *abstract nouns* to reflect the common usage of these terms in the literature). We use the term *word type* rather than *form class* to reflect this division in nouns.

The idea that words fall along a continuum from less to more perceptually accessible has important consequences for theories of vocabulary depth. It indicates that the types of semantic

information available for words along the continuum will be qualitatively different, so a learner's semantic network for a concrete noun will look different from his or her semantic network for an abstract verb. This also suggests that the instructional information that can be provided for different word types will also be different. In our study, we examine children's learning of words that fall at different points on the continuum, enabling us to determine the impact of word type on learning.

### **Semantic networks for concrete nouns.**

**Functional information.** For concrete nouns, functional information, or information about what an object does or is used for, has been found to be highly salient for preschool word learners (Booth, 2009; Greif, Nelson, Keil, & Gutierrez, 2006). Preschoolers were found to be more likely to learn words that were described in terms of their function (e.g., a shovel is used to dig) than in nonfunctional terms (e.g., a shovel has a part that is made out of metal; Booth, 2009; Nelson, O'Neil, & Asher, 2008).

**Hierarchical information.** Word knowledge also includes the understanding of hierarchical relationships among concrete nouns. This dimension involves the ability not simply to add nodes to the semantic network but also to categorize the relationships among nodes, such as the ability to identify superordinates and subordinates (Verhallen & Schoonen, 1993). Another type of hierarchical relationship among concrete nouns is that of part-whole relations (Henriksen, 1999; Verhallen & Schoonen, 1993). For a word such as fish, various characteristic component parts may be included in a child's semantic network, such as scales, fins, gills, and tail.

**Perceptual qualities.** For concrete nouns, the object's perceptual qualities constitute additional nodes in the semantic network. Although gleaned perceptual information about an



object is often thought of as only a first step toward gaining a deeper knowledge of word meaning (Hollich, Hirsh-Pasek, & Golinkoff, 2000), deciding which perceptual qualities are characteristic of particular nouns is a skill that reveals deeper conceptual knowledge (Booth & Waxman, 2002). Perceptual information about objects (e.g., cats have fur, armor is made out of metal, gold is yellow) provides important fodder for the sophisticated process of categorizing what type of object or material something is and how it can be differentiated from other similar objects or materials.

### **Semantic networks for concrete and abstract nouns, verbs, and adjectives.**

**Synonyms.** Another key aspect of a semantic network for concrete and abstract nouns, verbs, and adjectives is synonyms, or the core meaning(s) of a word. A synonym can be either a single word or a short, decontextualized definition when a single-word synonym does not exist (Miller & Fellbaum, 1991). A verb's core meaning is often expressed by using a synonym with a manner qualification, such as "To (verb 1) is to (verb 2) in some manner" (Miller & Fellbaum, 1991). For example, the meaning of the word gallop might be expressed in this way: "To gallop is to run fast."

Knowledge of synonyms is often the deciding factor in whether a child knows a word, demonstrating a decontextualized knowledge of word meaning. A student's ability to select synonyms for a word is a commonly used standardized measure of receptive vocabulary knowledge (e.g., Test of Word Knowledge; Wiig & Secord, 1992). In teaching vocabulary, giving synonyms or short, decontextualized word meaning explanations has been shown to help primary-grade students learn new words (Biemiller & Boote, 2006). However, Beck, McKeown, and Kucan (2002) cautioned that providing decontextualized word meaning information is

helpful only when couched in child-friendly language and also paired with more contextualized examples of how a word is used.

**Gestures.** Although knowledge about words is often thought of as only verbal, there is a growing awareness that embodied experiences of words may help support comprehension and that gestures serve as another way of representing the meanings of words (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004). Just as words serve as a label for underlying semantic information, gestures can serve as a similar kind of label, although they can also provide supplementary information about meaning in ways a word label does not (Göksun, Hirsh-Pasek, & Golinkoff, 2010). Gesture plays an important supporting role in communication because of its ability to clarify or supplement spoken language, especially when that language is complex (McNeil, Alibali, & Evans, 2000). Pairing language with gesture, rather than using language alone, was shown to improve comprehension for preschoolers (McNeil et al., 2000) and for first and second graders (Glenberg et al., 2004). McNeil and colleagues posited that gestures can act as a scaffold for verbal information, helping to guide children's attention to the semantic content of complex language. Gesture may also function as an alternative way for children to express knowledge before they can explain it verbally (Capone, 2007). Examining gestural representations of knowledge, therefore, may show increases in student learning that would otherwise be overlooked (Verhallen & Schoonen, 1993).

Gestures may be a particularly powerful way of teaching the meanings of verbs. Although concrete nouns have stable perceptual features, verbs are dynamic and require children to abstract the verbal essence of an action (Golinkoff et al., 2002). Gestures may be a useful way of representing verbs because they filter out the noise and distill an action into a limited, more meaningful essence.

**Contextual information.** Finally, rich word knowledge must include not only semantic information but also an ability to use a word in different contexts (Nagy & Townsend, 2012; Stahl & Fairbanks, 1986). Nagy and Scott (2000) argued that knowing a word means being able to do things with it, and the ability to use a word correctly in context shows a deep, applied knowledge of a word and its use. Like other elements of depth, the ability to appropriately use a word develops over time, progressing from a basic association with the word and its typical context of use (e.g., knowing that the word has something to do with \_\_\_; Dale, 1965), to being able to use the word in a single context, to eventually learning to use the word flexibly in a range of contexts (Clark, 2010).

### **Summary**

Deep word knowledge for concrete nouns includes functional, perceptual, and hierarchical information, synonyms, nonverbal information, and knowledge about how words are used. The concepts underlying abstract nouns and verbs are complex and relational and, therefore, more easily described with synonyms or gestures or through meaningful context. Adjectives lend themselves to networks of knowledge composed of synonyms, gestures, and use in context.

### **Measuring Depth of Word Knowledge**

Depth of word knowledge is an important construct that is rarely measured in vocabulary interventions. Instead, many assess only whether a child knows a word, using tasks such as asking children to select a picture from a group that matches a word spoken by the tester (Dunn & Dunn, 1997; Ouellette, 2006). These kinds of measures do not account for the fact that words can be known to a lesser or greater extent and that differences in depth of knowledge have consequences for reading comprehension (Perfetti, 2007).

The vocabulary measure used in the present study is a definition task in which we ask students to tell us what they know about words. Wasik and Hindman (2014) suggested that for preemergent readers who learn words through oral language, not print, a slightly more limited conception of depth is appropriate: a facility with basic phonological, semantic, syntactic, and contextual, but not orthographic, aspects of a word. Definition tasks capture the quality of the semantic and contextual information a child knows about a word and may also indirectly tap into the strength of his or her phonological and syntactic representation. Consistent with the lexical quality hypothesis, definition tasks test the main criterion for a high-quality representation: the ability to retrieve word identities.

The following questions are addressed:

1. Did children's depth of vocabulary knowledge for target words increase, as compared with their knowledge of exposure and control words?
2. Did increases in depth of knowledge for target words vary by word type?
3. How did the kind of semantic information learned vary by word type?

## **Methods**

### **Study Design and Research Participants**

The present study was conducted as part of Read, Play, Learn, a project aimed at increasing the vocabulary knowledge of preschoolers from low-income backgrounds through a combined method of book reading and play (Dickinson et al., 2013; Weisberg et al., 2015). The goal of the larger intervention was to test the efficacy of play combined with book reading as a method of vocabulary instruction; however, the present study focuses specifically on increases in depth of children's word knowledge by word type (concrete and abstract nouns, verbs, and

adjectives) during the course of this intervention, without examining the efficacy of the book-reading and play methods in detail.

This study uses a within-subjects research design in which children served as their own controls. To determine whether children learned a significant amount about words during the intervention, they were assessed on three kinds of words: target words, which were part of the book text, used in the play sessions, and explicitly defined; exposure words, which were not explicitly taught or defined but were in the book read-alouds and used in the play sessions; and control words, which were not used in the intervention at all. This design allowed us to test whether the effects of the intervention were due to merely hearing words used repeatedly or if additional teaching of the words made a significant difference in students' learning.

Data come from 240 four- and five-year-old students (mean age = 59.3 months, standard deviation = 4.8 months). Eighty-five of these students were enrolled in seven Head Start classrooms in a mid-Atlantic U.S. state, and 155 were enrolled in 11 preschool classrooms from a state-funded program for low-income families in a Southeastern U.S. city. The sample included only students who did not have an Individualized Education Plan and who understood enough English to be able to follow directions, as reported by their teacher. Table 1 summarizes the demographics for the sample. The intervention was delivered by nine female intervention specialists, all of whom possessed a bachelor's or master's degree plus experience in early childhood settings.

Table 1. *Demographic Information for Sample*

	<i>N</i>	Percent of Sample
Gender		
Female	130	54.2
Male	110	45.8
Race/Ethnicity		
African-American	131	54.6
Hispanic	55	22.9
Caucasian	33	13.8
Other	16	6.6
No response	5	2.1
English Language Learner		
No	204	85.0
Yes	36	15.0

*Note.* English language learner = children who speak a language other than English at home.

### **Materials: Book and Word Selection**

The book-reading and play intervention was developed around two themes (dragon and farm), which were chosen for their appeal to young children and opportunities for play. Two books per theme were read aloud to students: *The Knight and the Dragon* by Tomie dePaola and *Dragon for Breakfast* by Eunice McMullen, or *Farmer Duck* by Martin Waddell and *Pumpkin Soup* by Helen Cooper. All four books were comparable in terms of the pictorial representations of most target words, text complexity, and length.

Ten target words per book—abstract and concrete nouns, verbs, and adjectives—were selected using the following procedures. As an initial step, we identified words in the story that were considered Tier 2, or sophisticated words of high utility (Beck et al., 2002), and would therefore need additional explanation for children to understand them fully. Additional target words were inserted in the texts because all four books lacked 10 total Tier 2 words. Because some of the books had minimal text, these adaptations typically involved adding sentences with

Tier 2 words that described the action depicted in the book’s illustrations. For example, the book *Farmer Duck* includes several illustrations of the duck doing work around the farm without any text describing his actions. We added sentences such as “The duck noticed some weeds in the garden. He took his shovel and dug those weeds out,” thereby providing a fuller description of the book’s action without significantly altering the story line.

Next, we considered whether the words could be easily explained in child-friendly terms and were semantically and phonologically distinct from one another. We also cross-referenced our selection with Biemiller’s (2010) list of words, which are rated in terms of appropriateness for instruction by grade level. Nine target words did not appear on Biemiller’s list. Of the 31 target words that were on the list, 61% were characterized as at least level T2—high-priority words that are typically known by more advanced students by the end of second grade and not known by at-risk students. Approximately half of the target words also appeared on the Dale–Chall (Chall & Dale, 1995) list of words known by 80% of fourth-grade students.

Finally, to ensure that the target words were sufficiently difficult, words that over 30% of students from the previous iteration of the experiment identified correctly at pretest were replaced. We used the same methods to select 17 exposure words and 16 control words that were comparable in difficulty to the taught words and contained similar proportions of the four word types (concrete and abstract nouns, verbs, and adjectives). Because of the significant cognitive demand of the vocabulary measure used in this study, in which students were asked to define words, we randomly selected a subset of 21 target words, 10 exposure words, and eight control words for testing on this measure. See Appendix B for a list of all words assessed, along with their word type, difficulty, and descriptive information.

## Procedures

The intervention was conducted over a two-month period, from April to May 2012. All students were individually pretested and posttested by members of the research team for knowledge of the target vocabulary within one week prior to and following the intervention, respectively. Students were randomly assigned to one of three play conditions within classrooms, and classrooms were randomly assigned to one of the two themes. Books within each theme were counterbalanced. Intervention specialists read aloud to mixed-gender groups of three students in a quiet location outside the classroom for four consecutive days during the week. Immediately following each book reading, play sessions were conducted. The current article does not focus on the main effects of different intervention methods.

**Book reading.** Intervention specialists read two books aloud to students four times over the course of the intervention. Each target word was explained as part of every book reading, once during each reading as the words occurred in the text and once after each reading was completed as part of a vocabulary and plot review. The explanation consisted of the following:

- Drawing students' attention to a word by pointing to the picture, which also helps illustrate meaning (e.g., "Look, the king is wearing spectacles" while pointing to the glasses in the picture)
- Definitional information delivered in concise, child-friendly language (e.g., "Spectacles are glasses"), including perceptual, functional (e.g., "The spectacles help the king see better"), and hierarchical information whenever possible
- The use of gesture, whenever possible, to kinesthetically reinforce meaning (e.g., "Can you pretend you are wearing spectacles like this?" while the teacher makes spectacles with rounded fingers)



- An example of a word in a context other than the one used in the story (e.g., “Look, your teacher wears spectacles, too!”)

During the third and fourth readings, students’ verbal participation was encouraged to reinforce each word’s phonological and meaning representations (e.g., “What was the king wearing to help him see?”).

**Play Conditions.** A 10-minute play session immediately followed each book reading. Three play conditions were developed to test the effects of adult-supported play on students’ word learning. Because the effects of the different play conditions are not the focus of the present study, we included play condition as a covariate in our analyses, controlling for its effects.

### **Instrument and Scoring**

*New Word Definition Test—Modified.* To measure students’ depth of knowledge of target words, an experimenter-designed measure was developed and administered at pretest and posttest. This measure was adapted from Blewitt, Rump, Shealy, and Cook’s (2009) New Word Definition Test, which we renamed as the New Word Definition Test—Modified (NWDT–M) to reflect our adaptations to the coding scheme, namely, additional categories for gestures and contextual information. This definition task employs an informal rather than a formal definition task (Snow, Cancino, De Temple, & Schley, 1991). Our focus here is not the structure or form of children’s definitions, which may reveal more about their metalinguistic skills than their knowledge of words. Therefore, the NWDT–M does not track the extent to which children give adult-like definitions of words, a skill that requires practice with the form of definitions (Read, 2004), but instead codes for the amount of accurate semantic and contextual information that students provided for each word. The results of the NWDT–M allow for an understanding of the

kinds of information that preschool students learn about words ranging in perceptual accessibility.

Students were asked to define concrete and abstract nouns, verbs, and adjectives verbally or by using gestures. Students were not tested on all target, exposure, and control words on this measure due to time constraints and the cognitive demands on children. The NWDT–M test forms for the dragon and farm themes were similar in the number of items, number of words per word type, and difficulty of words. In a previous phase of this study, we identified words that were known by more than 30% of students at pretest, and eliminated those words for the present study to ensure that all words were of a similar difficulty level.

For each word, students were asked, “What is (a) \_\_\_?” and a follow-up question, “Can you show me or tell me anything else about \_\_\_?” If a student did not respond to a question, the tester moved on to the next word. All student responses were transcribed by testers and also video- or audiotaped. A coding scheme was developed (adapted from Blewitt et al., 2009) to categorize and score student responses for the number of information units given. Coding was conducted by research assistants, and 20% of all forms were randomly selected and checked for reliability against a master coder after every four forms were completed. Overall percentage agreement averaged 93.2%, with a mean Cohen’s Kappa value of .82.

**Coding Scheme.** We used seven information unit categories to score student responses for semantic content and contextual information: perceptual qualities, functional information, part/whole, synonyms, gestures, meaningful context, and basic context. Each information unit was worth 1 point except for basic context, which was worth 0.5 point. The first three categories were used for concrete nouns only. Perceptual qualities included properties such as how something looks, smells, tastes, feels, or sounds. Functional information included any process,

purpose, or use for concrete nouns and answers the question, “What do you do with it?”

Part/whole described a distinct part of a target word or the whole that the target word was a part of. The remaining categories were used for all word types. Synonyms included any word or short phrase that was equivalent to the word being explained, and provided decontextualized meaning information. Gestures included gestures, actions, or facial expressions (e.g., the teacher uses a scary face to illustrate the word *fierce*) that showed knowledge of the word’s meaning.

We also coded for two types of use in context. Meaningful context included responses that showed knowledge of the target word in a typical, meaningful context, along with semantic information. For example, if a student said, “A shovel is used to dig up weeds in a garden,” “used to dig” would be scored for function, and “weeds in a garden” would be scored for meaningful context, because the student used a typical example to explain how a shovel could be used, along with semantic information. Basic context, worth only 0.5 point, was a simple association between a target word and a typical context, without any use of semantic information. For example, students frequently said, “Santa Claus,” for *chimney*, a response that does not include any semantic information but still contains an association with a typical context in which the target word is used. Incorrect or irrelevant responses received a score of 0. See Appendix A for more examples of coded student responses.

## **Results**

We performed multilevel regression models to address each of our three research questions, testing for (1) overall growth in depth of knowledge, comparing students’ learning of target words to that of exposure and control words; and (2) growth in depth of knowledge for target words only by word type and (3) by word type and semantic information category.

Multilevel model procedures (Raudenbush & Bryk, 2002) were used to account for interdependency among study observations.

### Descriptive Statistics

The descriptive statistics in Table 2 show that there was variance in students' vocabulary knowledge. The distribution was skewed, so log transformations were performed to improve the normality of the distribution.

Table 2. NWDT-M *Descriptive Statistics: Number of Information Units Per Word by Level of Instruction* (n = 240)

Word Type	Test Period	Min.	Max.	<i>M</i>	SD	Percentage of Sample with Responses Above Floor-Level
All Words	Pretest	0	1.17	0.20	0.21	85.4]
	Posttest	0	2.02	0.47	0.37	96.7]
Target Words	Pretest	0	1.19	0.13	0.20	60.4]
	Posttest	0	2.15	0.55	0.48	90.8]
Exposure Words	Pretest	0	2.50	0.38	0.42	68.8]
	Posttest	0	2.50	0.49	0.45	80.8]
Control Words	Pretest	0	1.50	0.19	0.23	59.6]
	Posttest	0	1.25	0.25	0.26	64.6]

*Note:* Above floor-level responses are those coded as receiving a score greater than 0.

Because students were randomly assigned to a theme (dragon or farm) and the words tested for each theme were different, it was necessary to determine whether the farm and dragon NWDT–M test forms were comparable. An independent samples *t*-test on NWDT–M pretest and posttest scores for the dragon and farm themes indicated that there was no significant difference in mean pretest NWDT–M scores in the two themes,  $t(240) = 0.72, p = .473$ , 95% confidence interval  $[-0.55, 1.42]$ , but that there was a significant difference at posttest,  $t(240) = -4.00, p < .001$ , 95% confidence interval  $[-5.68, -1.93]$ . Because of this difference, we included theme as a covariate in all subsequent analyses.

### **Psychometric Properties of Measure**

Both of the test forms demonstrate acceptable internal consistency at pretest (farm theme: Cronbach's  $\alpha = .73$ ; dragon theme: Cronbach's  $\alpha = .76$ ) and posttest (farm theme: Cronbach's  $\alpha = .76$ ; dragon theme: Cronbach's  $\alpha = .86$ ). We also evaluated the validity of the NWDT–M by comparing the test scores on a concurrent measure of a highly related construct (Cronbach, 1971; Messick, 1989), in this case, the receptive vocabulary measure used in this study. Although the NWDT–M measures depth of word knowledge and the receptive measure evaluates breadth of knowledge, studies have shown that these two constructs are distinct but related. The correlations between the NWDT–M and receptive measure were statistically significant ( $r = .41$  at pretest,  $r = .62$  at posttest), demonstrating a moderately strong relationship between them but also indicating that they do not measure exactly the same construct.

### **Overall Growth in Depth of Knowledge**

Although our primary interest in this study was to examine children's growth in depth of knowledge by word type and semantic information category, it was necessary first to determine whether they learned a significant amount about taught words in general before we carried out

more detailed analyses. To answer our first research question about overall growth in depth of knowledge, we compared children’s learning of taught words with their learning of exposure and control words. Using the following multilevel regression model, we tested whether vocabulary gains varied by level of instruction (target, exposure, and control words):

$$\text{Posttest}_{ijk} = \gamma_{000} + (\gamma_{100} \times \text{Taught}_{ijk}) + (\gamma_{200} \times \text{Exposure}_{ijk}) + (\gamma_{300} \times \text{Pretest}_{ijk}) + (\gamma_{010} \times \text{Age}_{jk}) + (\gamma_{001} \times \text{Theme}_k) + (\gamma_{002} \times \text{Condition}_k) + U_{00k} + U_{0jk} + r_{ijk} \quad (1)$$

The model accounted for three nesting levels in the data: Level of instruction<sub>ijk</sub> (target, exposure, and control) was repeated within children<sub>jk</sub> ( $n = 240$ ), and children were nested within play groups<sub>k</sub> ( $n = 85$ ). For parsimony, the classroom random effects were aggregated at the reading playgroup level. Level of instruction was dummy coded with control words as the reference group, which were contrasted to taught ( $\gamma_{100}$ ) and exposure ( $\gamma_{200}$ ) words. To look at residualized gains, students’ pretest vocabulary scores ( $\gamma_{300}$ ) were included as a covariate, along with age at pretest ( $\gamma_{010}$ ), book theme ( $\gamma_{001}$ ), and play condition ( $\gamma_{002}$ ). Book theme was coded with the dragon theme as the reference group, and play condition was coded with free play as the reference group. Theme was included as a covariate because the words used in the two themes were different from one another.

Analysis indicated that after accounting for the model covariates of pretest vocabulary knowledge,  $\gamma_{300} = 0.59$ , standard error ( $SE$ ) = 0.04,  $p < .001$ ; age at pretest,  $\gamma_{010} = 0.003$ ,  $SE = 0.001$ ,  $p < .001$ ; theme,  $\gamma_{001} = -0.03$ ,  $SE = 0.01$ ,  $p < .001$ ; and play condition,  $\gamma_{002}$  (free play vs. guided play = 0.03,  $SE = 0.01$ ,  $p = .001$ , free play vs. directed play = 0.03,  $SE = 0.01$ ,  $p = .001$ ), students knew more taught words at posttest than control words,  $\gamma_{100} = 0.10$ ,  $SE = 0.01$ ,  $p < .001$ , and more exposure words than control words,  $\gamma_{200} = 0.04$ ,  $SE = 0.01$ ,  $p < .001$ . Post hoc pairwise

comparisons with least significant difference adjustments also confirmed that students knew more taught words than exposure words ( $p < .001$ ). On average, after controlling for covariates, students gave 4.68 more information units at posttest for the target words as a whole, or 0.42 more information units per target word. The pretest–posttest effect sizes were 1.22 for target words, 0.26 for exposure words, and 0.22 for control words.

### **Growth in Depth of Knowledge by Word Type**

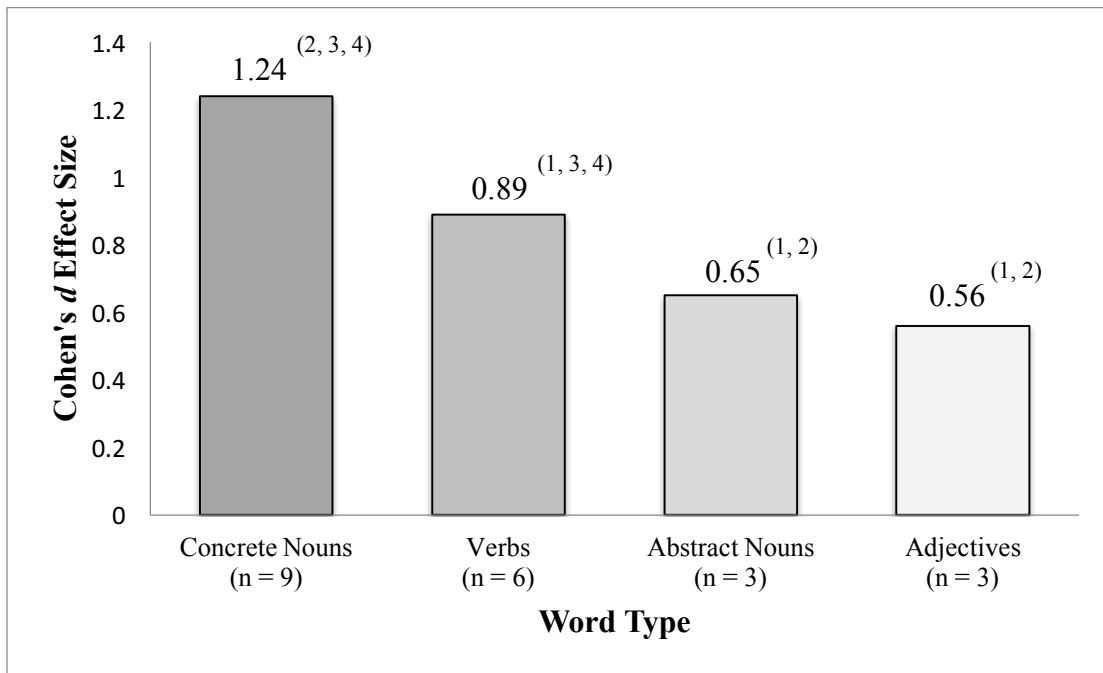
The first analysis determined that students had indeed shown significantly greater growth in their knowledge of taught words than exposure and control words. Further analyses were conducted only on target words and examined students’ growth in knowledge of target words in more detail. Our second research question proposed to investigate how students’ learning varied by word type. Using the following multilevel regression model, we tested whether vocabulary gains varied by word type (concrete nouns, abstract nouns, verbs, and adjectives).

$$\begin{aligned} \text{Posttest}_{ijk} = & \gamma_{000} + (\gamma_{100} \times \text{Verbs}_{ijk}) + (\gamma_{200} \times \text{AbstractNouns}_{ijk}) + (\gamma_{300} \times \\ & \text{Adjectives}_{ijk}) + (\gamma_{400} \times \text{Pretest}_{ijk}) + (\gamma_{010} \times \text{Age}_{jk}) + (\gamma_{001} \times \text{Theme}_k) + (\gamma_{002} \times \\ & \text{Condition}_k) + U_{00k} + U_{0jk} + r_{ijk} \quad (2) \end{aligned}$$

The model accounted for three nesting levels in the data; word type<sub>ijk</sub> (concrete nouns, abstract nouns, verbs, and adjectives) was repeated within children<sub>jk</sub>, and children were nested within play groups<sub>k</sub>. Word type was dummy coded with concrete nouns as the reference group, which was contrasted to verbs ( $\gamma_{100}$ ), abstract nouns ( $\gamma_{200}$ ), and adjectives ( $\gamma_{300}$ ).

This analysis indicated that students showed significantly greater growth in their knowledge of concrete nouns as compared with verbs,  $\gamma_{100} = -0.05$ ,  $SE = 0.01$ ,  $p < .001$ ; abstract nouns,  $\gamma_{200} = -0.12$ ,  $SE = 0.01$ ,  $p < .001$ ; and adjectives,  $\gamma_{300} = -0.13$ ,  $SE = 0.01$ ,  $p < .001$ . Post hoc pairwise comparisons with least significant difference adjustments revealed that students

also learned significantly more about verbs than abstract nouns ( $p < .001$ ) and adjectives ( $p < .001$ ). There was no significant difference in the learning of adjectives and abstract nouns. Students showed significant growth in knowledge for each of the four word types from pretest to posttest ( $p < .001$ ). Figure 1 shows the pretest–posttest effect sizes for each word type and the significant contrasts in learning by word type. Table 3 lists descriptive information for target words by word type.



*Figure 1.* Cohen's  $d$  pre-post effect sizes for concrete nouns, verbs, abstract nouns, and adjectives.

Reference group is level of instruction at time 1 (pretest). <sup>(1)</sup> significant difference ( $p < .001$ ) from concrete nouns; <sup>(2)</sup> significant difference ( $p < .001$ ) from verbs; <sup>(3)</sup> significant difference ( $p < .001$ ) from abstract nouns; <sup>(4)</sup> significant difference ( $p < .001$ ) from adjectives.



Table 3. NWDT-M Descriptives for Target Words by Form Class (n = 240)

Word Class	Test Period	Min.	Max.	<i>M</i>	SD	Percentage of Sample with Responses Above Floor-Level
Concrete Noun	Pretest	0.00	2.33	0.18	0.32	43.8]
	Posttest	0.00	3.00	0.78	0.70	85.0]
Verb	Pretest	0.00	2.50	0.22	0.38	45.4]
	Posttest	0.00	2.00	0.57	0.55	75.4]
Adjective	Pretest	0.00	2.00	0.06	0.23	7.9 ]
	Posttest	0.00	2.00	0.24	0.42	31.7]
Abstract Noun	Pretest	0.00	2.00	0.05	0.22	9.6 ]
	Posttest	0.00	2.33	0.30	0.49	35.8]

Note: Above floor-level responses are those coded as receiving a score greater than 0.

### Growth in Depth of Knowledge by Word Type and Semantic Information Category

Our third research question asked how the kind of semantic information learned varied by word type. To address this question, we tested whether significant gains were made from pretest to posttest for each semantic information category by word type, using the following multilevel regression model:

$$\text{Vocab}_{ijk} = \gamma_{000} + (\gamma_{100} \times \text{Observation}_{ijk}) + (\gamma_{010} \times \text{Age}_{jk}) + (\gamma_{001} \times \text{Theme}_k) + (\gamma_{002} \times \text{Condition}_k) + U_{00k} + U_{0jk} + r_{ijk} \quad (3)$$

The model accounted for three nesting levels in the data; assessment observation<sub>ijk</sub> was repeated within children<sub>jk</sub>, who were nested within play groups<sub>k</sub>. The independent variable of interest is Observation ( $\gamma_{100}$ ) and represents the contrast of pretest and posttest scores for each semantic information category by word type. Separate models were conducted for each semantic information category by word type. Because separate tests were run for each semantic information type (19 tests), we used a Bonferroni-adjusted  $\alpha$  level of .003 per test (.05/19) to determine significance. The kind of semantic information that students learned differed by word type (see Table 4).

Table 4. *Unstandardized Coefficients (Standard Errors) and Effect Sizes of Growth in Knowledge of Target Words by Word Type and Semantic Information Type Used to Teach Words*

Word Type	Semantic Information Category	% of Words Taught Using Semantic Information Category	Coefficient ( <i>b</i> )	<i>d</i>
Concrete Noun				
	Function	89.0	0.21 (0.02)	0.96*
	Meaningful Context	100.0	0.15 (0.01)	0.84*
	Part/Whole	33.0	0.12 (0.01)	0.83*
	Synonym	100.0	0.07 (0.01)	0.65*
	Gesture	11.0	0.04 (0.01)	0.53*
	Perceptual Quality	56.0	0.05 (0.01)	0.40*
	Basic Context		0.03 (0.01)	0.25*
Verb				
	Synonym	100.0	0.12 (0.01)	0.79*
	Meaningful Context	100.0	0.09 (0.01)	0.53*
	Gesture	67.0	0.05 (0.01)	0.42*
	Basic Context		0.04 (0.01)	0.29*
Abstract Noun				
	Synonym	100.0	0.05 (0.01)	0.49*
	Meaningful Context	100.0	0.04 (0.01)	0.38*
	Basic Context		<0.01 (0.01)	0.04
	Gesture	0.0	<0.01 (0.02)	0.03
Adjective				
	Synonym	100.0	0.05 (0.01)	0.46*
	Basic Context		0.02 (0.01)	0.15
	Meaningful Context	100.0	0.01 (0.01)	0.13
	Gesture	67.0	<0.01 (0.02)	0.11

Note. *N* = 240. *d* = Cohen's *d* standardized mean-difference effect size. Reference group is level of knowledge at time 1 (pretest).

\* *p* < .001

Students showed significant growth in their knowledge of all semantic information categories for concrete nouns, learning functional information best, followed by meaningful context, synonyms, part-whole relations, gestures, perceptual qualities, and basic context. Students also showed growth in all semantic information categories for verbs, learning synonyms best, followed by meaningful context, gestures, and basic context. For abstract nouns, students showed significant growth only in their knowledge of synonyms and meaningful context. They

showed no growth in knowledge of the basic context and gesture categories for abstract nouns. Finally, students showed growth in knowledge of synonyms for adjectives. Although meaningful context was taught for all of the adjectives in the study, and gesture was taught for 67% of them, students did not show significant growth in those categories. Table 4 shows the growth in depth of knowledge from pretest to posttest for taught words in all word types by semantic information category, along with the percentages of words in each category taught using that kind of semantic information and effect sizes for each category

### **Discussion**

Depth of knowledge is an important and distinct facet of vocabulary knowledge that supports reading comprehension (Ouellette, 2006). Because children who already have rich vocabulary knowledge are better able to acquire more rich vocabulary knowledge, and those who lack that knowledge fall further behind (Neuman & Celano, 2006), there is a pressing need for efforts that focus on building vocabulary depth in young children. However, there is very little information about the kind of instruction that fosters this learning. We know that increased encounters with words build depth (Beck & McKeown, 2007), but there has been little research addressing the question of which specific kinds of information about words are best learned by children, therefore adding to their depth of knowledge. The present study addresses this gap by showing that certain kinds of input are especially helpful in fostering depth and that these kinds of input vary by word type.

The words we taught in this study fell at different points along the conceptual continuum, ranging in their perceptual accessibility. Concrete nouns such as *handkerchief* were easy to observe and individuate, had a consistent shape, and were highly imageable. In contrast, verbs such as *returning* and abstract nouns such as *plan* had no consistent shape, were difficult to

picture, and could not be physically manipulated or as easily observed in the world. Over the course of the intervention, all of these words were taught by providing a short verbal definition and contextual information, but beyond these common features, we did not systematically control the kind of information supplied about words, because different words lent themselves to different kinds of supportive information. Indeed, it would be nearly impossible to fully vary word type by information type. For example, although words such as *returning* and *plan* can be defined, it is difficult to use an iconic gesture to represent these words or to supply functional information for them. This variability in instruction and in word types provides an opportunity to examine children's depth of word learning in a detailed way, looking at their relative learning of words by both word type and the categories of semantic information that students were able to learn for each word type.

### **Results by Word Type**

Our results are consistent with Maguire and colleagues' (2006) theory that perceptual accessibility, which includes the factors of shape, imageability, concreteness, and individuation, predicts the ease with which words are learned. The words that showed the most growth in learning were those that were the most perceptually accessible. Students learned the most about concrete nouns (Cohen's  $d = 1.24$ ), followed by verbs (of which all but one were perceptually accessible; Cohen's  $d = 0.89$ ), abstract nouns (Cohen's  $d = 0.65$ ), and finally adjectives (Cohen's  $d = 0.56$ ). This finding applies not only to form class differences but also to differences within form classes; in this study, the highly abstract concepts labeled by certain nouns were more difficult to learn than concrete verbs.

## Semantic Information Categories

The conceptual continuum theory posits that perceptual accessibility is the primary determinant of ease of learning. Our examination of the kinds of information that we could naturally supply about different words made evident that the types of information that can be provided about word meanings may be a second factor that affects word learning and may help explain differential learning of varying types of words.

We found that certain types of semantic information were more often retained than others. The most perceptually accessible category of words, concrete nouns, was taught with a synonym, an explanation of what the word meant in context, and a reference to a picture in the book (see Table 4). All but one of the concrete nouns were also explained using functional information (e.g., “We use nostrils to breathe”), about half were explained using perceptual qualities (e.g., “Nostrils look like little holes”), and a third of the words were explained by pointing out a part of the object or the whole of which the object is a part (e.g., “Scales are on a fish’s body”; see Table 4). Students’ learning of concrete nouns reflects this instruction. They showed significant growth in their knowledge of function, context, synonym, part–whole relationships, gesture, and perceptual qualities. Interestingly, function was the category that showed the most growth (Cohen’s  $d = 0.96$ ) for concrete nouns, and although only a third of the words were explained using part–whole relationships, this type of semantic information also had a large effect size (Cohen’s  $d = 0.83$ ).

The less perceptually accessible words in the study—abstract nouns, verbs, and adjectives—were also taught with a synonym and an explanation of what the word meant in context. Pictures from the book were referenced for two thirds of the abstract nouns and verbs and about a third of the adjectives. Gesture was another important element of instruction for

these words: Two thirds of the verbs and adjectives were labeled with a gesture illustrating the word's meaning. Consistent with the instruction given, synonym was the best learned category for verbs (Cohen's  $d = 0.79$ ), abstract nouns (Cohen's  $d = 0.49$ ), and adjectives (Cohen's  $d = 0.46$ ). The meaningful context category also showed significant growth for all word types except adjectives. Finally, the gesture category showed significant growth for verbs (Cohen's  $d = 0.42$ ) but not for adjectives, in spite of the instruction given.

These results demonstrate that not only were the concrete nouns better learned because of their perceptual accessibility, but they also naturally lent themselves to fuller, more varied kinds of instructional information. The less perceptually accessible words, in contrast, were not only less imageable, less concrete, more difficult to individuate, and without a consistent shape but were also more difficult to define in terms of function or discuss as a part or whole. Therefore, our data suggest that speed of learning of words may reflect the converging effects of both their perceptual accessibility and the type of information that can be provided.

### **Educational Implications**

Our results can help inform vocabulary instruction by guiding the type of information that teachers use to explain new words to preschoolers. For the concrete nouns in our study, students found functional information highly salient. This finding is in line with Booth's (2009; Perfetti, 2007) findings that providing functional information for objects is a powerful way to help increase preschoolers' depth of word knowledge. To support word learning for concrete nouns, then, teachers should not only exploit their perceptual accessibility by using pictures and pointing out important parts of the object but should also explain an object's function. Perceptual information can also serve as a gateway to conceptual information (Booth, 2009): for example,

describing what armor looks like (e.g., hard, made out of metal) naturally leads into a description of what it is used for (to protect a person's body in a fight).

These results also make it clear that verbs (even the concrete ones used in this study), abstract nouns, and adjectives simply have fewer readily describable features as compared with concrete nouns. In spite of this limitation, the verbs taught in this study were well learned (Cohen's  $d = 0.89$ ), most likely because they are relatively accessible perceptually. Because of these features, we were able to teach students simple, easy-to-remember gestures for many of the verbs. This suggests that teaching verbs through gesture or other forms of embodied learning can indeed serve as a helpful scaffold for the verbal information provided (McNeil et al., 2000).

The growth in learning for abstract nouns (Cohen's  $d = 0.65$ ) and adjectives (Cohen's  $d = 0.56$ ) was modest but still significant. Given that these types of words are abstract, have few describable features, and are difficult to explain, how can we help children learn them? One important takeaway from these results is that both synonyms and meaningful context were well learned for almost every word type (students did not show growth in meaningful context for adjectives). The large effect sizes for synonyms suggests that children can learn and provide brief, simple definitions, further justifying the use of clear word meaning explanations when new vocabulary words are introduced to children (Biemiller & Boote, 2006). This result held for all form classes in the study, showing that even when words are fairly abstract, children are able to learn something about a word's essential meaning through instruction.

The meaningful context category also showed significant growth for all word types except adjectives, suggesting that children not only need clear semantic information about words but also remember and use information about the typical contexts in which words are used. This finding supports instructional methods that emphasize the importance of both giving definitions

and teaching vocabulary in context (Beck et al., 2002; Biemiller & Boote, 2006; Coyne, Simmons, Kame'enui, & Stoolmiller, 2004). This may be especially important for words that are difficult to learn. Hearing a difficult, highly abstract word (e.g., *plan* in this study) used in context multiple times (in this study, at least eight times) allows children to progressively refine their knowledge of the nuances of its meaning. However, average growth per word was somewhat limited, suggesting that eight exposures may not be sufficient. Bolger et al. (2008) found that adult learners had higher quality knowledge of words when encountering them in multiple varied contexts as opposed to a single context multiple times. In this study, preschool students' encounters with words in related but different contexts, such as the book-reading and play settings, may have had additional benefits in helping students refine their word knowledge.

### **Limitations**

The number of words for each word type here is small, particularly for the adjectives and abstract noun categories, and the findings here may not be applicable to adjectives or abstract nouns that are significantly more concrete or more abstract than the ones used in this study. We have provided the specific words used (see Appendix B) to help guide interpretation.

Further studies should also look at the learning of abstract verbs to more fully explore the impact of certain types of instruction on words along the conceptual continuum.

It is also important to note that students' increases in word knowledge were relatively small (about 0.42 information units per word), given that students could theoretically score a nearly unlimited number of points for each word (although the highest score for an individual word was 6 points). However, the demands of the definition task are significant, and it is meaningful that preschool students learned and expressed semantic information about Tier 2 words with only a verbal prompt.



Furthermore, we did not test students at a later date for maintenance of vocabulary knowledge. Further studies exploring the instructional implications of different kinds of vocabulary instruction should explore whether more extensive types of vocabulary instruction lead to better retention of knowledge as opposed to brief, less comprehensive instruction.

## **Conclusions**

The present study adds to the research on children's language acquisition by examining the factors that lead to depth of learning by word type. We respond to the need in the field for reports of vocabulary interventions that discuss not only how many words children have learned but also how much, what kind of information has been learned about different types of words, and how this information can be used to better tailor vocabulary instruction. Furthermore, studies do not always report the type of words taught. Given that depth of learning may interact with word type, our study reinforces the importance of attending to word type when planning and reporting results of vocabulary interventions. Our results suggest not only that some words are learned more quickly and with greater depth because they are more perceptually accessible but also that these perceptually accessible words also lend themselves to a greater variety of highly salient instructional information types. Highly abstract words, then, are not only more difficult for children to learn on their own but also more difficult to learn through instruction. Students must gain a deep knowledge of highly abstract, conceptually complex words to achieve academic success (Snow & Uccelli, 2009), and our results demonstrate that clear information about meaning and use of words in meaningful contexts can help support learning. Further efforts must concentrate on ways to foster depth of knowledge for the words that students will need most as they progress through school.

## 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.
- Beck, I.L., & McKeown, M.G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*(3), 251–271. doi:10.1086/511706
- Beck, I.L., McKeown, M.G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guilford.
- Biemiller, A. (2010). *Words worth teaching: Closing the vocabulary gap*. Columbus, OH: SRA/McGraw-Hill.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44–62. doi:10.1037/0022-0663.98.1.44
- Bion, R.A.H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126*(1), 39–53. doi:10.1016/j.cognition.2012.08.008
- Blewitt, P., Rump, K.M., Shealy, S.E., & Cook, S.A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Education & Psychology, 101*(2), 294–304. doi:10.1037/a0013844
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

- Bolger, D.J., Balass, M., Landen, E., & Perfetti, C.A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122–159. doi:10.1080/01638530701792826
- Booth, A.E. (2009). Causal supports for early word learning. *Child Development, 80*(4), 1243–1250. doi:10.1111/j.1467-8624.2009.01328.x
- Booth, A.E., & Waxman, S.R. (2002). Word learning is ‘smart’: Evidence that conceptual information affects preschoolers’ extension of novel words. *Cognition, 84*(1), B11–B22. doi:10.1016/S0010-0277(02)00015-X
- Cain, K., Oakhill, J., & Bryant, P. (2004). Children’s reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31–42. doi:10.1037/0022-0663.96.1.31
- Capone, N.C. (2007). Tapping toddlers’ evolving semantic representation via gesture. *Journal of Speech, Language, and Hearing Research, 50*(3), 732–745. doi:10.1044/1092-4388(2007/051)
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.
- Chall, J.S., & Dale, E. (1995). *Readability revisited: The new Dale–Chall readability formula*. Cambridge, MA: Brookline.
- Christ, T. (2011). Moving past “right” or “wrong” toward a continuum of young children’s semantic knowledge. *Journal of Literacy Research, 43*(2), 130–158. doi:10.1177/1086296X11403267
- Clark, E.V. (2010). *First language acquisition* (2nd ed.). Cambridge, UK: Cambridge University Press.

- 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. doi:10.1207/s15327035ex1203\_3
- Cronbach, L.J. (1971). Test validation. In R.L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp. 443–507). Washington, DC: American Council on Education.
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English*, *42*(8), 895–901, 948.
- Dickinson, D.K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills and environmental supports of early literacy. *Learning Disabilities Research & Practice*, *16*(4), 186–202. doi:10.1111/0938-8982.00019
- Dickinson, D.K., Turner, K.A., Collins, M.F., Nicolopoulou, A., Golinkoff, R.M., Hirsh-Pasek, K., ... Rivera, B.L. (2013, April). *More word learning occurs when book reading is followed by teacher-supported play*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Dollaghan, C. (1985). Child meets word: “Fast mapping” in preschool children. *Journal of Speech and Hearing Research*, *28*(3), 449–454. doi:10.1044/jshr.2803.454
- Dunn, L.M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test* (3rd ed.). Circle Pines, MN: American Guidance Service.
- Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.
- Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S.C. Levinson (Eds.), *Language, culture, and cognition: Vol. 3. Language*

- acquisition and conceptual development* (pp. 215–256). New York, NY: Cambridge University Press.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176. doi:10.1016/S0010-0277(99)00036-0
- Glenberg, A.M., Gutierrez, T., Levin, J.R., Japuntich, S., & Kaschak, M.P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology*, *96*(3), 424–436. doi:10.1037/0022-0663.96.3.424
- Göksun, T., Hirsh-Pasek, K., & Golinkoff, R.M. (2010). How do preschoolers express cause in gesture and speech? *Cognitive Development*, *25*(1), 56–68.  
doi:10.1016/j.cogdev.2009.11.001
- Golinkoff, R.M., Chung, H.L., Hirsh-Pasek, K., Liu, J., Bertenthal, B.I., Brand, R., ... Hennon, E. (2002). Young children can extend motion verbs to point-light displays. *Developmental Psychology*, *38*(4), 604–614. doi:10.1037/0012-1649.38.4.604
- Greif, M.L., Nelson, D.G.K., Keil, F.C., & Gutierrez, F. (2006). What do children want to know about animals and artifacts? Domain-specific requests for information. *Psychological Science*, *17*(6), 455–459. doi:10.1111/j.1467-9280.2006.01727.x
- Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Syntax and semantics in the acquisition of locative verbs. *Journal of Child Language*, *18*(1), 115–151.  
doi:10.1017/S0305000900013325
- Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly*, *51*(2), 181-198. doi:10.1002/rrq.130

- Henriksen, B. (1999). Three dimensions of vocabulary development. *Studies in Second Language Acquisition*, 21(2), 303–317. doi:10.1017/S0272263199002089
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R.M. (2000). II. The emergentist coalition model. *Monographs of the Society for Research in Child Development*, 65(3), 17–29. doi:10.1111/1540-5834.00092
- Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R.M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, 79(4), 979–1000. doi:10.1111/j.1467-8624.2008.01171.x
- Landi, N., & Perfetti, C.A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, 102(1), 30–45. doi:10.1016/j.bandl.2006.11.001
- Maguire, M.J., Hirsh-Pasek, K., & Golinkoff, R.M. (2006). A unified theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364–391). New York, NY: Oxford University Press.
- McNeil, N.M., Alibali, M.W., & Evans, J.L. (2000). The role of gesture in children's comprehension of spoken language: Now they need it, now they don't. *Journal of Nonverbal Behavior*, 24(2), 131–150. doi:10.1023/A:1006657929803
- Medina, T.N., Snedeker, J., Trueswell, J.C., & Gleitman, L.R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22), 9014–9019. doi:10.1073/pnas.1105040108
- Messick, S. (1989). Validity. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York, NY: American Council on Education/Macmillan.

- Miller, G.A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition*, *41*(1–3), 197–229. doi:10.1016/0010-0277(91)90036-4
- Morgan, P.L., Fuchs, D., Compton, D.L., Cordray, D.S., & Fuchs, L.S. (2008). Does early reading failure decrease children’s reading motivation? *Journal of Learning Disabilities*, *41*(5), 387–404. doi:10.1177/0022219408321112
- 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.
- Nagy, W.E., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, *47*(1), 91–108. doi:10.1002/RRQ.011
- Nation, K., & Snowling, M.J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85–101. doi:10.1006/jmla.1998.2564
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Washington, DC: National Institute for Literacy.
- Nelson, D.G.K., O’Neil, K.A., & Asher, Y.M. (2008). A mutually facilitative relationship between learning names and learning concepts in preschool children: The case of artifacts. *Journal of Cognition and Development*, *9*(2), 171–193. doi:10.1080/15248370802022621
- Neuman, S.B., & Celano, D. (2006). The knowledge gap: Implications of leveling the playing field for low-income and middle-income children. *Reading Research Quarterly*, *41*(2), 176–201. doi:10.1598/RRQ.41.2.2

- Ordóñez, C.L., Carlo, M.S., Snow, C.E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology, 94*(4), 719–728. doi:10.1037/0022-0663.94.4.719
- Ouellette, G.P. (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. doi:10.1037/0022-0663.98.3.554
- Paivio, A., Yuille, J.C., & Madigan, S.A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology, 76*(1, Pt. 2), 1–25. doi:10.1037/h0025327
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. doi:10.1080/10888430701530730
- Proctor, C.P., Silverman, R.D., Haring, J.R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing, 25*(7), 1635–1664. doi:10.1007/s11145-011-9336-5
- Quinn, J.M., Wagner, R.K., Petscher, Y., & Lopez, D. (2015). Developmental relations between vocabulary knowledge and reading comprehension: A latent change score modeling study. *Child Development, 86*(1), 159–175. doi:10.1111/cdev.12292
- Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language:*



- Selection, Acquisition, and Testing* (pp. 209-227). Philadelphia, PA: John Benjamins Publishing Company.
- Roth, F.P., Speece, D.L., & Cooper, D.H. (2002). A longitudinal analysis of the connection between oral language and early reading. *The Journal of Educational Research, 95*(5), 259–272. doi:10.1080/00220670209596600
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25*(2), 211–236. doi:10.1177/0265532207086782
- Snow, C.E., Cancino, H., De Temple, J., & Schley, S. (1991). Giving formal definitions: A linguistic or metalinguistic skill? In E. Bialystock (Ed.), *Language processing in bilingual children* (pp. 90–112). Cambridge, UK: Cambridge University Press.
- Snow, C.E., & Uccelli, P. (2009). The challenge of academic language. In D.R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112–133). New York, NY: Cambridge University Press.
- 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. doi:10.3102/00346543056001072
- 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. doi:10.1037/0012-1649.38.6.934
- 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. doi:10.1080/10888430709336632

- Verhallen, M., & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied Linguistics*, *14*(4), 344–363. doi:10.1093/applin/14.4.344
- Wasik, B.A., & Hindman, A.H. (2014). Understanding the active ingredients in an effective preschool vocabulary intervention: An exploratory study of teacher and child talk during book reading. *Early Education and Development*, *25*(7), 1035–1056. doi:10.1080/10409289.2014.896064
- Weisberg, D.S., Ilgaz, H., Hirsh-Pasek, K., Golinkoff, R., Nicolopoulou, A., & Dickinson, D.K. (2015). Shovels and swords: How realistic and fantastical themes affect children’s word learning. *Cognitive Development*, *35*, 1–14. doi:10.1016/j.cogdev.2014.11.001
- Wiig, E.H., & Secord, W. (1992). *Test of Word Knowledge*. San Antonio, TX: Psychological.
- Yu, C., & Smith, L.B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review*, *119*(1), 21–39. doi:10.1037/a0026182

### CHAPTER III

#### MEASURING YOUNG CHILDREN'S WORD KNOWLEDGE: A CONCEPTUAL REVIEW

To assess word knowledge, one must first determine what it means to know a word. The determination made by many vocabulary assessments used for young children is that a child “knows” a word if she can correctly identify an image representing that word. These known word items can then be tallied to give a rough estimate of a child’s vocabulary size, often referred to as breadth. A key question emerges from this emphasis on vocabulary breadth, however: what is the *quality* of knowledge for known words, and how might further investigating this aspect of vocabulary enrich our understanding of children’s word learning?

We argue here that early childhood vocabulary assessment could benefit from a richer, more multifaceted theoretical foundation that takes into account the complexity of what it means to know a word. We focus specifically on the understudied dimension of vocabulary depth, or the quality of one’s word knowledge. For example, a child might be able to select a picture that represents a word on a vocabulary measure, but such an assessment leaves open the question of what the child is really able to *do* with that word beyond identifying it. Are they able to also pronounce that word correctly, use it in a sentence, or understand its meaning in a text? Do they know multiple meanings for that word? Could they identify a larger category to which the word belongs? Do they know situations in which this word typically would or would not be used? Given that these are the functional vocabulary skills that we wish to foster in children if they are to be proficient readers, writers, and speakers, vocabulary depth must be measured if we are to better understand and support young children’s language and literacy development.

The purpose of this paper is to lay the groundwork for a more robust theoretical framework that can be used to both evaluate and design vocabulary measures for young children.

Using this framework, we review the major assessments currently in use with preschool to 1<sup>st</sup> grade students, in the context of vocabulary interventions or studies on the relationship between oral language development and reading comprehension. We end by making recommendations for how future measures might be designed to give a more comprehensive picture of vocabulary knowledge.

### **Vocabulary Assessment in Early Childhood**

Recent articles (Hoffman, Teale, & Paciga, 2014; Pearson, Hiebert, & Kamil, 2007) have expressed concern about the impoverished state of vocabulary assessment. Pearson and colleagues (2007), describing primarily print-based vocabulary measures, characterized this field as “grossly undernourished, both in its theoretical and practical aspects,” and called for measures that “are as conceptually rich as the phenomenon (vocabulary knowledge) they are intended to measure” (2007, pp. 282-283). Hoffman et al. (2014) were similarly critical in their review of early childhood vocabulary assessments, concluding that the majority of these measures are convenient for use by researchers but hold little practical significance.

Like Hoffman et al. (2014), we focus here on the challenge of assessing oral vocabulary in young children. This subfield of vocabulary assessment, if anything, is even more susceptible to the criticism of undernourishment made by Pearson et al. (2007). While vocabulary knowledge is a complex, multi-dimensional phenomenon, most commonly used oral vocabulary measures fail to capture this complexity, instead relying on breadth as the main indicator of word knowledge.

The potential pitfalls of this limited range of options can be illustrated through the findings of the National Early Literacy Panel’s (NELP; 2008) meta-analysis of relationships between early abilities and later conventional literacy skills. Many were surprised that the NELP

report found that early (preK-K) vocabulary was a relatively weak predictor of reading comprehension in kindergarten-2<sup>nd</sup> grade ( $r = 0.25$ ) (Dickinson, Golinkoff, & Hirsh-Pasek, 2010). A closer look, however, revealed that the definition of vocabulary used in the report was quite narrow, encompassing only vocabulary breadth (as measured by receptive assessments). Vocabulary depth (as measured by giving definitions of words), on the other hand, was a significantly stronger predictor of reading comprehension ( $r = 0.45$ ), along with other complex oral language skills such as grammar and listening comprehension. Additional research corroborates the NELP report's (2008) findings: depth has been found to play a unique and important role in the relationship between vocabulary and reading comprehension (Roth, Speece, & Cooper, 2002), even when controlling for breadth (Ouellette, 2006). While using breadth measures to assess vocabulary in early childhood studies is the norm, these results indicate that this approach tells an incomplete story about the relationship between vocabulary and reading comprehension.

Further evidence suggests that the measures used to capture growth in vocabulary over a short period of time knowledge are similarly inadequate. In a vocabulary intervention meta-analysis performed by Marulis & Neuman (2010), nearly half of the studies included used only standardized assessments, even though such measures are not well-suited for tracking growth because of their insensitivity to small increases in knowledge (National Reading Panel, 2000). While the use of such measures may test for generalization, they leave no way of knowing whether, or how well, children actually learned the specific words that were taught in the intervention. Furthermore, nearly all of the standardized measures used were picture vocabulary measures such as the PPVT or EOWPVT, which are intended to estimate breadth, but not depth, of vocabulary knowledge. Of the interventions that did use a researcher-created measure to track

increases in knowledge for specific words, only about half used a measure that tapped depth. This focus on breadth likely misses the nature and extent of word knowledge acquired by children that could better inform instructional approaches.

### **Dimensions of Vocabulary Knowledge: Breadth & Depth**

In the following section, we discuss more fully what we mean by the terms breadth and depth. Although we give definitions of each of these terms separately below, our intention is not to suggest that they are two completely separate constructs. Rather, they are interrelated aspects of vocabulary knowledge that likely grow in tandem. As a word-learner adds new words to her vocabulary, she also gains more knowledge about how those words relate to one another, adding to the overall quality of her word knowledge. We argue here that measuring both dimensions of word knowledge can lead to a richer, more detailed portrait of word-learning.

#### **Defining Breadth**

Vocabulary breadth can be considered from two perspectives: 1) general vocabulary knowledge, or a person's lexicon as a whole, and 2) word-level vocabulary knowledge, or knowledge of individual words. *General breadth* is an estimate of the overall number of items in one's lexicon, without specific attention paid to how well each item is known. *Word-level breadth* is a less useful concept, as it is difficult to understand how knowledge of a single word might be "broad," but in the literature, breadth at the word-level is often used to denote shallower, less comprehensive understanding of individual words (e.g., Hoffman et al., 2014).

The breadth of young children's vocabularies has been shown to be predictive of their language and literacy achievement well into elementary school. Vocabulary size at age two is a significant predictor of literacy outcomes such as phonological awareness, letter identification, and reading comprehension up to grade 5 (Lee, 2011), and vocabulary size in kindergarten is a

significant predictor of reading comprehension in grades 3 and 4 (Senechal, Ouellette, & Rodney, 2006). Important research has also demonstrated that vocabulary size varies by socioeconomic status (SES), with large samples of low-SES preschoolers scoring from 1 to 1.5 standard deviations below the norm on standardized breadth measures (Hindman & Wasik, 2013; Huaqing Qi, Kaiser, Milan, & Hancock, 2006).

As these studies demonstrate, measuring vocabulary in terms of numbers of words known has been a very fruitful direction. This perspective on vocabulary is necessary, but as we argue here, not sufficient.

### **Defining Depth**

There has been much confusion in the literature around the concept of vocabulary depth (Schmitt, 2014). Why is depth such a difficult construct to define and operationalize? One problem is a lack of specificity. The classic, highly influential definition of depth comes from Anderson & Freebody's work on the relationship between vocabulary and reading comprehension, which defines breadth as how many words are known and depth as how well those words are known (1981). This definition is powerful in terms of its simplicity and clarity, and has often been repeated in the literature (e.g., Read, 2004). When it comes to operationalizing depth, however, this definition is somewhat empty. What exactly does it mean to know a word well or less well?

Depth is also often defined as a learner's knowledge of multiple aspects of a word, including its phonological, orthographic, pragmatic, syntactic, and semantic features (Silverman & Hartranft, 2015). This definition adds much-needed specificity to conceptualizations of depth, but also becomes somewhat unwieldy as various definitions of depth include not only the aspects mentioned above, but also features such as morphology, collocations, and grammatical functions

(Nation, 2001). A more cohesive idea of depth is needed that has a strong theoretical position about how these elements relate to one another and what constitutes “deeper” knowledge for each.

Here, we draw on Perfetti’s Lexical Quality Hypothesis (2007) and Kintsch’s construction-integration model (1988, 2005) to address many of these issues and to define depth at both the word-level and general level.

**Depth at the word-level.** Depth at the word-level can be defined as the quality of individual items in one’s lexicon. Our understanding of depth at the word-level is rooted in Perfetti’s Lexical Quality Hypothesis (LQH; 2007), which views reading comprehension as dependent on the ability to efficiently retrieve word identities. This ability is in turn based on the lexical quality of a word, or how much knowledge a reader has about the form (orthography, phonology, and grammar) and meaning of a particular word, as well as how closely these form and meaning elements are connected to one another. The LQH presents depth as a continuum of knowledge, with individual words ranging from low to high lexical quality. High-quality representations are those that have stable representations of spelling and sound and a generalized sense of word meaning that can transfer to multiple contexts. These elements are also closely bound to one another in a single package, making the retrieval of the complete word identity quick and effortless. In contrast, low-quality lexical representations are those in which the phonological representation is less stable, the full range of form classes are not represented, and meaning may be bound to a single context. The elements of form and meaning are also only loosely connected to one another, causing inefficient or inaccurate retrieval of a word identity (for example, hearing a word and mistakenly retrieving the meaning of a similar-sounding word).



Therefore, low quality representations threaten a reader's retrieval speed and ability to comprehend a passage.

The LQH is built on an understanding of the word learning process as incremental, with each additional encounter with a word improving the quality of its representation. As a learner has more experience with a word in a variety of contexts, its phonological representation becomes more accurate, more inflections of the word are learned, and meaning becomes more refined. This incremental process of word-learning generally begins with a quick, relatively crude mapping of a word to an object or action (Hollich, Hirsh-Pasek, & Golinkoff, 2000). Research on early language development shows that children can gain a shallow understanding of a word, thus adding it to their lexicon, without fully understanding the multiple facets of its meaning (Lahey, 1988). This initial, limited learning of some aspects of a new word is termed "fast mapping" (Carey, 1978). 1.5-2 year-olds are able to establish very minimal representations of word forms, consisting mainly of phonological (Estes, Evans, Alibali, & Saffran, 2007; Swingley, 2007) or syntactic (Yuan & Fisher, 2009) representations of words that can be "filled in" later with more comprehensive semantic information. These initial, shallow representations of words are quite fragile (Bion, Borovsky, & Fernald, 2013): in one study two-year-olds initially "learned" words for novel objects, but had forgotten them again after only a five-minute delay (Horst & Samuelson, 2008).

Children who have a fast-mapped understanding of a word may be able to successfully identify its image on a receptive vocabulary measure, but lack deeper conceptual and pragmatic knowledge of the word. Therefore, while in some sense children may "know" fast-mapped words, these shallow lexical representations may not support real-world comprehension or use. Fast-mapped words would be considered low quality lexical representations in the LQH, and

would contribute relatively little to comprehension because of the delay involved in retrieving these word identities efficiently.

How, then, do learners acquire high quality word knowledge? Proceeding further along the continuum to higher-quality representations of words takes time and more encounters with words (Bion et al., 2013; Yu & Smith, 2012), a process sometimes described as “slow-mapping” (Swingley, 2010). The Instance-Based Learning Framework (Bolger, Balass, Landen, & Perfetti, 2008), provides an incremental model for depth of word learning in which each encounter with a word encodes a trace of its form and context in the memory. Over time, encountering a word in multiple varied contexts leads to the building of an abstracted representation of the word (Yurovsky, Fricker, Yu, & Smith, 2014). In a study with adult readers, Bolger et al. (2008) found that depth of learning was supported not only through the use of varied contexts, but also by supplying explicit information about word meaning. Frishkoff, Perfetti, and Collins-Thompson (2011) measured adults’ incremental learning of new words, finding that the definitions they gave increased in accuracy with each additional word-learning episode. Gains in word knowledge were also greater when words were presented in sentences that provided clues to the word’s meaning. These studies indicate that depth is built slowly and incrementally, but that this process can be expedited by providing word meaning information and using varied, supportive contexts.

The work discussed so far provides a useful theoretical framework for deep word learning, but has been primarily focused on reading and therefore validated mainly with older children and adults. However, the LQH is equally relevant to early childhood learners in that it highlights the importance of building high quality word meanings, a protracted, incremental process that must begin early if later reading comprehension is to proceed smoothly. To apply

the LQH to these young learners requires an understanding of the term “form” that refers to its phonological and grammatical aspects rather than the orthographic features that are emphasized in most discussions of lexical quality.

Several researchers have suggested that being able to use a word in an appropriate way is another important aspect of lexical quality (Nagy & Scott, 2000; Read, 2004; Silverman & Hartranft, 2015). Use is not one of the main constituents of depth in the LQH, although Perfetti sees quick retrieval of a word identity (i.e., the ability to remember or use a word) as the key indicator of high quality word knowledge. Defining lexical quality in terms of “use” is even more important when studying young children, whose knowledge of words is particularly tied to context (Snow, Cancino, De Temple, & Schley, 1991).

Therefore, we define depth at the word-level here as the richness of one’s representations of a word, including grammatical, phonological, and semantic information, the ability to use a word, and how tightly bound these elements are to one another (see Ordóñez, Carlo, Snow, & McLaughlin, 2002). In other words, lexical quality is comprised of form, meaning, and use (Nation, 2001). We consider these constituents and their development separately below, with the understanding that, in practice, they are closely bound to one another. According to Perfetti (2007), as knowledge of each element grows, so does the connections between these elements (what Perfetti terms “constituent binding”). The overall representation of the word becomes more stable and able to be retrieved as a single unit, with both form and meaning constituents retrieved at the same time.

**Form.** The first important constituent of lexical quality is that of form. For young children who are not yet reading print, form refers to the phonological representation of the word, or the storage of information about the sounds in a particular word (Ainsworth,

Welbourne, & Hesketh, 2015). Like other aspects of lexical quality, the quality of a phonological representation can range from low to high. Quality is determined by the accuracy of the representation (whether the sounds in a word are stored correctly), its stability (whether sounds in the word can be consistently and accurately produced or identified), as well as the degree of segmentation of these representations (whether items are represented as a whole word or are more finely differentiated at the syllable, onset-rime, or phoneme level) (Ainsworth et al., 2015). The Lexical Restructuring Hypothesis (Walley, Metsala, & Garlock, 2003) suggests that phonological representations are largely holistic during infancy, but that as children acquire more words that sound similar to one other, they must more finely differentiate these words' phonological representations in order to store them accurately. Acquiring more vocabulary words, then, leads to more segmentation in children's phonological representations (Metsala & Walley, 1998). A preschool-aged child with a high-quality phonological representation, then, should be able to not only accurately produce all of the word's phonemes, but will likely also have a representation that is segmented at least at the onset-rime level for high-frequency words (Ainsworth et al., 2015).

Another aspect of form includes the grammatical features of a word. According to Perfetti (2007), high quality knowledge of a word's form means knowing all grammatical classes of the word (e.g., knowing both *anger* and *angry*) and being able to manipulate the word to reflect changes in tense, mood (e.g., the conditional mood: I *would have* eaten there), person, number, and gender. A lower-quality representation of a word would mean that the knowledge of form classes was incomplete or the learner was unable to use inflected forms of a word consistently and appropriately.

**Meaning.** The second constituent of lexical quality to be considered is that of meaning. In Perfetti's LQH (2007), having higher-quality semantic information about a word allows a person to distinguish between closely related words. For example, in order to discriminate between a *shovel* and a *spade*, it is necessary to understand what each is used for, their relative sizes, and other perceptual features of each. Therefore, a learner must have a great deal of semantic information about individual words in order to retrieve the appropriate word rather than one that is semantically related.

The kind of semantic information it is possible to learn about words varies by word type, with different kinds of semantic information available for highly imageable, concrete nouns that label objects more so than for more abstract, less imageable words such as adjectives or nouns that label ideas or qualities (Hadley, Dickinson, Hirsh-Pasek, Golinkoff, & Nesbitt, 2016). Concrete nouns name parts of the world that are naturally individuated, whereas most verbs, adjectives, and abstract nouns label more diffuse, relational concepts (Gentner, 2006). Therefore, the kind of meaning information that can be learned about these words is qualitatively different.

In particular, concrete nouns have available a much wider array of semantic information categories than do other word types because of their perceptual accessibility. Perceptual information is sometimes considered to be a gateway to deeper conceptual understanding of words (Booth & Waxman, 2002), leading to the kind of refined, precise knowledge that is characteristic of high lexical quality. For example, feeling that a *helmet* is hard helps children to understand its function (to protect someone's head), while seeing that a creature has four legs helps a child to categorize that object as an *animal*. These examples illustrate two of the types of conceptual information available for concrete nouns: functional information, what something

does or is used for (e.g., a *helmet* protects one's head), and category membership (a cat is a kind of *animal*).

Other types of semantic information apply both to concrete nouns and more abstract words. For example, synonyms, or a core meaning for a word, as well as antonyms, or the opposite of a word, are available for many words across word types (Miller & Fellbaum, 1991). There are also some semantic features that are unique to word types other than concrete nouns. For example, for three-year-olds learning action verbs, causation, or who or what caused the action, is the most salient feature (Forbes & Farrar, 1993). The affective association for many abstract words, including abstract nouns (*freedom, anger*) and adjectives (*happy*), is an important element of these words' semantic representations, and have even been posited to help children acquire these words more easily (Kousta, Vigliocco, Vinson, Andrews, & Del Campo, 2011).

Another source of semantic information available for most word types is the embodied experience children have with words, including the information they can glean from gestures (Glenberg, Gutierrez, Levin, Japuntich, & Kaschak, 2004). An emerging body of research suggests that young children may be able to infer word meanings from gestures alone, without being given any other information about words (Goodrich & Kam, 2009). Gestures can be used to represent words in nearly any form class, but they may be especially helpful in conveying dynamic concepts, such as those labeled by verbs (Goodrich & Kam, 2009) and spatial words (Cartmill, Pruden, Levine, & Goldin-Meadow, 2010).

**Use.** Use refers to the ability to put word knowledge into action, such as appropriately using a word in multiple contexts to convey meaning (Silverman & Hartranft, 2015), as well as awareness of a word's connotations, typical registers, idiomatic or rhetorical uses. This aspect of word knowledge is sometimes seen as the true marker of high quality word knowledge, where

“knowing a word means being able to do things with it” (Nagy & Scott, 2000, p. 237). Under the LQH, the rapid retrieval and use of a word is also the hallmark of high quality semantic and phonological knowledge, as the ability to use a word is dependent on the amount of knowledge one has about form and meaning.

The ability to appropriately *use* a word develops slowly over time, progressing from comprehension of a word, to use of a word in a single context, to eventually using the word across a range of contexts (Clark, 2010). Children often have surprisingly restricted contexts of use for words that they appear to “know.” Seston and colleagues (Seston, Golinkoff, Ma, & Hirsh-Pasek, 2009) found that half of six-year-olds tested were unable to understand common verbs used in unfamiliar contexts (e.g., someone *sweeping* dirt away with their feet or *writing* with their finger), even when the new context included a number of details to help support comprehension. These results show that while young children may “know” common verbs, they require further exposure and support to understand and use these words in proficient, flexible ways across a range of contexts.

Low quality knowledge of use includes a range of situations: when a child has only a memory trace of a word used in a certain context (Bolger et al., 2008), but is not yet able to use the word herself, or a child who has memorized a dictionary definition for a word, but whose actual use of that word is odd and/or incorrect. For example, McKeown (1993) writes of a fifth grader who learned the definition of *devious* then, when asked to use it in a sentence, wrote, “Some drivers *devious* of the road.” As knowledge of use develops, a child may be able to use a word correctly in a single context. High-quality use would include being able to use the word in several contexts, as well as generalizing use to new contexts.

**Summary.** We review these aspects of form, meaning, and use to indicate the range of quality that is possible for individual word items along each of these dimensions. For purposes of assessment, this perspective demonstrates that knowing a word is not an all or nothing proposition, but rather can range from low to high in a variety of ways.

**Depth at the general level.** Depth does not exist only at the word-level, with individual lexical representations operating as islands of form, meaning, and use that are disconnected from other lexical representations. Depth can also be conceptualized at the general level by “zooming out” to take a wider view of how lexical representations relate to one another. We draw here on several descriptions of reading comprehension that conceptualize general depth as a rich semantic network of interconnected word identities. Anderson and Freebody’s (1981) knowledge hypothesis argues that reading comprehension depends not on knowledge of individual word meanings, but on rich networks of conceptual knowledge. Word labels in this view are just the tip of the iceberg, signaling the interconnected networks of knowledge beneath the surface that are necessary for making meaning from text. These schemas provide a filter for reading text that allows learners to “fill in the gaps” as they read and make inferences beyond what is on the page. Anderson & Freebody’s account does not explore in detail, however, the components of these conceptual networks.

More recent work, most notably Kintsch’s (1988, 2005) construction-integration (CI) model, extends and specifies the general account given by Anderson & Freebody. Both Anderson & Freebody (1981) and Kintsch (1988) see knowledge as essential to comprehension. But rather than viewing knowledge as organized into static schema, the CI model employs a more flexible representation of knowledge called associative networks, a net of related concepts that has been built through prior experience with words. Concepts that have been encountered



together in the past are linked together more strongly than those that have not. During the construction phase of the CI model, an initial mental representation of the text is formed, using the words and phrases from the text as well as parts of the learner's associative network of knowledge for those words and phrases. The resulting "text base" can itself be conceptualized as a chaotic, incoherent associative network in which many different meanings of words are represented, whether or not they fit the context at hand. In the integration phase, associations or meanings that are not appropriate to the context are deactivated, and the text base is edited and shaped into a coherent mental representation of the text.

Both stages of this process, construction and integration, rely on learner's deep word knowledge. In the construction phase, all of a learner's knowledge about word meanings and associations between words is activated to construct an initial mental representation of the text (Kintsch & Mangalath, 2011). If the learner does not have sufficient knowledge about words to draw on, comprehension is impaired. Likewise, the integration phase selects only the aspects of word meaning that are relevant to the particular context, a process that depends on a learner's prior experience with words in context to operate correctly. Text comprehension, then, relies on a deep, flexible knowledge of words, their multiple meanings, and how they might be used in context. The construction-integration model highlights the importance of a learner's general vocabulary depth, particularly the relationships between words and how they are used in context. Drawing on both Perfetti's view of lexical quality at the word level and Kintsch's theory of knowledge nets, we conceptualize general depth as a **knowledge network**. This knowledge network is made up of individual word identities comprised of form, meaning, and use, as discussed in the previous section, which are also connected to one another in associative networks. Lexical representations are linked to one another in networks when they are frequently

heard or encountered together, and these links strengthen with every additional joint use. Parts of these associative networks are then activated when language is encountered.

### **Review of Selected Vocabulary Measures**

In the following section, we will apply our theoretical framework to measures currently used in the field. The psychometric issues involved in many vocabulary measures have been discussed elsewhere (e.g., Hoffman et al., 2014), so here we address the theoretical questions of: 1) How much, and what kinds, of knowledge do measures assess? 2) What aspects of knowledge are *not* assessed? With these questions in mind, we review the major vocabulary assessments currently in use with the population most of interest to this paper, preschool to 1<sup>st</sup> grade students, in the context of vocabulary interventions or studies on the relationship between oral language development and reading comprehension. We do not include assessments that are used mainly by speech language pathologists in clinical settings for screening of language difficulties, unless those assessments have also been used with typically developing children in the types of studies of interest here.

#### **How Much, and What Kind, of Knowledge is Assessed**

We address our guiding questions by highlighting which of the three aspects of depth the measure taps (form, meaning, and use), and to what extent. Form is divided into the two major aspects discussed earlier, phonology (are children asked to recognize or produce an accurate phonological representation?) and grammar (are children asked to recognize or produce inflections of a word?). The meaning category is comprised of 1) perceptual information, measures that tap knowledge of a word's perceptual features, 2) conceptual information, measures that tap knowledge of conceptual information, such as function and category membership, and 3) knowledge networks, measures that tap general depth by exploring the

strength of connections between related words. The use in context category includes 1) familiar, or measures that only test children's understanding or production of a word in a familiar context, and 2) novel, or measures that test children's understanding or production of a word in a new context. Note that none of these categories are mutually exclusive – a single measure could theoretically test every aspect listed above.

Table 1 is not intended to suggest that the ideal measure *should* assess all three aspects of form, meaning, and use to their fullest extent, however – such a measure would likely be impractical and too taxing for young children. Perfetti's theory of constituent binding (2007), the idea that as the different aspects of word knowledge grow in quality, the connections between them grow tighter, can be used to justify the measurement of only one of these three aspects of depth. If a student shows high quality semantic knowledge of a word, we can theoretically infer that their knowledge of form and ability to use the word are similarly high quality. However, not all aspects of depth, particularly phonology, are necessarily indicative of high-quality word knowledge in general. Students can have high-quality phonological representations of words, but still struggle with reading comprehension because of poor-quality semantic representations (Nation & Snowling, 1998; Richter, Isberner, Naumann, & Neeb, 2013).

We also categorize measures in terms of whether they assess receptive or expressive knowledge. Receptive measures, which test understanding of words, are generally understood to be less difficult than expressive measures, which test a learner's ability to correctly retrieve a word identity and provide the word label (Melka, 1997). Receptive measures are often also **breadth** measures, as the format is well-suited to assessing a large number of words. Expressive measures vary in their difficulty: some require only that children produce the target word, while others ask them to use target words in a sentence. Some, but not all, expressive measures are also

**depth** measures, depending on the goal of the assessment (a count of, or the quality of knowledge for, words).

Finally, we include columns on ease of administration and scoring. These columns are intended to address some of the real-world concerns that researchers and teachers must face in assessing children’s word knowledge, and to acknowledge some of the trade-offs that must be made in terms of comprehensiveness of assessments vs. time and energy.

### **Specific vs. General Assessments**

Following a distinction made by Hoffman et al. (2014), we found that the assessments we reviewed could generally be categorized as either 1) **specific** (i.e., used to track the learning of specific words); or 2) **general** (i.e., used to gauge the general size or depth of one’s vocabulary). Each of these larger categories contain both **breadth** and **depth** measures (see Table 1).

Knowledge of specific words can range from a low quality lexical representation – perhaps just a fast-mapped understanding of form and a context-bound meaning – to a very high quality one. **Specific measures** tap levels of word knowledge at various points along this continuum, varying in the type and quality of knowledge students must demonstrate in order to get the word “right.” In this review, we include measures that require only lower-quality lexical representations as well as those that tap higher-quality representations. The measures that tap lower-quality lexical representations are typically described as **breadth measures** in the literature (e.g., Hoffman et al., 2014), and we follow that convention here, with the understanding that word knowledge lies on a continuum and there is no absolute rule that separates a specific “breadth” measure from a specific “depth” measure in terms of their cognitive demand. The main difference between **specific breadth** and **specific depth measures**

is one of intention: breadth measures are intended to assess *size* while depth measures are intended to assess *quality*. **Specific breadth measures**, then, are those that are intended to track the learning of specific words, but which have as their goal a *count* of words learned, and therefore tap only superficial knowledge of words so as to assess as many words as possible. **Specific depth measures** also track the learning of specific words, but are more concerned with the amount or *quality* of knowledge learned, and therefore assess a smaller sample of words more intensively.

**General breadth measures** are those that are intended to gauge the overall size of one's lexicon. To meet this goal, words are often assessed in a more "shallow" manner in order to get the largest possible sample of items. **General depth measures** assess the overall quality of one's lexicon, typically by gauging the quality of word knowledge a learner has for a sample of particular words, then generalizing those results to the learner's vocabulary as a whole. Ideally, these measures would also be able to not only assess knowledge of individual words, but also tap the richness of connections to other words in the knowledge network.

Table 1

*General and Specific Vocabulary Assessments in Use with Preschool-1<sup>st</sup> Grade Learners*

General/ Specific	Type of Measure	Receptive/ Expressive	Name of Measure and Ages	Aspects of Word Knowledge Assessed						Ease of Admin.	Ease of Scoring	
				Form	Meaning			Use in Context				
				Phonology	Grammar	Perceptual	Conceptual	Knowledge Networks	Familiar	Novel		
General	Breadth	Receptive	PPVT (2:6+)	X		X					E	E
		Expressive	EOWPVT (2:0+)	X		X					E	E
		Expressive	EVT (2:6+)	X		X					E	E
		Both	WJ-III Picture Vocabulary subtest (2:0+)	X		X					E	E
		Both	PLS-4 (birth+)	X		X					E	E
General	Depth	Expressive	TOLD-P:4, Oral Vocabulary subtest (4:0-8:11)	X		X	X				M	M
		Expressive	TOWK, Word Definitions subtest (5-17)	X		X	X				M	M
		Expressive	CELF (Formulated Sentences) (5-21)	X	X	X	X			X	M	M
Specific	Breadth	Receptive	PPVT-like formats	X		X					E	E
		Expressive	EOWPVT-like formats	X		X					E	E
		Expressive	EVT-like formats	X		X					E	E
Specific	Depth	Receptive	Closed-ended questions	X		X	X			X	E	E
		Expressive	Story-retelling and description	X	X	X	X		X <sup>b</sup>	X <sup>c</sup>	M	H
		Expressive	Context integration	X	X					X	E	M
		Expressive	Definition task	X	X	X	X	X			M	H

<sup>b</sup> Penno, Wilkinson, & Moore, 2002<sup>c</sup> McKeown & Beck, 2014

E = easy, M = medium, H = hard

## **Specific Measures**

We first review specific measures, those which are intended to assess the learning of specific words. All of the measures reviewed in this section were used in the context of vocabulary interventions with young children, and were created by researcher(s) to track the learning of the words taught. The measures discussed here are generally ordered from lower-quality to higher-quality representations, keeping in mind the receptive-productive dimension as well as the extent to which the measures demand knowledge of form, meaning, and use. This ordering should not be considered a strict ranking: rather, we intend to give a general impression of the range of specific measures and their demand in terms of quality of representation.

**Specific breadth measures.** A number of specific measures are modeled after general measures, especially pictorial assessments such as the Peabody Picture Vocabulary Test (PPVT, now in version IV; Dunn & Dunn, 2007), the Expressive Vocabulary Test (EVT, now in version II; Williams, 2007) and the Expressive One-Word Picture Vocabulary Test (EOWPVT, now in version 4; Brownell, 2010). The most common type of pictorial assessment is modeled after the PPVT, in which the examiner states a word and asks the child to select the referent from 3-4 illustrations. This type of researcher-created measure is widely used in early childhood vocabulary interventions to assess children's learning of target words (e.g., Pollard-Durodola et al., 2011; Roskos et al., 2008; Senechal, 1997; Wasik & Bond, 2001). PPVT-like specific measures tap children's knowledge of a word at a relatively shallow level; in order to get each answer "correct," children must have mapped a word's perceptual features to its label, but may not have deeper conceptual knowledge about the word.

Similar pictorial assessments mimic the EVT and EOWPVT in that children are shown a picture of the target word and asked to name the picture (e.g., Blewitt, Rump, Shealy, & Cook,

2009; Hargrave & Sénéchal, 2000; Whitehurst et al., 1994). This type of assessment is more demanding than PPVT-like measures, because children must provide a correct phonological representation for the word. Therefore, both a limited understanding of form and meaning, although not use, are tapped by this measure.

We can also examine these measures in terms of what they do *not* assess. The most striking feature of specific breadth measures is that they tap knowledge at the shallow end of the lexical quality continuum. This has several advantages: these measures can detect fairly minimal amounts of knowledge, which can be helpful when assessing children who have very low levels of vocabulary knowledge, or when assessing children who are very reticent or reluctant to speak to an examiner. It is also helpful when measuring learning from exposure alone, when we might not expect children's knowledge to proceed beyond a fast-mapped representation of new words. These types of measures can quickly yield a rough count of words that have been "learned" at least partially. Finally, given that many early childhood vocabulary interventions use similar measures, use of pictorial assessments allows for some comparability across studies. The value of this comparability should not be overstated, however, as difficulty of these measures ranges widely depending on choice of target words and foils and decisions about these factors are not always clearly reported.

Specific breadth measures' inability to tap deeper semantic knowledge has drawbacks when used in the context of an intervention in that they cannot distinguish between low and high quality lexical representations. For example, if a child is able to recognize a picture of the word *shield* at the beginning of an intervention, she would likely be assessed as "knowing" that word on a PPVT-like pretest. If after the intervention she can not only recognize *shield* but can also say that it is used by knights to protect themselves, she will still simply get a single point on the



PPVT-like posttest. As measured by the pictorial assessment, her knowledge of the word *shield* has remained static. Pictorial assessments' lack of sensitivity is also problematic when comparing the relative benefits of two or more instructional methods. For example, a book-reading and play vocabulary intervention for preschoolers compared children's learning of words that were explicitly taught ("target words") to words that were used in the story but not explained ("exposure words") (Dickinson et al., 2015). Children's learning of target and exposure words was measured on both a specific breadth and a specific depth measure. There was no significant difference in children's learning of target and exposure words on the breadth measures, but there was a significant difference and a meaningful effect size ( $d = 0.50$ ), between learning of target and exposure words on the depth measure. If only the breadth measure had been used, it would appear that there was no difference between children's learning of words they were merely exposed to versus the words they were explicitly taught.

Pictorial assessments can also be limited in the types of semantic knowledge they tap simply because they are picture-based and all words tested must be visually portrayed. It is much more feasible to create test items for words that are highly concrete and imageable, such as concrete nouns and concrete verbs. This is problematic because we know that children need to learn words from a variety of form classes, such as abstract verbs, spatial terms, and adjectives (Harris, Golinkoff, & Hirsh-Pasek, 2011), and words that will be helpful across a variety of academic domains, such as Tier Two words (Beck, McKeown, & Kucan, 2013). These words are also quite abstract and difficult to depict. For example, *contradict*, *precede*, and *auspicious* (Beck et al., 2013) are all Tier Two words, and none of these words immediately suggest a clear, representative image that could be used on a pictorial assessment.

The specific breadth measures reviewed here are also limited in their assessment of form and use, in that they do not assess for the flexibility of grammatical forms (e.g., can children use a verb in different tenses) or for a range of uses.

**Specific depth measures.** Measures that are intended to track the quality of knowledge for specific words taught in interventions (specific depth measures) are less commonly used than specific breadth measures. Capturing the multiple facets of depth has proved difficult from an assessment perspective, as it is impractical to intensively assess the quality of each phonological, syntactic, semantic, and pragmatic representation of words. Instead, most assessments attempt to capture one or two aspects of depth, reasoning that more advanced knowledge in one category may also demonstrate basic knowledge in the others (Read, 2004). There is no single approach for assessing depth which dominates the field, so we highlight here a handful of specific depth assessments that have been used with young children, generally ordering them in terms of how much information about form, meaning, and use they tap.

***Closed-ended questions.*** In this method, children are asked several questions about the meaning and typical contexts for each target word (e.g., Beck & McKeown, 2007; Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009). For example, such questions might ask, “Does *extraordinary* mean very hungry?” and “Would it be *extraordinary* to see a monkey at school?” (Beck & McKeown, 2007). Children are typically assessed as “knowing” a word if they answer the majority of questions for that word correctly (e.g., McKeown & Beck, 2014).

This measure assesses children’s conceptual knowledge of words as well as their ability to understand typical contexts of use. Children are not required to produce the word, its meaning, or use it in a sentence: they only need to recognize a correct use and correct meaning in order to earn full points. Administration and scoring of this task is quick and straightforward.

***Story-retelling and description measures.*** Story retelling tasks, in which children’s use of target vocabulary words is evaluated for accuracy as they retell the narrative of a book, have also been used as a measure of high quality word knowledge (Penno, Wilkinson, & Moore, 2002) because they reveal students’ ability to retrieve, pronounce, and use a word correctly in context. This measure is somewhat problematic in that children may know, but choose not to use the target vocabulary. It is also unclear from this measure whether children’s knowledge of a word has generalized beyond the storybook context. Moreover, the scoring of this task is more difficult in that it requires transcription and coding.

A more targeted version of a story-retelling measure has been used by McKeown & Beck (2014), in which students are shown a picture for each target word (e.g., two girls with their arms around each other for the word *inseparable*). Children are asked specific questions about the picture (e.g., “What can you see by looking at the girls?”) This version of a storytelling measure taps higher-quality knowledge of use because it asks children to transfer their knowledge of a word to a new context.

Both of these versions of story-telling measures are quite demanding because they test children’s ability to retrieve an appropriate target word and use it correctly in context, a skill that demonstrates a high quality, coherent lexical representation with well-established semantic and phonological information for words.

***Context integration tasks.*** McKeown & Beck (McKeown & Beck, 2014) have used a highly demanding vocabulary measure called a context integration task that tests the quality of kindergarten children’s lexical representations at the higher end of the lexical continuum. Children are asked a question that probes for understanding of a target word in context. For example, the following question was asked for the word *insist*: “Jim had to *insist* that Freddy go

on the merry-go-round. How did Freddy feel about the merry-go-round?” Children’s responses were given 1 point if they reflected knowledge of the target word (e.g., “he didn’t like it”) and 0 points if they were incorrect (e.g., “it was fun”). This measure is especially demanding because the sentences used encouraged children to interpret the target word incorrectly. The child must have a highly developed, conceptually rich semantic representation that is able to “crowd out” the alternative explanation suggested by the sentence context to answer correctly. This measure does not directly assess children’s ability to produce the phonological representation of a word or different grammatical forms. One possible issue with this measure is that it requires strong listening comprehension and self-regulation. While McKeown & Beck (2014) found that kindergarten and 1<sup>st</sup>-grade children were able to respond to these items during pilot testing, this task may be simply too difficult for preschool children. While relatively easy to administer, this measure requires coding of student responses.

***Definition tasks.*** In definition tasks, students are typically asked what they know about a word, and their responses can either be scored 1) along a continuum, so that fewer points are given for more connotative or contextual responses and more points for decontextualized responses (Biemiller & Boote, 2006; Coyne et al., 2009; Leung, 2008), or 2) for completeness of definition, in which a point is earned for each unit of semantic information given (Blewitt et al., 2009). For example, the Blewitt et al. study awarded a point for each of the following types of semantic information: superordinate category membership, synonyms, perceptual or functional properties, and parts. In some studies, points are also awarded for using the target word in a typical context or for representing the word with a gesture (Hadley et al., 2016). Most definition tasks, therefore, focus mainly on capturing the amount or quality of the semantic information a child knows about a word, but also indirectly test the quality of his or her phonological

representation, particularly if the student uses the word in their definition. Since young children have little experience with formal definitions (Snow et al., 1991), they often demonstrate their knowledge of a word by giving an example in context (e.g., for the target word *fetch*, one child responded, “I throw the ball to my dog and he *fetches* it and gives it to me,” Hadley et al., 2016), therefore showing an understanding of grammatical class and use. Definition tasks are highly demanding and tap knowledge of a word at the high end of the lexical quality continuum.

One drawback of definition tasks is that they require a great deal of oral language proficiency and children do not, or cannot, always express all of the semantic knowledge that they have about a word. While administration is easy and the format of the test is highly flexible and requires little advance preparation, this task does require recording and coding of children’s responses.

### **General Measures**

General measures are intended to gauge the overall size or quality of children’s vocabularies. The measures reviewed here are all standardized on a national sample so that results can be used to determine how a child’s vocabulary size compares to their peers’. General measures are typically used for the following purposes: 1) to determine whether the size of learners’ vocabularies is within normal range for their age, and to compare them to other children who are the same age (Farkas & Beron, 2004); 2) to assess long-term growth in vocabulary knowledge over the course of a year or more — for example, to investigate how school instruction has impacted vocabulary during a school year (Silverman & Crandell, 2010); or 3) in correlational studies that explore relationships between vocabulary and its impact on later reading achievement (Cunningham & Stanovich, 1997; Dickinson & Porche, 2011). They are

sometimes also used in short-term vocabulary interventions to measure whether the intervention affected general vocabulary growth (Hargrave & Sénéchal, 2000).

We answer the questions of “how much and what kind of knowledge is measured” by again tracking the extent to which these measures tap form, meaning, and use, as well as whether they allow for any testing of the connections between words, consistent with the knowledge network view of depth that we have proposed here.

**General breadth measures.** General breadth measures assess the overall size of one’s vocabulary. The most widely used general vocabulary breadth measure is the PPVT, in which a child is shown a page with four pictures and asked to point to the image of a target word. The EVT and EOWPVT are general vocabulary breadth measures in which children are asked to name what they see in a single picture. Other general breadth measures include the Woodcock-Johnson III Picture Vocabulary Subtest (Woodcock, McGrew, & Mather, 2001) and the Preschool Language Scale (PLS, now in version 4; (Zimmerman, Steiner, & Pond, 2002).

Like the researcher-created versions that are modeled after them, these assessments tap semantic knowledge at a shallow level so that they can get the largest possible sample of items. However, some of the most difficult items on the PPVT (mainly those intended for older children and adults) can test more finely differentiated conceptual knowledge by including foils that are closely related to the target item. These general breadth measures do not tap knowledge of use in context or a child’s grasp of various grammatical forms of a word.

**General depth measures.** General depth measures assess the overall quality of one’s vocabulary by gauging the amount or quality of word knowledge a learner has for a sample of particular words, then generalizing those results to the learner’s vocabulary as a whole. General

depth measures have most commonly been used to explore the impact of depth on later reading comprehension (National Early Literacy Panel, 2008).

**General definition tasks.** Similar to specific definition tasks, this type of measure asks children to provide definitions for words and then scores their responses for amount and quality of information given. Standardized definition tasks include the Oral Vocabulary subtest of the Test of Language Development-Primary (TOLD:P, now in version 4, Hammill & Newcomer, 2008) and the Word Definitions subtest of the Test of Word Knowledge (TOWK, Wiig & Secord, 1992). Definition tasks are particularly demanding in their assessment of the semantic aspect of depth, requiring children to think about their lexicon and express their knowledge of word meaning explicitly (Roth et al., 2002).

**Clinical Evaluation of Language Fundamentals (CELF), Formulated Sentences subtest.** For the Formulated Sentences subset of the CELF (Semel, Wiig, & Secord, 2003), children are shown a picture and asked to generate a sentence describing it while using a target word. Scoring ranges from 0-2 and responses are explicitly scored for whether the use of the target word is syntactically, pragmatically, and semantically correct. This assessment has been used as measure of depth for its ability to capture multiple dimensions of word knowledge (Proctor, Silverman, Harring, & Montecillo, 2012), as it tests children's knowledge of form (particularly grammatical aspects), conceptual knowledge, and correct usage. This measure is unique in its ability to quickly assess all three aspects of depth, and highly promising for the same reason, although it has not been widely used outside of clinical settings.

**General depth measures for older children.** A number of other standardized depth measures have been used in studies with older children, such as semantic association tasks in which students must select the words that are related to a target word from among several

choices (e.g., the Word Association Test, validated for ages 9-12,(Schoonen & Verhallen, 2008). Other options for older children include subtests of the Language Processing Test - Revised (LPT-R, validated for ages 5-11;(Richard & Hanner, 1995) which ask children to provide a synonym for a word used in a sentence or to provide as many attributes as possible for a noun. Some of these measures could be adapted for younger children. For example, a semantic associations task for preschool and kindergarten-aged children might ask them to sort several pictures of words or objects into groups that “go together” in order to assess the richness of their knowledge networks for these words.

### **Conclusions and Recommendations**

This paper began by citing Pearson et al.’s call for more conceptually rich measures in the field of vocabulary assessment. In looking more specifically at vocabulary measures currently in use with the early childhood population, it is clear that the field is indeed, as both Pearson et al. (2007) and Hoffman et al. (2014) conclude, “undernourished.”

One factor contributing to the undernourishment of the field is that many of the measures most commonly used are not explicitly grounded in a theoretical perspective on the word-learning process and how knowledge of words develops over time. Here, we have used Perfetti’s LQH (2007) as our primary conceptual framework, while also drawing on Kintsch’s (1988, 2005) and Anderson & Freebody’s (1981) view of knowledge networks to explore how commonly used measures assess different aspects of form, meaning, and use, as well as varying levels of lexical quality. In practice, however, many vocabulary interventions create custom measures or choose general measures without a similar theoretical grounding. The goal of this paper has been to provide a theoretical framework that allows for the selection and design of



measures that are not simply convenient, but are also informed by theory. We review our main findings and provide some more specific recommendations below.

### **Specific Measures: Tracking Growth in Knowledge**

Specific measures are created by researchers to track the learning of words taught through direct instruction. However, the field is currently dominated by the use of pictorial breadth assessments, which are not sensitive enough to track small increases in knowledge (Coyne et al., 2009; Dickinson & Brady, 2005). Because interventions are typically short-term and/or often seek to compare different instructional conditions, more fine-grained measures are needed that are capable of detecting relatively small differences in knowledge. These more detailed measures are also capable of providing nuanced information about children's word learning, such as the reporting of results by word type and kind of semantic information learned from instruction (Hadley et al., 2016).

Moreover, the goals of interventions are often not well-matched to the specific assessments chosen. Vocabulary interventions must be clear about the aspects of vocabulary knowledge they wish to foster and the measures that will best assess children's progress towards these goals. The measures reviewed in Table 1 provide different kinds of information about children's word learning. If the primary intervention goal is to assess whether children have gained flexible, high quality representations of new words, the results of a pictorial assessment should not be used as evidence that this goal has been met. On the other hand, if the goal is to teach for fast-mapped knowledge, teaching a lot of words at once so that an initial representation can be later developed, than a pictorial assessment might be appropriate if carefully designed. Results should be stated in a way that reflects what the assessment shows that children can do

with words (e.g., if children can label words, it does not necessarily mean that they have “learned” them or will be able to understand them in context).

One promising response to Pearson et al.’s (2007) call for “conceptual richness” in the field of vocabulary assessment is the use of multiple specific measures to assess children’s word learning. These multiple measures must provide different kinds of information about learning, though; many of the studies reviewed here use both a receptive and productive pictorial assessment, but as both of these measures tap only shallow lexical representations, the results may be redundant and represent a missed opportunity to test for higher quality lexical representations. Similarly, McKeown & Beck (2014) note that many vocabulary interventions use only one specific vocabulary measure and a general comprehension task, but that the results from such studies leave a wide gulf in understanding children’s growth in word knowledge.

The benefit of using multiple vocabulary measures that tap knowledge along the lexical quality continuum can be illustrated by a study reported by Coyne et al. (2009) in which two instructional conditions are compared, one (“Rich Instruction”) more intensive than the other (“Embedded Instruction”). This study found that the additional supports provided by the Rich Instruction condition led to more refined word knowledge, while the Embedded Instruction condition gave students only a partial knowledge of words. This finding was only possible because a battery of four measures, ranging in difficulty, was used to gauge the growth in the quality of students’ word knowledge. Interestingly, the measures that tapped more surface-level knowledge of words showed no difference in learning between the two conditions: only the more sensitive, demanding measures showed that Rich Instruction was more effective.

This study, along with McKeown and Beck (2014), who used a similar battery of assessments, represents important steps in a promising direction. Both studies articulate a

theoretically grounded view of vocabulary knowledge, set out clear goals for learning, and use multiple measures that tap learning at various points along the lexical quality continuum.

### **General Measures**

General measures are typically used for three purposes in early childhood research: 1) determining whether the size of children's vocabularies is within normal range for their age, 2) assessing growth in vocabulary knowledge, and 3) exploring relationships between vocabulary and reading comprehension. These measures are well-suited for the first purpose, but must be employed thoughtfully when used for the second and third.

**Assessing long-term growth in vocabulary knowledge.** Some vocabulary interventions (McKeown & Beck, 2014) or observational studies (e.g., Wasik & Bond, 2001) use general measures at pretest and posttest to assess global growth in vocabulary knowledge. These measures can serve as an indication of whether increases in vocabulary knowledge have generalized beyond the specific words learned in an intervention, although they may not be sensitive enough to detect small changes in vocabulary knowledge resulting from short-term or less intensive interventions (e.g., Silverman et al., 2014). The majority of vocabulary interventions that include general vocabulary measures use only breadth, not depth, measures (e.g., in the vocabulary intervention meta-analysis performed by Marulis & Neuman, 2010, 42 out of 48 studies used general breadth measures only). This represents a missed opportunity, as investigating not only how the *size*, but also the *quality*, of children's general vocabulary has changed over time may give additional information about the effectiveness of interventions. Future research should investigate the relative benefit of including general depth and/or breadth measures as a complement to the specific measures used in interventions. Feasibility must also

be considered, given that depth measures are often more time-consuming to score and require more developed language abilities from children.

**Exploring relationships between vocabulary and reading comprehension.** General depth measures have been shown to be of value in predicting later reading comprehension (National Early Literacy Panel, 2008), but the use of such measures is not yet widespread. Future research should explore how the inclusion of general depth measures can add to our understanding of the relationship between reading comprehension and vocabulary.

### **Recommendations**

For vocabulary interventions, we suggest the use of a battery of **specific** measures: 1) a researcher-created breadth measure that requires only lower-quality, fast-mapped knowledge and allows for a rough count of words “learned” (a receptive measure such as a carefully designed pictorial assessment or closed-ended question measure); 2) a researcher-created depth measure that emphasizes children’s ability to *use* a word accurately (such as McKeown & Beck’s picture description task, 2014 or a researcher-created version of the CELF-Formulated Sentences task, Semel et al., 2003), and 3) a depth measure that tests for very high quality knowledge of words and emphasizes knowledge of *meaning* (such as a definition task, Hadley et al., 2016). This collection of measures ensures not only that low levels of knowledge are captured, but also that relatively subtle differences in learning can be pinpointed by the highly sensitive depth measures.

We also suggest the use of the PPVT or other **standardized** breadth measure at pretest, not to test for generalization in word knowledge, but to gauge the size of children’s vocabularies in relation to their peers and to allow for comparability of samples across studies.

We also hope future research explores the use of other measures, standardized or researcher-created, that test for generalization of knowledge. While it is an important goal of

vocabulary interventions to determine if, and to what extent, children have learned taught words, the ultimate goal is to actually improve children's general vocabularies and perhaps even word-learning abilities. Innovative assessments that can detect changes in vocabulary abilities beyond breadth are needed. For example, assessments such as the Diagnostic Evaluation of Language Variation (DELV, Seymour, Roeper, & de Villiers, 2005) contain tasks that could be adapted to test for changes in children's word-learning ability, such as the Fast Mapping test, which tests a child's ability to infer the meaning of a new verb after hearing it in context. Similarly, Neuman, Newman, & Dwyer (2011) developed a task that tested children's ability to categorize new words after participating in an intervention in which they were taught words in taxonomic categories. More widespread use of similar assessments that test children's fast-mapping or lexical organization could provide a novel view into not only whether vocabulary interventions can change what children *know*, but whether they can change how they *learn*.

Finally, we also see a notable gap in general vocabulary measures in the lack of depth assessments that are aligned with a view of knowledge networks. The majority of general depth measures are definition tasks that assess the quality of knowledge of individual items. A general depth measure for young children that assessed the organization of children's semantic networks, perhaps by testing the ability to successfully group similar words, would be a theoretically sound and potentially fruitful addition to the general vocabulary measures currently available.

## References

- Ainsworth, S., Welbourne, S., & Hesketh, A. (2015). Lexical restructuring in preliterate children: Evidence from novel measures of phonological representation. *Applied Psycholinguistics, FirstView*, 1-27. doi:doi:10.1017/S0142716415000338
- Anderson, R. C., & Freebody, P. (1981). *Vocabulary knowledge*. Newark: International Reading Association.
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107*(3), 251-271. doi:http://dx.doi.org/10.1086/511706
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. New York: The Guilford Press.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*(1), 44-62.  
doi:http://dx.doi.org/10.1037/0022-0663.98.1.44
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition, 126*(1), 39-53.  
doi:http://dx.doi.org/10.1016/j.cognition.2012.08.008
- Blewitt, P., Rump, K. M., Shealy, S. E., & Cook, S. A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Education Psychology, 101*(2), 294-304. doi:http://dx.doi.org/10.1037/a0013844
- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes, 45*(2), 122-159. doi:10.1080/01638530701792826

- Booth, A. E., & Waxman, S. R. (2002). Word learning is 'smart': evidence that conceptual information affects preschoolers' extension of novel words. *Cognition*, 84(1), B11-B22. doi:10.1016/s0010-0277(02)00015-x
- Brownell, R. (2010). Expressive One-Word Picture Vocabulary Test, Fourth Edition (EOWPVT): Pearson.
- Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.
- Cartmill, E., Pruden, S. M., Levine, S. C., & Goldin-Meadow, S. (2010). The role of parent gesture in children's spatial language development. In K. Franich, K. M. Iserman, & L. L. Keil (Eds.), *Proceedings of the 34th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Clark, E. V. (2010). *First language acquisition* (2nd ed.). Cambridge: Cambridge University Press.
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, 110(1), 1-18. doi:10.1086/598840
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33(6), 934-945. doi:http://dx.doi.org/10.1037/0012-1649.33.6.934
- Dickinson, D. K., & Brady, J. P. (2005). Toward effective support for language and literacy through professional development. In M. Zaslow & I. Martinez-Beck (Eds.), *Critical issues in early childhood professional development* (pp. 141-170). Baltimore, MD, US: Paul H Brookes Publishing.

- Dickinson, D. K., Golinkoff, R. M., & Hirsh-Pasek, K. (2010). Speaking out for language: Why language is central to reading development. *Educational Researcher*, 39(4), 305-310.
- Dickinson, D. K., Nesbitt, K. T., Collins, M. F., Hadley, E. B., Newman, K. M., Rivera, B. L., . . . Golinkoff, R. (2015). Effects of intensity of instruction of different word types during preschool book reading considering initial vocabulary
- Dickinson, D. K., & Porche, M. V. (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. doi:10.1111/j.1467-8624.2011.01576.x
- Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test, IV.*: Circle Pines, MN: American Guidance Service, Inc.
- Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science*, 18(3), 254-260. doi:10.1111/j.1467-9280.2007.01885.x
- Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: differences by class and race. *Social Science Research*, 33(3), 464-497.  
doi:http://dx.doi.org/10.1016/j.ssresearch.2003.08.001
- Fisher, A. V., Godwin, K. E., Matlen, B. J., & Unger, L. (2015). Development of category-based induction and semantic knowledge. *Child Development*, 86(1), 48-62.  
doi:10.1111/cdev.12277
- Forbes, J. N., & Farrar, M. J. (1993). Children's initial assumptions about the meaning of novel motion verbs: Biased and conservative? *Cognitive Development*, 8(3), 273-290.  
doi:http://dx.doi.org/10.1016/S0885-2014(93)80002-B



- Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2011). Predicting robust vocabulary growth from measures of incremental learning. *Scientific Studies of Reading, 15*(1), 71-91.
- Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Cambridge: Oxford University Press.
- Glenberg, A. M., Gutierrez, T., Levin, J. R., Japuntich, S., & Kaschak, M. P. (2004). Activity and imagined activity can enhance young children's reading comprehension. *Journal of Educational Psychology, 96*(3), 424-436. doi:<http://dx.doi.org/10.1037/0022-0663.96.3.424>
- Goodrich, W., & Kam, C. L. H. (2009). Co-speech gesture as input in verb learning. *Developmental Science, 12*(1), 81-87. doi:[10.1111/j.1467-7687.2008.00735.x](https://doi.org/10.1111/j.1467-7687.2008.00735.x)
- Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly, 51*(2), 181-198. doi:[10.1002/rrq.130](https://doi.org/10.1002/rrq.130)
- Hammill, D., & Newcomer, P. (2008). Test of Language Development-Primary, fourth edition (TOLD-P:4): Pro-Ed.
- Hargrave, A. C., & Sénéchal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: the benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15*(1), 75-90. doi:[10.1016/s0885-2006\(99\)00038-1](https://doi.org/10.1016/s0885-2006(99)00038-1)
- Harris, J., Golinkoff, R. M., & Hirsh-Pasek, K. (2011). Lessons from the crib for the classroom: How children really learn vocabulary. In S. B. Neuman & D. K. Dickinson (Eds.), *Handbook of Early Literacy Research* (pp. 49-66). New York: Guilford Press.

- Hindman, A. H., & Wasik, B. A. (2013). Vocabulary learning in Head Start: Nature and extent of classroom instruction and its contributions to children's learning. *Journal of School Psychology, 51*(3), 387-405. doi:http://dx.doi.org/10.1016/j.jsp.2013.01.001
- Hoffman, J. L., Teale, W. H., & Paciga, K. A. (2014). Assessing vocabulary learning in early childhood. *Journal of Early Childhood Literacy, 14*(4), 482-509.  
doi:10.1177/1468798413501184
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). II. The Emergentist Coalition Model. *Monographs of the Society for Research in Child Development, 65*(3), 17-29.  
doi:10.1111/1540-5834.00092
- Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*(2), 128-157. doi:10.1080/15250000701795598
- Huaqing Qi, C., Kaiser, A. P., Milan, S., & Hancock, T. (2006). Language performance of low-income African-American and European American preschool children on the PPVT-III. *Language, Speech & Hearing Services in Schools, 37*(1), 5-16. doi:10.1044/0161-1461(2006/002)
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.
- Kintsch, W. (2005). An overview of top-down and bottom-up effects in comprehension: The CI perspective. *Discourse Processes, 39*(2-3), 125-128.  
doi:10.1080/0163853X.2005.9651676
- Kintsch, W., & Mangalath, P. (2011). The construction of meaning. *Topics in Cognitive Science, 3*(2), 346-370.

- Kousta, S.-T., Vigliocco, G., Vinson, D. P., Andrews, M., & Del Campo, E. (2011). The representation of abstract words: why emotion matters. *Journal of Experimental Psychology: General*, 140(1), 14.
- Lahey, M. (1988). *Language disorders and language development*. Needham, MA: Macmillan.
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(01), 69-92.  
doi:doi:10.1017/S0142716410000299
- Leung, C. B. (2008). Preschoolers' acquisition of scientific vocabulary through repeated read-aloud events, retellings, and hands-on science activities *Reading Psychology*, 29, 28.  
doi:http://dx.doi.org/10.1080/02702710801964090
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300-335.
- McKeown, M. G. (1993). Creating effective definitions for young word learners. *Reading Research Quarterly*, 28(1), 17-31. doi:10.2307/747814
- McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly*, 29(4), 520-530.
- Melka, F. (1997). Receptive versus productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary : description, acquisition and pedagogy* (pp. 84-102). Cambridge: Cambridge University Press.
- Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early

- reading ability. In J. L. M. L. C. Ehri (Ed.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Miller, G. A., & Fellbaum, C. (1991). Semantic networks of English. *Cognition*, *41*(1-3), 197-229. doi:[http://dx.doi.org/10.1016/0010-0277\(91\)90036-4](http://dx.doi.org/10.1016/0010-0277(91)90036-4)
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3, pp. 269-284). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nation, K. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, K., & Snowling, M. J. (1998). Semantic processing and the development of word-recognition skills: Evidence from children with reading comprehension difficulties. *Journal of Memory and Language*, *39*(1), 85-101.
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Retrieved from Washington, DC: Available at <http://www.nifl.gov/earlychildhood/NELP/NELPreport.html>
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Retrieved from Washington, D.C.:
- Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly*, *46*(3), 249-272.

- Ordóñez, C. L., Carlo, M. S., Snow, C. E., & McLaughlin, B. (2002). Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology, 94*(4), 719-728. doi:<http://dx.doi.org/10.1037/0022-0663.94.4.719>
- Ouellette, G. P. (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. doi:<http://dx.doi.org/10.1037/0022-0663.98.3.554>
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly, 42*(2), 282-296.
- 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. doi:<http://dx.doi.org/10.1037/0022-0663.94.1.23>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi:<http://dx.doi.org/10.1080/10888430701530730>
- Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., . . . Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children, 77*(2), 161-183.
- Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and Writing, 25*(7), 1635-1664. doi:<http://dx.doi.org/10.1007/s11145-011-9336-5>
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a Second Language:*

- Selection, Acquisition, and Testing* (pp. 209-227). Philadelphia, PA: John Benjamins Publishing Company.
- Richard, G. J., & Hanner, M. A. (1995). *Language Processing Test - Revised (LPT-R)*. Austin, TX: LinguiSystems.
- Richter, T., Isberner, M.-B., Naumann, J., & Neeb, Y. (2013). Lexical Quality and Reading Comprehension in Primary School Children. *Scientific Studies of Reading, 17*(6), 415-434. doi:10.1080/10888438.2013.764879
- Roskos, K., Ergul, C., Bryan, T., Burstein, K., Christie, J., & Han, M. (2008). Who's learning what words and how fast? Preschoolers' vocabulary growth in an early literacy program. *Journal of Research in Childhood Education, 22*(3), 275-290. doi:10.1080/02568540809594627
- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research, 95*(5), 259-272. doi:10.1080/00220670209596600
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 1-39*. doi:10.1111/lang.12077
- Schoonen, R., & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing, 25*(2), 211-236. doi:10.1177/0265532207086782
- Semel, E., Wiig, E. H., & Secord, W. (2003). *Clinical Evaluation of Language Fundamentals - Fourth Edition*: Pearson.

- Senechal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of Child Language*, 24(1), 123-138.  
doi:<http://dx.doi.org/10.1017/S0305000996003005>
- Senechal, M., Ouellette, G., & Rodney, D. (2006). The misunderstood giant: On the predictive role of early vocabulary to future reading. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of Early Literacy Research* (Vol. 2, pp. 173-182). New York: Guilford Press.
- Seston, R., Golinkoff, R. M., Ma, W., & Hirsh-Pasek, K. (2009). Vacuuming with my mouth?: Children's ability to comprehend novel extensions of familiar verbs. *Cognitive Development*, 24(2), 113-124. doi:10.1016/j.cogdev.2008.12.001
- Seymour, H. N., Roper, T. W., & de Villiers, J. (2005). *Diagnostic Evaluation of Language Variation--Norm Referenced (DELV)*: Pearson.
- Silverman, R. D., & Crandell, J. D. (2010). Vocabulary practices in prekindergarten and kindergarten classrooms. *Reading Research Quarterly*, 45(3), 318-340.  
doi:<http://dx.doi.org/10.1598/RRQ.45.3.3>
- Silverman, R. D., & Hartranft, A. M. (2015). *Developing vocabulary and oral language in young children*. New York: Guilford Press.
- Silverman, R. D., Proctor, C. P., Harring, J. R., Doyle, B., Mitchell, M. A., & Meyer, A. G. (2014). Teachers' instruction and students' vocabulary and comprehension: An exploratory study with English monolingual and Spanish–English bilingual students in grades 3–5. *Reading Research Quarterly*, 49(1), 31-60. doi:10.1002/rrq.63
- Snow, C. E., Cancino, H., De Temple, J., & Schley, S. (1991). Giving formal definitions: A linguistic or metalinguistic skill? In E. Bialystock (Ed.), *Language Processing in Bilingual Children* (pp. 90-112). Cambridge: Cambridge University Press.

- Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology, 43*(2), 454-464. doi:10.1037/0012-1649.43.2.454
- Swingley, D. (2010). Fast mapping and slow mapping in children's word learning. *Language Learning and Development, 6*(3), 179-183.
- Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing, 16*(1), 5-20. doi:10.1023/a:1021789804977
- Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Education Psychology, 93*(2), 243-250.
- Whitehurst, G. J., Arnold, D. S., Epstein, J. N., Angell, A. L., Smith, M., & Fischel, J. E. (1994). A picture book reading intervention in day care and home for children from low-income families. *Developmental Psychology, 30*(5), 679-689. doi:http://dx.doi.org/10.1037/0012-1649.30.5.679
- Wiig, E. H., & Secord, W. (1992). *Test of Word Knowledge*: Psychological Corporation.
- Williams, K. (2007). Expressive Vocabulary Test, second edition (EVT): Pearson.
- Woodcock, R., McGrew, K. S., & Mather, N. (2001). Woodcock Johnson III test of cognitive abilities. Rolling Meadows, IL: Riverside Publishing.
- Yu, C., & Smith, L. B. (2012). Modeling cross-situational word-referent learning: Prior questions. *Psychological Review, 119*(1), 21-39. doi:http://dx.doi.org/10.1037/a0026182
- Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts About verbs by listening. *Psychological Science, 20*(5), 619-626. doi:10.1111/j.1467-9280.2009.02341.x



Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review*, 21(1), 1-22.

doi:10.3758/s13423-013-0443-y

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). *Preschool Language Scale*, fourth edition (PLS-4): Pearson.

## CHAPTER IV

### BUILDING SEMANTIC NETWORKS: THE IMPACT OF A VOCABULARY INTERVENTION ON PRESCHOOLERS' DEPTH OF WORD KNOWLEDGE

To make meaning from text, children draw on a wealth of accumulated knowledge about words and the concepts they signify. Comprehension requires not just that children have broad vocabularies (i.e., a large number of words in their lexicon), but also that those words signal rich, interconnected networks of conceptual knowledge (Anderson & Freebody, 1985; Kintsch, 1998; Perfetti, 2007). These networks of word knowledge, often referred to as vocabulary depth, have been shown to play a unique and particularly powerful role in supporting children's ability to understand what they read (Ouellette, 2006; Roth, Speece, & Cooper, 2002; Tannenbaum, Torgesen, & Wagner, 2006). The National Early Literacy Panel (2008) found that children's abilities to give definitions for words (a measure of depth) was a significantly stronger predictor of later decoding and reading comprehension than receptive vocabulary (a measure of breadth). Moreover, vocabulary depth has been shown to predict reading comprehension above and beyond the association explained by breadth (Ouellette, 2006). Unlike fast-mapped, shallow knowledge about words, deep word knowledge slowly accumulates over time (Bloom, 2002; Bolger, Balass, Landen, & Perfetti, 2008), and intentional efforts at fostering this knowledge in classrooms should begin early. However, the available literature on supporting depth of knowledge in early childhood learners is sparse, with limited information about which features of instruction might support the building of semantic networks. The present study examines the impact of a vocabulary intervention designed to support depth in preschool children. I also investigate the effect of specific instructional strategies on depth, namely teaching words across multiple contexts and in conceptually-related categories. Finally, I investigate additional factors that may contribute to depth by examining the relationship between children's growth in word

knowledge and 1) types of adult-child interaction, and 2) the kinds of support provided for new words.

### **Theoretical Framework**

The term vocabulary depth refers to the quality, rather than the quantity, of words known (Anderson & Freebody, 1985). Perfetti's Lexical Quality Hypothesis (LQH, 2007) supplies theoretical framing to this construct, defining depth as the quality of a lexical representation for an individual word, or how much knowledge a learner has about its form and meaning. When considering oral language, form refers to the grammatical and phonological representation of a word, and meaning to its semantic representation. Lexical quality is built incrementally over time, with each new encounter with a word encoding a trace representation in one's memory (Bolger et al., 2008). As encounters accumulate, the representation's quality increases and it is accessed more quickly. Low-quality representations, consisting of minimal phonological (Estes, Evans, Alibali, & Saffran, 2007; Swingley, 2007) or syntactic (Yuan & Fisher, 2009) representations of words with little to no semantic information, are retrieved slowly, impairing comprehension. Higher quality representations, consisting of a stable phonological representation and a generalized, flexible sense of meaning, are retrieved quickly, facilitating a more effortless comprehension process. The key indicator of high quality lexical representations, then, is the ability to quickly retrieve (i.e., remember or use) a word.

The LQH defines depth primarily in terms of the lexical quality for individual words (i.e., at the micro-level). It can be extended by adopting a macro-level perspective that considers how individual lexical representations are connected to one another to form semantic networks. Language researchers have applied network science, an approach that draws on graph theory to examine complex systems such as social networks and the internet (Börner, Sanyal, &

Vespignani, 2007), to the problem of how knowledge is organized in the mental lexicon. Using tools from network science, word knowledge can be modeled as semantic networks in which words are represented as nodes and semantic relationships as connections between those nodes (Hills, Maouene, Maouene, Sheya, & Smith, 2009; Steyvers & Tenenbaum, 2005). These semantic networks have a “small world” structure, meaning that there is a relatively small distance between any two words, and words tend to form clusters more than would be expected by chance (Steyvers & Tenenbaum, 2005). Semantic networks are also “scale-free,” meaning that only a small number of words are highly connected to other words, with most words having only a few connections (Steyvers & Tenenbaum, 2005). These structural properties are believed to support efficient language processing and word retrieval (Borovsky, Ellis, Evans, & Elman, 2016b; Griffiths, Steyvers, & Firl, 2007; Steyvers & Tenenbaum, 2005; Vitevitch, 2008).

The small-world, scale-free structure of semantic networks likely emerges as children’s vocabularies grow, with reorganization of networks occurring as new words are added. Semantic networks expand through the principle of preferential attachment: when new words are added to the semantic network, they are more likely to connect to words that are already highly connected (Steyvers & Tenenbaum, 2005), creating the characteristic scale-free structure. The principle of preferential attachment has important consequences for theories of word-learning: it suggests that new words are added to the semantic network by further differentiating existing networks. Furthermore, it implies that children may be more likely to learn new words that are semantically related to known words than those that are unrelated. That is, when children encounter a variety of new words in their environment, they may be more likely to acquire and retain the words that have ready-made semantic “hooks” to existing networks. Borovsky and colleagues tested this hypothesis with two-year-olds, finding that children were more accurate at recognizing novel

words when they knew more about the category to which the words belonged, as opposed to words with low category knowledge (Borovsky, Ellis, Evans, & Elman, 2016a). These findings indicate that dense semantic neighborhoods may help to leverage word learning because of the knowledge children already have about semantically similar words in the network.

Growth in depth can therefore be considered not only as an increase in the quality of individual lexical representations (Perfetti, 2007), but also the increased semantic differentiation and reorganization of semantic networks that occur as new words are added to the lexicon (Steyvers & Tenenbaum, 2005). The LQH suggests that depth may be fostered through repeated encounters with words that help children learn information about the meaning, form, and use of words (Perfetti, 2007). Semantic network theory indicates that depth can also be supported by building networks of conceptually-linked knowledge so that new, semantically similar words can be acquired more readily (Borovsky et al., 2016a).

### **Factors that Support Depth of Knowledge**

#### **Repeated Encounters with Words**

Children are able to glean some information about a word's form, meaning, and use from only a single exposure. To do so, they draw on a wide range of cues in their environment, including gestures, intonation, eye gaze, perceptual features, and syntactic information (Hollich, Hirsh-Pasek, & Golinkoff, 2000). A single initial encounter with a word can result in a fast-mapped lexical representation, consisting of minimal phonological (Estes et al., 2007; Swingley, 2007) or syntactic (Yuan & Fisher, 2009) information with little to no semantic information. The quality of the lexical representation increases with each additional encounter with a word (Frishkoff, Perfetti, & Collins-Thompson, 2011). This process can be quite protracted, with six and eight-year-old children showing difficulty in understanding common verbs (*write, scrub*)

when extended to new contexts (Seston, Golinkoff, Ma, & Hirsh-Pasek, 2009), signaling relatively low lexical quality.

Encountering words in settings that provide cues for meaning can expedite the word-learning process (Frishkoff et al., 2011). In a study where K-2<sup>nd</sup> graders heard a book read four times but were not given any extra-textual information about words, children showed gains of 15% of word meanings simply from hearing them used repeatedly in the book (Biemiller & Boote, 2006). Multiple readings of a book provides rich information about new words, as these words are: 1) embedded in engaging narratives or informational text that allow for inferencing about word meaning, 2) used in syntactic frames that help children determine aspects of word meaning and encounter typical contexts of use, and 3) supported visually through illustrations or adult gestures that support knowledge of a word's perceptual features (Dickinson, Griffith, Golinkoff, & Hirsh-Pasek, 2012).

Building depth through repeated encounters with words is most effective if those encounters are planned purposefully (e.g., repeated readings of a book). Relying solely on incidental exposures for word-learning is unlikely to give children sufficient experience with low-frequency words. It may be also difficult for children to build depth for less imageable, more abstract words without explicit information about word meaning (Hadley, Dickinson, Hirsh-Pasek, Golinkoff, & Nesbitt, 2016).

### **Explicit Information About Word Meanings**

Providing explicit information about the meaning of words has been shown to support depth of knowledge beyond the contribution of repeated encounters alone (Bolger et al., 2008). K-2<sup>nd</sup> grade children gained 15% of target word meanings from repeated readings of a book, but gains increased to 22% when brief definitions of words were included (Biemiller & Boote,

2006). Similarly, preschoolers had significantly greater depth in knowledge ( $d = 0.41$ ) for words taught with definitions vs. words simply heard during repeated readings of books (Dickinson et al., 2016). Providing explicit information about word meaning highlights the word in question so that children attend to it more carefully. It also helps children to construct higher-quality lexical representations by supplying them with explicit semantic information, examples of use in context, and a strong phonological model to imitate.

Combining both factors, repeated encounters with words and explicit information about meaning, appears to be especially powerful in building depth. In the Biemiller & Boote study (2006), word meaning gains increased from 22% to 41% when an additional review of words and their meanings was added. In another study, K-1<sup>st</sup> graders learned significantly more about words that were taught with six days of intensive instruction vs. three (Beck & McKeown, 2007). These findings argue for instructional support for depth that includes: 1) repeated encounters with words, 2) explicit information about meaning, and 3) the use of a supportive context that provides multiple cues about new words' form, meaning, and use.

### **Integrating Knowledge Across Contexts**

The key indicator of high-quality lexical knowledge is the ability to quickly retrieve a word identity and apply it flexibly across a range of contexts (Perfetti, 2007). (Note that “context” is used here in a linguistic sense, to mean the words or phrases surrounding the word in question). This is a difficult task for young children, who often limit a word's use to the context it was learned in (e.g., using *healthy* only to refer to food, rather than a person's physical condition) (Beck, McKeown, & Kucan, 2013). One approach to supporting this aspect of depth is to build in encounters with words in multiple contexts. However, there are mixed findings about the efficacy of this approach: some research indicates that children learn more about new words,

particularly verbs and adjectives, when they are presented consistently in a single context (Goldberg, Casenhiser, & Sethuraman, 2004). For example, preschool children who learned new words during the same book read three times learned and retained significantly more words than those who learned the same new words during three different books (Horst, Parsons, & Bryan, 2011).

Other research has found that diverse contexts are helpful for word-learning across a range of ages, with adults learning words more quickly and deeply when they were presented in varied sentence contexts (Bolger et al., 2008), and 5 to 7-year-olds fast-mapping more words when they were presented in more diverse contexts (Suanda, Mugwanya, & Namy, 2014). Hills et al. (2010) also found that the contextual diversity of words, or the number of unique word types that co-occurred with that word in adult speech, predicted how early they were acquired by young children. In an attempt to reconcile these findings, I hypothesize that in an instructional setting, it may be beneficial for children to learn words first in a single, supportive context to establish at least fast-mapped knowledge, then push for generalization of knowledge through exposure to additional contexts.

Several vocabulary interventions with kindergartners have taken this approach, teaching target words first in a single context through book-reading, then engaging in discussion about words that incorporate a variety of contexts (e.g., asking questions such as, “What are some things you might be *reluctant* to do?”) after reading (Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; McKeown & Beck, 2014). McKeown & Beck (2014) compared depth of knowledge for kindergartners who learned words through repeated readings of a book, i.e., a single context (“repetition condition”), with those who heard the book read only once, then engaged in discussion about target words using a variety of contexts (“interactive condition”). The children



in the interactive condition had significantly higher scores on measures that tapped high-quality word knowledge. These results suggest that helping children to encounter words in multiple contexts may support depth more than restricting encounters to a single context, although the differences shown could also be due to the increased child engagement in the interactive condition.

In the present study, I taught one set of words in a single unit, through book-reading and play focused on either *vegetables* or *flowers*, and a second set of words in both units (*vegetables and flowers*). By giving children substantial experience with words in one unit before teaching them again in the second unit, children could theoretically establish a strong initial representation for words before integrating that knowledge into a new and different context. I contrasted the difference in learning between the words taught in one unit vs. two, controlling for exposure, to isolate the effect of learning words in more than one context on children's growth in vocabulary knowledge.

### **Relationships Between Words**

Semantic network theory suggests that supporting children's knowledge of the semantic relationships between words may also foster depth, as new words are thought to hook into the network more readily when they are semantically related to known words (Borovsky et al., 2016a; Steyvers & Tenenbaum, 2005). One approach to doing so, pioneered by Neuman and colleagues (Neuman, Newman, & Dwyer, 2011), is to teach words in conceptually-related categories, using words' common features to build semantic networks more efficiently than would otherwise be possible.

The practice of teaching words in categories draws on language research that suggests that improving the quality of word knowledge involves not only adding more information about

individual concepts, but also changes in how concepts are organized (A. V. Fisher, Godwin, Matlen, & Unger, 2015). While even infants show a limited ability to categorize (Delle Luche, Durrant, Floccia, & Plunkett, 2014), the ability to more finely differentiate these categories, and group categories into nested hierarchies, continues to develop as children gain more knowledge about the world around them (A. V. Fisher et al., 2015; Hills et al., 2009). In particular, there appears to be a developmental shift from categorizing based on thematic to taxonomic relations (Cronin, 2002; A. V. Fisher et al., 2015; Markman, 1989).

**Thematically-related words.** Thematically-related words are those that are involved in the same event (rain and umbrella) or are spatially or causally related (boy and baseball). Thematically related words do not share inherent characteristics and are not things of the same type (Markman, 1989). When children learn about concepts in thematic groups, they gain an understanding of semantic relationships between words such as causal or spatial relationships (Markman, 1989). Many early childhood curricula capitalize on the learning possibilities of thematic categories by organizing instruction around themes. For example, a “farm, markets, & food” theme (Shine Early Learning, 2016) involves instruction about growing, purchasing, and cooking food, thereby building a rich semantic network of words that co-occur in the same context.

Previous studies have designed vocabulary instruction around words grouped in thematic categories (Pollard-Durodola et al., 2011; Wasik & Bond, 2001). Pollard-Durodola and colleagues’ (2011) Words of Oral Reading and Language Development (WORLD) intervention presented new words in thematic groups so that children could make connections between concepts and build more extended semantic networks. Researchers chose four narrative and informational texts for a water theme (e.g., *The Snowy Day*, *Amazing Water*), then selected

“lexical sets” of thematically related words such as *raindrop*, *liquid*, *frozen*, and *drain* for instruction. Instruction before, during, and after reading focused on supporting word meaning, and a weekly review sought to integrate knowledge across books. Children in the WORLD intervention condition showed significantly greater growth in vocabulary depth than those in the control condition.

**Taxonomically-related words.** Words in taxonomies are hierarchically related, organized in a nested structure so that each higher-order category is increasingly general. Taxonomies allow for inference-making, so that an animal with five digits can be categorized as a primate, which in turn supports inductions that are not perceptually available (e.g., that the animal likely sees in color and is warm-blooded). Taxonomic knowledge therefore provides a short-cut for acquiring information about the world. There is evidence that taxonomic organization and semantic knowledge are reciprocally related, with deeper knowledge supporting children’s ability to categorize, and more differentiated taxonomic organization in turn leveraging children’s word-learning (A. V. Fisher et al., 2015; Kaefer & Neuman, 2013). Using taxonomies also exhibits features of academic language such as organizing information into a hierarchical structure (Snow & Uccelli, 2009), and is central to academic discourse in disciplines such as science and social studies (Richardson Bruna, Vann, & Perales Escudero, 2007; Wignell, Martin, & Eggins, 1989). Gaining proficiency with this form of conceptual organization, then, can help support students’ ability to acquire and communicate knowledge using the language of schooling (Schleppegrell, 2012).

Neuman and colleagues’ World of Words (WOW) intervention (Neuman & Dwyer, 2011; Neuman et al., 2011; Neuman, Pinkham, & Kaefer, 2015) was the first to teach words in taxonomies as a way of promoting vocabulary growth, while also including additional supports

for conceptual knowledge such as multimedia, informational texts, and writing activities (Neuman et al., 2011). 3 and 4-year-olds children in the WOW intervention condition learned significantly more words than control children and could use their knowledge of categories to identify new words (Neuman & Dwyer, 2011). Growth in both vocabulary and category knowledge was sustained at a 6-month posttest.

**Thematically vs. taxonomically-related words.** While interventions such as WOW and WORLD indicate the value of teaching words in both thematically- and taxonomically-related groups, teaching words in taxonomies may be of particular advantage because of its efficient nature – i.e., knowing something about a category can be applied to all the exemplars in that category, thereby supporting depth without an undue investment of instructional time (Kaefer & Neuman, 2013). To our knowledge, however, an intervention study has not examined the relative benefits of teaching words to preschoolers in taxonomic vs. thematic categories. The present study adds to the research on teaching words in categories by explicitly contrasting preschool children’s learning of thematically vs. taxonomically-related words.

### **Adult-Child Interactions**

Young children learn information about the form, meaning, and use of words through verbal interactions with more skilled language partners. The features described above – repeated encounters with words, explicit information about words, integrating knowledge across contexts, and explaining relationships between words – are all instantiated in the context of these adult-child interactions, the different forms of which can significantly influence children’s vocabulary development (Heath, 1982; Hoff, 2006). I focus here on three main types of interactions: 1) *instructional*, 2) *responsive*, and 3) *active processing*, that have been shown to be helpful for

children's language development and support vocabulary knowledge, but differ in their relative emphases on child participation and engagement.

*Instructional* talk is initiated by an adult with the primary intention of giving information about target words or instructional topics. The term instructional talk encompasses explicit instructional strategies such as providing definitions for words, giving examples of usage, and supplying conceptual information, and has shown to be a predictor of language growth in preschoolers (McCartney, 1984). Instructional talk has many of the benefits reviewed in the section on explicit information about word meaning above, but is primarily child-directed rather than interactional, meaning that children may not be fully interested or engaged. Instructional conversations can have a negative effect on children's language growth (Dickinson, 2001; Gámez & Lesaux, 2015) if not appropriately tailored to children's language level, if they preclude opportunities for children to use language themselves, or if instructional talk during shared book-reading sessions detracts from children's understanding of the story.

*Responsive* interactions are those in which a child signals an interest or need and an adult responds in a way that recognizes and/or extends the child's offering (Landry et al., 2012). The use of responsive language strategies is predicated on the theory that language develops from a foundation of joint attention and engagement, where both language partners are attending to the same topic or object (Tomasello & Farrar, 1986). Specific responsive strategies include modeling language use, restating and expanding children's utterances, and noticing and responding to children's verbal and nonverbal communication cues (Kaiser, Neitfeld, & Roberts, 2010). Preschool teachers' use of responsive interaction strategies such as following children's lead in conversations was associated with vocabulary growth in preschoolers (Cabell et al., 2011), and children in daycare centers whose caregivers engaged in more responsive interactions talked

more and in more complex utterances than their peers (Girolametto, Weitzman, & Greenberg, 2003). Like instructional interactions, responsive interactions are a source of rich information about words: children hear their own use of target vocabulary words echoed by adults, but used with an expanded meaning and different syntactic structure. Responsive interactions are particularly powerful for word-learning because they are initiated by children, who are therefore actively thinking about the concept at hand and receptive to adult input. Responsive interactions also give children the opportunity to put their emerging word knowledge into action, testing out how a new word “works,” building syntactic and phonological representations, and receiving feedback and scaffolding from an adult.

*Active processing* is a term used by McKeown & Beck (2014) to refer to learners’ hands-on experiences analyzing the semantic meaning and connotations of words. During *active processing* interactions, adults prompt learners to interact with and analyze word meanings. In other words, children are asked to “do things” with words (Nagy & Scott, 2000), such as generate examples and non-examples of use or whether a word fits in a particular context (McKeown & Beck, 2014). These interactions have been shown to increase vocabulary depth in kindergartners (McKeown & Beck, 2014) and were associated with significant gains in word knowledge in a meta-analysis of vocabulary instruction (Stahl & Fairbanks, 1986). Active processing interactions ask children to quickly retrieve their emerging lexical representations for words and apply them in a novel context. Active processing interactions are initiated by adults and are the most cognitively demanding of the three interaction types in that they ask children to generalize their word knowledge with little adult scaffolding. Moreover, active processing interactions supply minimal information about word form, meaning, or use, unless the adult explicitly provides this information in follow-up responses to children.

### **Activity Settings: Book-reading and Play**

Another promising approach for supporting depth is to teach words in more than one activity setting during the school day (i.e., Wasik & Bond, 2001), as this approach builds in frequent encounters with words and allows for different types of teacher-child interactions. Carefully choosing targeted settings for vocabulary instruction is important, though, as efforts that have sought to more generally increase sophisticated vocabulary use across the day have had limited success (Dickinson, 2011). The most successful vocabulary interventions have taken place in discrete, language-rich settings that require relatively small modifications to teacher's existing practice, such as shared book-reading. There is a wealth of research on preschool shared book-reading interventions aimed at supporting oral language development (Mol, Bus, & de Jong, 2009), and meta-analyses estimate that these interventions have a moderate effect on vocabulary knowledge ( $d = 0.60$ , National Early Literacy Panel, 2008). However, there is some concern that interventions designed to support vocabulary learning must become more potent to build the deep word knowledge important for later reading comprehension (Beck & McKeown, 2007; Neuman et al., 2011; Roskos & Burstein, 2011).

One approach to boosting the impact of book-reading is to add play (or playful activities) to book-reading sessions (Hadley et al., 2016; Roskos & Burstein, 2011; Weisberg et al., 2015). Like book-reading, play is an established and developmentally appropriate part of the early childhood curriculum, and longitudinal research has found that preschoolers whose teachers extended talk on topics, used cognitively challenging talk, and allowed them to talk more during play had higher scores on kindergarten language assessments (Dickinson, 2001). A newer line of research has explored the learning possibilities of guided play, a method in which teachers play with children while scaffolding them towards specific learning aims such as learning new words

(Hirsh-Pasek & Golinkoff, 2011). The small number of vocabulary interventions that have employed guided play methods have found significant growth in vocabulary knowledge (Christakis, Zimmerman, & Garrison, 2007; K. R. Fisher, Hirsh-Pasek, Newcombe, & Golinkoff, 2013; Han, Moore, Vukelich, & Buell, 2010; Weisberg et al., 2015).

There are very few studies on the benefits of combining book-reading and play, but an observational study found that preschoolers naturally bring ideas, characters, and even dialogue from books into their free play, and this play enriches and clarifies children's understanding of books (Rowe, 1998). This suggests that book-reading and play can have a complementary relationship wherein concepts learned during book-reading can be further explored and assimilated during play.

Pairing book-reading and play shows promise for fostering depth of knowledge, as combining these activity settings builds in repeated encounters with words, rich semantic information about form, meaning, and use, and opportunities to interact with words. Shared book-reading sessions can serve as a foundation for later play, as children gain a fast-mapped understanding of new words and a shared narrative from the book to draw on as play fodder. Guided play can further deepen semantic representations as new words are indexed to play props (e.g., using a small chair toy to learn *throne*) or illustrated through play characters' actions and feelings. Play also serves as an opportunity for children to interact with and integrate different contexts of use (Bolger et al., 2008; McKeown & Beck, 2014).

These activity settings also emphasize different, complementary emphases on the types of adult-child interactions shown to be supportive of language learning. Book-reading is the most common activity setting for instructional talk in preschool classrooms (Gest, 2006), and provides an ideal context for supplying information about words over multiple readings. Once children



have gained initial representations of new words, active processing interactions can be used to probe nuances of meaning during post-book-reading discussions (Coyne et al., 2009; McKeown & Beck, 2014). Book-reading is a less common site for responsive interactions, as reading a text in classroom settings is typically an adult-led activity. Guided play complements the more teacher-led nature of book-reading, as responsive interactions occur more often during play than in other early childhood activity settings (Girolametto & Weitzman, 2002), including book-reading. Active processing interactions can also be incorporated into guided play, such as asking open-ended questions about words' meaning (Toub et al., 2017).

### **Present Study**

The primary goal of the present study was to examine the effects of an intervention designed to support preschool children's depth of vocabulary knowledge, as well as to examine specific features of instruction that supported depth. Based on the review of research above, I made five hypotheses about growth in word knowledge:

- 1) children will show significant growth in their knowledge of target words on breadth and depth measures, as compared to their knowledge of exposure and control words;
- 2) children will show significantly greater increases in knowledge for words taught in two units vs. one;
- 3) children will show significantly greater increases in knowledge for taxonomically-related vs. thematically-related words;
- 4) the frequency of instructional, responsive, and active processing interactions will all be significant predictors of growth in word knowledge, with differential patterns across book-reading and play activity settings. Specifically, I predict:

- a. the frequency of instructional interactions during book-reading will have the strongest positive relationship with vocabulary growth in this setting, given that definitional information has been shown to promote growth in word-learning during book-reading;
  - b. the frequency of responsive and active processing interactions during play will have the strongest positive relationship with vocabulary growth in this setting because children will be able to draw on and further refine their lexical representations through these hands-on experiences with words,
- 5) the frequency of definitions, target word use in the book text, and provision of visual supports will all be significant predictors of growth in word knowledge.

Further, I hypothesize that for #2-5 above, differences in learning will be more apparent on the more sensitive vocabulary depth measure, which is able to capture small increments of word knowledge, as opposed to the vocabulary breadth measure.

## **Methods**

### **Research Participants**

Data come from 30 children enrolled in three preschool classrooms from a state-funded program for low-income families in a Southeastern U.S. city. The sample included only children who did not have an Individualized Education Plan and who understood enough English to be able to follow directions, as reported by their teacher. The average age for the children at pretest was 59.6 months ( $SD = 3.1$  months). The sample was approximately 43% male, and based on teacher report, 76.7% percent of the sample children were African-American, 6.7% Hispanic, 10% Caucasian, and 6.6% were designated biracial or of another ethnicity. Within each classroom, children were randomly assigned to a playgroup of three children. Playgroups were

mixed-gender: three playgroups were predominately male and seven predominately female. Children remained in the same playgroup for the duration of the intervention. I delivered the intervention to children, and am an experienced classroom teacher and trained educational researcher.

### **Materials: Book and Word Selection**

Two commercially available information texts were chosen that were well-suited for building knowledge about flowers (*Planting a Rainbow* by Lois Ehlert) and vegetables (*Vegetables in the Garden* by Pascale de Bourgoing and Gallimard Jeunesse). These books contained information about the plant-growing process as well as descriptions of different category members, such as types of vegetables or flowers. Half of the ten playgroups (50% of children), were randomly assigned to start with the flower book while the other half were assigned to start with the vegetable book.

Eight target words were selected for each book (16 total). These words included the category name (*vegetables* or *flowers*), five words for category members (e.g., *artichoke*, *tiger lily*), and two theme words that were conceptually, but not taxonomically, related to the category (e.g., *vines* for the vegetable book, *petals* for the flower book). Additionally, five target words (*stem*, *bulb*, *seeds*, *soil*, *roots*) were taught in both books. These words were intended to help children integrate the categories of vegetables and flowers into the larger category of “growing things,” therefore creating a more comprehensive semantic network. Three exposure words were selected for each book (6 total). These words appeared in the book but were not explicitly defined. Eight control words, equivalent in difficulty to the target and exposure words, were also selected. Approximately half of the target words, half of the exposure words, and 38% of the control words appeared on the Dale-Chall (1995) list of common words (known by 80% of

fourth-grade students), meaning that roughly half of words in each group were rare words and half were common words. When possible, we also categorized the words based on Biemiller's *Words Worth Teaching* (2010) list, with the majority of words falling in category E, words known by 80% of children by the end of 2<sup>nd</sup> grade. Most of the words were concrete nouns (with two exceptions: *sprouting* and *raw*). See Appendix C for a complete list of words, along with Biemiller (2010) and Dale-Chall (1995) ratings.

### **Procedures**

The intervention was conducted over a two-month period, from February to April 2013. The intervention was organized under the general theme of “growing things,” and included one book on vegetables and another on flowers. Activities based on each book lasted for four days. Mixed-gender playgroups of three children left their classroom to participate in intervention activities in a quiet space. During each of two weeks, children participated in four consecutive days of back-to-back book reading and play sessions, for a total of eight days of intervention activities. The book was read first, and then children engaged in 10 minutes of book-related, adult-guided play. Each book reading and play session lasted for approximately 20 minutes. All children were individually pretested and posttested by members of the research team for knowledge of vocabulary words within one week prior to and following the intervention, respectively.

**Book-reading.** I read each book aloud to children four times over the course of a week. Before each reading, the properties of each category were discussed using each theme word (*e.g.*, *stem*, *bulb*). Then children were shown pictures of various plants and other “growing things,” asked to decide whether the picture was, or was not, a category member, and explain their answer.

During the book-reading, each book word was explained when it occurred in the text. The explanation consisted of the following:

- Pointing at a corresponding illustration in the book to help support word meaning, and also showing a card that depicted a photograph of the word in order to support conceptual knowledge and ensure that the perceptual features of the object were clear (e.g., “These are *radishes*. Here’s another picture of some *radishes* growing in the ground.”)
- Definitional information delivered in concise, child-friendly language, including:
  - taxonomy membership (e.g., “*Radishes* are vegetables.”)
  - taxonomy non-membership (e.g., “*Radishes* don’t have seeds, so they’re not a fruit.”)
  - perceptual features (e.g., “*Radishes* are red on the outside and white on the inside. They are kind of spicy.”)
  - conceptual information (e.g., “*Radishes* are the *root* of the plant, so they grow underground.”)
  - functional information (e.g., “People usually eat *radishes* raw.”)

During the first and second readings, children were encouraged to repeat the word in order to reinforce the word’s phonological representation (e.g., “Can you say *radish*?”), and in the third and fourth readings, children were given a definition and asked to supply the word (e.g., “What is the vegetable that grows underground, and is red on the outside and white on the inside?”).

This extra-textual talk was listed on prompt cards that were used during reading to ensure children in different playgroups received similar information about words.

**Play.** A ten-minute play session immediately followed each book-reading. There was a collection of toys for each book with props related to target vocabulary. For the vegetables book,

this included a farmhouse, farmer figurines, small toy vegetables, seeds, as well as cooking implements and larger toy vegetables. For the flower book, the same farmhouse, farmer figurines, and seeds were used, but the collection also included a variety of toy plant beds, clay used to represent dirt, and gardening implements such as a watering can, hose, rake, and shovel.

During the first two days of play, I used an adult-directed method of play in which each child was each given 2-3 props, and I instructed children to enact key concepts from the book. For example, after the vegetable book, children were each given farmer figurines and instructed to act out planting seeds in the soil, watering the plants, harvesting and cooking the vegetables. This play also involved some sort of threat or conflict to foster a sense of playfulness and fun: animals coming to eat the plants, a tornado ruining the crop, or some other difficulty involving growing conditions. I used target vocabulary words in each “scene,” along with a definition. This adult-directed play was intended to serve as a model for children’s play, demonstrating ways to use the props and incorporate concepts from the book into their play.

During the second two days of play, a more child-led method of play was used in which the children initiated the play and I followed their lead, building on their play ideas and encouraging the other children to join in. I also took on one of the character roles (farmer, chef) during this play, and focused on incorporating target words whenever possible, as well as capitalizing on opportunities for developing conceptual knowledge as they arose (e.g., talking about why the seeds won’t grow if we plant them in the farmer’s hat).

Throughout all four days of play, several language support strategies were used in order to increase children’s depth of knowledge of target words. These strategies included 1) encouraging children to use the book and theme words, 2) expanding children’s utterances and 3) asking open-ended questions to help develop conceptual understanding. Pre-written questions

were listed on prompt cards for reference, and a checklist was used to ensure that all target words were used during play. See Appendix D for selections from transcripts of play sessions that illustrate these strategies.

### **Measures and Variables of Interest**

**Coding for target word use.** I developed a coding system to identify and describe all adult uses of target words during book-reading and play. Children's use of target words was not coded because children were not always visible or audible on videotapes. All book-reading and play sessions were video-recorded, and half of all videos were selected for coding – two videos from each book for each playgroup, or a total of four videos for each playgroup.

We selected the videos from days two and three as most representative of the range of instruction used in the intervention, as day two was designed to be more instructional and day three more responsive and interactive. In three instances, a video from day one or four was substituted because the day two or three video was missing or incomplete. The average video length was 21.06 minutes (median 21.75 minutes) and ranged from 12-33 minutes. Analyses controlled for the number of target word exposures as a way of equalizing intensity of exposure to the intervention. Exposure was controlled for in this manner, rather than including length of video as a covariate, because some of the longer videos were longer because of factors such as restroom breaks, behavior management, or other non-instructional interruptions, not because more instruction took place.

Coders recorded each use of a target word by the adult. The coder then filled out the following fields for each use of the target word: the setting (book-reading or play), interaction type, and supports for understanding the word. These codes are described in more detail below. An education master's student was trained to criterion (90% agreement) on the coding scheme

and coded all of the selected videos. To establish interrater reliability, 20% of the videos were double-coded by the author; interrater reliability was high (95.7%). Interrater reliability for specific categories is given below.

***Number of exposures.*** Because book-reading and play sessions were designed to be responsive to children's interests and questions, the procedures did not strictly control for the number of times each word was used. The coding of videos counted each use of the target word so that a statistical control could be created for this factor. Interrater reliability was high for this category (90%).

***Adult-child interactions.*** The coder selected one of three mutually exclusive options for the type of adult-child interaction in which a target word was used: instructional, responsive, and active processing. Instructional interactions were those intended to teach or transmit knowledge, rather than respond to children's cues, while using a target word. Types of instructional interactions included giving a word's definition, labeling a picture with the target word, reading the book text, or directing children's play while using a target word (e.g., "why don't you plant some *seeds*?"). Responsive interactions were those that responded to something a child did or said, while using a target word. Types of responsive interactions included answering a child's question, expanding or recasting children's utterances, or building on a child's play idea. Active processing interactions were those that asked children to synthesize or analyze word meaning. Types of active processing interactions included asking children about nuances of word meaning (e.g., "how is *sprouting* different from growing?") or asking open-ended questions that probed category membership (e.g., "how do you know that an *artichoke* is a vegetable and not a fruit?"). More examples of these codes are given in Appendix E. Interrater reliability was high for this category (95.6%).



**Word supports.** Coders selected from six non-exclusive codes to describe the kinds of supports used to teach word meanings. These codes were as follows: 1) definition – definitional information is given about the word, 2) part of book text – word is read aloud as part of the book; 3) book picture – adult points to a picture in the book to illustrate word meaning; 4) picture card – adult holds up or points at the picture card for the word; 5) gesture – teacher performs a gesture that illustrates word’s meaning in conjunction with verbal use of the word; 6) prop – target word is indexed to a toy/prop. Because codes 3-6 were not significant individually and provided similar types of support, a composite “visual support” variable was created by adding codes 3-6 together. Interrater reliability was high for this category (96.6%).

**Peabody Picture Vocabulary Test, Version IV.** To assess general vocabulary breadth and language abilities of the sample as compared to their age group peers, we administered the Peabody Picture Vocabulary Test (PPVT, Dunn & Dunn, 2007) before the intervention began. For this sample, the mean standard score was slightly lower than the normative mean of 100 ( $M = 97.0$ ) and the standard deviation was slightly higher than the normative standard deviation of 15 ( $SD = 16.05$ ).

**Vocabulary Breadth Measure.** To measure children’s receptive understanding of target words, an experimenter-designed measure was modeled after the PPVT-IV and administered at pretest and posttest. This measure captures vocabulary breadth in that it taps relatively shallow knowledge of target words, and was included so that even minimal knowledge of target words (i.e., knowledge that children are not yet able to verbalize) could be captured. Similar multiple choice tests have been widely used to assess target word comprehension (Blewitt, Rump, Shealy, & Cook, 2009; Penno et al., 2002; Senechal, 1997). The examiner stated a word and asked the child to select the referent from three illustrations, including a correct referent, a foil from the

same taxonomy (e.g., a marigold for the target word *tulip*) and a foil from the overall theme of “growing things” (e.g., a fern for the target word *hyacinth*). For the target word *artichoke*, the taxonomically related foil was a picture of a cucumber, and the thematically related foil was a picture of a lemon. Pictures of the target words were selected that were different from those used during the intervention, and four practice items depicting familiar objects were used at the beginning of the test to be certain that children understood the task. The test was comprised of 18 target, 5 exposure, and 8 control words (see Appendix C). Four words were omitted from the test and assessed only on the NWDT-M measure (see below for a description of this measure) due to limited imageability (e.g., *raw*), or because they were high-frequency words (e.g., *vegetables*, *flowers*) that children likely had at least minimal knowledge about, and were therefore best measured on the NWDT-M, which allows for measurement of incremental increases in knowledge. There were two versions of the test (version A and B) with the items in different orders; the order in which these versions were given to children was counterbalanced.

**Vocabulary Depth Measure, New Word Definition Test—Modified (NWDT-M).** To measure children’s depth of knowledge of target words, an experimenter-designed measure was developed and administered at pretest and posttest. This measure was adapted from Blewitt, Rump, Shealy, and Cook’s (2009) New Word Definition Test, which we renamed as the New Word Definition Test—Modified (NWDT–M) to reflect our adaptations to the coding scheme, namely, additional categories for gestures and contextual information. This definition task employs an informal rather than a formal definition task (Snow, Cancino, De Temple, & Schley, 1991), coding for the amount of accurate semantic and contextual information that children provided for each word, rather than their ability to give adult-like definitions of words. It taps

vocabulary depth by assessing multiple aspects of word knowledge and probing for relatively high-quality knowledge of words.

Children were asked to define words verbally or by using gestures. Children were tested on a representative subset of the total number of target, exposure, and control words on this measure (23 out of 35 words; see Appendix C) due to time constraints and the cognitive demands on children. For each word, children were asked, “What is (a) \_\_\_?” and a follow-up question, “Can you show me or tell me anything else about \_\_\_?” If a student did not respond to a question, the tester moved on to the next word. All student responses were transcribed by testers. Two forms of the test (A and B) listed words in different orders, and the order in which these forms were administered was counterbalanced.

A coding scheme was developed (adapted from Blewitt et al., 2009) to categorize and score student responses for the number of information units given. Coding was conducted by a research assistant, and 20% of all forms were randomly selected and checked for reliability against a master coder after every four forms were completed. Overall percentage agreement averaged 97.6%, with a mean Cohen’s Kappa value of .97.

***Coding Scheme.*** We used eight information unit categories to score student responses for semantic content and contextual information: category information (naming the taxonomy or a taxonomy member), perceptual qualities, functional information, part/whole, synonyms, gestures, meaningful context, and basic context. Each information unit was worth 1 point except for basic context, which was worth 0.5 point. The first four categories were used for concrete nouns only. Perceptual qualities included properties such as how something looks, smells, tastes, feels, or sounds. Functional information included any process, purpose, or use for concrete nouns and answers the question, “What do you do with it?” Part/whole described a distinct part of a

target word or the whole that the target word was a part of. The remaining categories were used for all word types. Synonyms included any word or short phrase that was equivalent to the word being explained, and provided decontextualized meaning information. Gestures included gestures or actions that showed knowledge of the word's meaning (e.g., curling up in a ball and then gradually standing up to represent *sprouting*).

We also coded for two types of use in context. Meaningful context included responses that showed knowledge of the target word in a typical, meaningful context, along with semantic information. For example, one student said, “*Seeds* grow. They grow into a red tree.” In this example, “grow” would be scored for function, and “into a red tree” would be scored for meaningful context, because the student used an example to illustrate what seeds might grow into, along with semantic information. Basic context, worth only 0.5 point, was a simple association between a target word and a typical context, without any use of semantic information. For example, several children said, “monkey,” for *vines*, a response that does not include semantic information but still contains an association with a typical context in which the target word is used. Incorrect or irrelevant responses received a score of 0. See Appendix F for examples of student responses and scoring.

### **Data Analysis**

This study uses a within-subject design in which children serve as their own controls, and their learning of one kind of words is compared to their learning of another (e.g., their learning of taught words to exposure and control words). The advantages of using a within-subjects design are that it controls for classroom and demographic factors, and has the power to detect small effects. We used multilevel regression models to account for the nested nature of our data, in which measurement occasions are nested within children, and children are in turn nested within

playgroups. In our analyses, we examine children's residualized gains (posttest vocabulary knowledge controlling for pretest vocabulary knowledge) in vocabulary knowledge. Unless otherwise noted below, all post hoc pairwise comparisons were conducted using Fisher's Least Significant Difference (LSD) test, and effect sizes are presented as Cohen's *d* (Cohen, 1988).

## **Results**

Table 1 provides mean raw scores and standard deviations for both measures and all word types examined in hypotheses 1-3 at pretest and posttest.

Table 1  
*Depth and Breadth Measure Unadjusted Means (Standard Deviations)*

Variable	Pretest	Posttest
Hypothesis 1: Growth in Vocabulary Breadth and Depth		
Breadth measure		
Target Words	0.48 (0.14)	0.85 (0.11)
Exposure Words	0.49 (0.20)	0.42 (0.19)
Control Words	0.47 (0.23)	0.51 (0.21)
Depth measure		
Target Words	0.36 (0.23)	1.18 (0.54)
Exposure Words	0.17 (0.26)	0.16 (0.31)
Control Words	0.17 (0.31)	0.17 (0.32)
Hypothesis 2: Teaching Words in Two vs. One Units		
Breadth measure		
One Unit Words	0.41 (0.15)	0.83 (0.13)
Two Unit Words	0.64 (0.18)	0.89 (0.13)
Depth measure		
One Unit Words	0.40 (0.28)	1.23 (0.60)
Two Unit Words	0.30 (0.24)	1.15 (0.54)
Hypothesis 3: Teaching Taxonomy vs. Theme Words		
Breadth measure		
Taxonomy Words	0.41 (0.18)	0.85 (0.12)
No-Tax Words	0.55 (0.17)	0.85 (0.15)
Depth measure		
Taxonomy Words	0.58 (0.36)	1.70 (0.79)
No-Tax Words	0.23 (0.23)	0.87 (0.45)

*Note.* Breadth measure values indicate the proportion of items that were answered correctly. Depth measure values indicate the average number of information units children provided for each word.

### **Growth in Vocabulary Breadth and Depth**

Our first hypothesis is related to the main effect of the intervention, predicting that children’s learning of target words would be greater than that of exposure and control words. In separate multilevel regression models for the two vocabulary measures, we tested whether vocabulary gains varied by level of instruction (taught, exposure, and control words):

$$(1) \text{ Posttest}_{ij} = \gamma_{00} + (\gamma_{10} * \text{Exposure}_{ij}) + (\gamma_{20} * \text{Control}_{ij}) + (\gamma_{30} * \text{Pretest}_{ij}) + U_{0j} + e_{ij}$$

This model accounted for two nesting levels in the data: level of instruction<sub>ij</sub> (target, exposure, and control) was nested within children<sub>j</sub> ( $n = 30$ ). For parsimony, the playgroup random effects were aggregated at the child level. Level of instruction was dummy coded with target words as the reference group, which were contrasted with exposure ( $\gamma_{10}$ ) and control ( $\gamma_{20}$ ) words. To look at residualized gains, children’s pretest vocabulary scores ( $\gamma_{30}$ ) were included as a covariate.

Results for both the breadth and depth measures indicated that children learned significantly more about target words than either exposure or control words (see Table 2). Post-hoc pairwise comparisons also indicated that there was no significant difference between children’s learning of control and exposure words on the depth measure ( $p = 0.091$ ,  $d = 0.400$ ), but that control words were learned significantly better than exposure words on the breadth measure ( $p = 0.012$ ,  $d = 0.436$ ).

Table 2  
*Parameter Estimates (Standard Errors) for Growth in Word Knowledge (Top Panel) and Effect Sizes (Bottom Panel)*

Parameters	Vocabulary Breadth	Vocabulary Depth
Main Effects of Intervention		
Level 1, Level of Instruction		
Intercept,	0.651 (0.499)**	0.848 (0.079)**
Pre-test score,	0.414 (0.084)**	0.926 (0.135)**
Target versus Exposure,	-0.336 (0.037)**	-0.840 (0.077)**
Target versus Control,	-0.336 (0.037)**	-0.715 (0.077)**
Cohen’s $d$ Effect Sizes		
Target versus Exposure,	2.783**	1.906**
Target versus Control,	2.004**	1.609**

Note. Standard errors adjusted for interdependency of level of instruction nested within children. Target Words are the reference group for the comparison (negative estimates indicate that target words had larger covariate adjusted post-test scores).

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

## Teaching Words in One vs. Two Units

Our second hypothesis examines whether target words that were taught in two units (vegetable **and** flower book & play) were learned better than target words taught in only one unit (vegetable **or** flower book & play). In separate multilevel regression models for the two vocabulary measures, we tested whether vocabulary gains were greater in two book/play units vs. one:

$$(2) \text{ Posttest}_{ijk} = \gamma_{000} + (\gamma_{100} * \text{OneUnit}_{ijk}) + (\gamma_{200} * \text{Pretest}_{ijk}) + (\gamma_{001} * \text{Number}_k) + U_{00k} + U_{0jk} + e_{ijk}$$

This model accounts for three nesting levels in the data:  $\text{unit}_{ijk}$  is nested within  $\text{children}_{ij}$ , who are nested in  $\text{playgroups}_j$ . In other words, children have separate sets of pretest and posttest scores for words taught in one unit and in two units, which are nested within children, who are nested within playgroups. Number of units was dummy-coded with two units as the reference group, which was contrasted to one unit ( $\gamma_{100}$ ). The number of uses variable ( $\gamma_{001}$ ) represents the number of times one-unit words and two-unit words were used, respectively, and was included as a covariate so that the effect of learning a word in one vs. two units was isolated by holding the number of times a word was heard constant.

Results for both the breadth and depth measures indicated that there was no significant difference between children's learning of words taught in two units vs. one (see Table 3), but effect sizes indicate a moderate effect favoring words taught in one unit ( $d = 0.400$  on breadth measure,  $d = 0.734$  on depth measure).



Table 3

*Parameter Estimates (Standard Errors) for Teaching Words in One vs. Two Units (Top Panel) and Effect Sizes (Bottom Panel)*

Parameters	Vocabulary Breadth	Vocabulary Depth
Effect of One Unit vs. Two Units		
Level 1, Number of Units		
Intercept,	0.651 (0.094)**	1.343 (0.292)**
Pre-test score,	0.417 (0.086)**	0.612 (0.220)**
One Unit vs. Two Units	0.052 (0.064)	0.419 (0.274)
Level 3, Playgroup		
Number of exposures to target words	-0.009 (0.064)	-0.159 (0.105)
Cohen's <i>d</i> Effect Sizes		
One Unit vs. Two Units	0.400	0.734

Note. Standard errors adjusted for interdependency of context nested within children and children nested within playgroup. Two Units is the reference group for the comparison (positive estimates indicate that words taught in One Unit had larger covariate adjusted post-test scores).

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

### Teaching Taxonomy vs. Theme Words

The third hypothesis predicts that children will learn more about taxonomically-related than thematically-related target words, controlling for the number of exposures. We tested whether taxonomy words were learned better than theme words using separate multilevel regression models for the two vocabulary measures similar to Equation 2. Number of exposures ( $\gamma_{001}$ ), the number of times taxonomy and theme words were each used, was included as a covariate because some of the theme words were taught in more than one book and therefore received roughly double the exposures of other words. This covariate allowed us to hold exposures constant and isolate the effect of teaching in taxonomies v. themes on word-learning. Word type was dummy-coded with theme words as the reference group, which was contrasted to taxonomy ( $\gamma_{100}$ ).

Analyses revealed that children learned significantly more about taxonomy words than theme words on the depth, but not the breadth measure (see Table 4).

Table 4

*Parameter Estimates (Standard Errors) for Effect of Taxonomy Words (Top Panel) and Effect Sizes (Bottom Panel)*

Parameters	Vocabulary Breadth	Vocabulary Depth
Effect of Taxonomy Words		
Level 1, Word type		
Intercept,	0.741 (0.121)**	1.106 (0.415)*
Pre-test score,	0.268 (0.093)**	0.859 (0.230)**
Taxonomy v. Theme	0.053 (0.051)	0.720 (0.231)**
Level 3, Playgroup		
Number of exposures to target words	-0.001 (0.001)	-0.001 (0.002)
Cohen's <i>d</i> Effect Sizes		
Taxonomy v. Theme	0.390	1.120**

*Note.* Standard errors adjusted for interdependency of word type nested within children and children nested within playgroup. Theme words are the reference group for the comparison (positive estimates indicate that words taught in Taxonomy had larger covariate adjusted post-test scores).

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

### **Correlations Between Teacher-Child Interactions, Word Supports, and Pre-Post Target Word Knowledge**

Table 5 provides the means and correlation matrices for vocabulary measures, types of teacher-child interactions, and word supports. The pretest scores on both the vocabulary breadth and depth measures were negatively associated with many of the interaction and word support types (most strongly with definitions and instructional interactions during book-reading), suggesting that children who knew more about words at pretest received less information about words during sessions. Posttest scores showed similar, although weaker, negative relationships with the number of interactions and word supports provided.

Other relationships of note included a positive correlation between responsive play interactions (e.g., the adult expanding child utterances during play) and instructional book interactions (e.g., the adult supplying definitions during book-reading). Given that play interactions were designed to be responsive and book interactions were designed to be more instructional, this indicates that when book interactions went “as planned,” play interactions did

as well, perhaps because children were especially attentive during both. The same could be said of the strong positive relationships between 1) visual supports for words (a planned part of the method) and responsive play interactions, where the use of more visual supports was related to the use of more responsive play interactions, and 2) visual supports and instructional book interactions, where the use of more visual supports was related to the use of more instructional interactions during book-reading. These factors likely are associated with each other as they represent a book-reading and play session going according to plan, in which substantial visual supports, instructional book interactions, and responsive play interactions were all present.

Table 5  
*Correlation Matrices for Outcome Measures, Interaction Types, and Word Supports*

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11	12	13
1. VB (pre)	0.48	0.23	1												
2. VB (post)	0.85	0.54	0.501**	1											
3. VD (pre)	0.36	0.14	0.432*	0.230	1										
4. VD (post)	1.18	0.11	0.655**	0.594**	0.612**	1									
5. RSP-B	38.60	12.76	0.103	0.182	0.099	0.096	1								
6. RSP-PL	128.50	22.78	-0.441*	-0.205	-0.368*	-0.279	0.049	1							
7. INS-B	346.70	43.58	-0.451*	-0.390*	-0.514**	-0.332	0.122	0.789**	1						
8. INS-PL	83.60	21.59	-0.159	0.075	-0.214	-0.196	-0.271	0.099	-0.137	1					
9. ACT-B	22.70	10.07	0.115	0.060	0.235	0.077	0.565**	0.343	0.156	0.017	1				
10. ACT-PL	5.20	2.44	-0.378*	-0.365*	-0.099	-0.359	-0.017	0.685**	0.552**	0.092	0.150	1			
11. DEF	107.80	23.30	-0.504**	-0.193	-0.532**	-0.372*	0.202	0.655**	0.722**	-0.014	-0.023	0.423*	1		
12. TEXT	55.80	6.64	0.005	0.203	-0.236	0.180	-0.229*	0.073	0.081	-0.083	-0.219	-0.534**	-0.079	1	
13. VIS	560.10	59.85	-0.456*	-0.241	-0.451*	-0.296	0.325	0.834**	0.907**	0.010	0.396*	0.473**	0.774**	0.095	1
14. EXP	476.50	103.43	-0.384*	-0.126	-0.254	-0.210	0.198	0.888**	0.590**	0.157	0.470*	0.764**	0.617**	-0.213	0.740**

Note. VB = vocabulary breadth score for target words; VD = vocabulary depth scores for target words; RSP-B = responsive interactions – book; RSP-PL = responsive interactions – play; INS-B = instructional interactions – book; INS-PL = instructional interactions – play; ACT-P = active processing – pre; ACT-B = active. The number of inter processing – book; ACT-PL = active processing – play; DEF = total number of target words defined; TEXT = total number of target words read from book text; VIS = total number of visual supports for target words; EXP = total number of exposures for target words. Breadth measure values indicate the proportion of items that were answered correctly. Depth measure values indicate the average number of information units children provided for each word. Values for variables 5-14 represent the average number of interactions or word supports summed across four book-reading/play sessions.

## Associations Between Teacher-Child Interactions and Growth in Vocabulary Knowledge

Hypothesis 4 examines the associations between types of teacher-child interactions (responsive, instructional, and active processing, in both book-reading and play settings) and vocabulary learning.

We tested the association between the six playgroup-level predictors ( $\gamma_{01}$ -- $\gamma_{06}$ ) for teacher-child interaction and the two vocabulary outcome measures, accounting for the nesting of children<sub>ij</sub> in playgroups<sub>j</sub>:

$$(3) \text{ Posttest}_{ij} = \gamma_{00} + (\gamma_{10} * \text{Pretest}_{ij}) + (\gamma_{01} * \text{BookResp}_j) + (\gamma_{02} * \text{BookIns}_j) + (\gamma_{03} * \text{BookActPro}_j) + (\gamma_{04} * \text{PlayResp}_j) + (\gamma_{05} * \text{PlayIns}_j) + (\gamma_{06} * \text{PlayActPro}_j) + U_{0j} + e_{ij}$$

Table 6 shows the results from this analysis. All variables were entered into the model simultaneously, and the coefficients can therefore be interpreted as the unique effect of a given interaction type, holding all others equal. The number of responsive interactions during play showed a positive and statistically significant association with growth in vocabulary depth and breadth, while the active processing interactions during play were negatively and significantly associated with both vocabulary depth and breadth (see Table 6). There were significant associations between interaction type and vocabulary growth during book-reading on the breadth measure only, with responsive interactions positively predicting growth and instructional interactions negatively predicting growth.

We estimated effect sizes for all linear relationships by multiplying the coefficient of the predictor by its standard deviation, then dividing by the standard deviation of the dependent variable (this approach is also used in Levya et al., 2015; NICHD ECCRN & Duncan, 2003; Burchinal et al., 2010). This effect size indicates the change in the outcome variable in standard deviation units when the predictor increases by a standard deviation, and is equivalent to

Cohen's *d*. See Table 6. While not statistically significant, there was a moderate negative effect of active processing interactions during book-reading on the breadth measure ( $d = -0.366$ ), and a moderate negative effect of instructional interactions during book-reading on the depth measure ( $d = -0.323$ ).

Table 6  
*Parameter Estimates (Standard Errors) for Interaction Types (Top Panel) and Effect Sizes (Bottom Panel)*

Parameters	Vocabulary Breadth	Vocabulary Depth
Interaction Types		
Level 1, Child		
Intercept,	0.661 (0.206)**	0.722 (0.922)
Pre-test score,	0.378 (0.125)**	1.56 (0.370)**
Level 2, Playgroup		
Book-reading Interactions		
Responsive	0.003 (0.002)*	0.008 (0.007)
Instructional	-0.001 (0.001)*	-0.004 (0.003)
Active Processing	-0.004 (0.002)	-0.008 (0.010)
Play Interactions		
Responsive	0.004 (0.001)**	0.015 (0.007)*
Instructional	0.001 (0.001)	-0.001 (0.004)
Active Processing	-0.017 (0.009)*	-0.112 (0.039)**
Effect Sizes		
Book-reading Interactions		
Responsive	0.348*	0.189
Instructional	-0.396*	-0.323
Active Processing	-0.366	-0.149
Play Interactions		
Responsive	0.828**	0.633*
Instructional	0.196	-0.040
Active Processing	-0.377*	-0.506**

*Note.* Standard errors adjusted for interdependency of children nested within playgroups. Effect sizes were calculated by multiplying the predictor's coefficient by its standard deviation, then dividing by the standard deviation of the outcome variable. The resulting effect size is equivalent to Cohen's *d*.

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

### Associations Between Word Supports and Growth in Vocabulary Knowledge

Hypothesis 5 investigates the relationship between three predictor variables representing support for target words: 1) giving definitions, 2) hearing target words read as part of the book text, and 3) visual supports (gestures, pointing at the book picture, props, and using picture

cards) and the vocabulary outcome variables. The model used was similar to Equation 3, but the playgroup-level predictors of interaction type were replaced with word support predictors. A covariate was also included to control for the number of target word uses.

Analysis indicated that hearing target words read as part of the book text was significantly and positively associated with growth in vocabulary breadth and depth, controlling for the other words supports and number of target words used (see Table 7 for parameter estimates and effect sizes). No other word support predictors were significantly associated with growth in word knowledge, but there was a moderate, although nonsignificant, negative effect of visual supports on vocabulary breadth ( $d = -0.544$ ).

Table 7  
*Parameter Estimates (Standard Errors) for Word Supports (Top Panel) and Effect Sizes (Bottom Panel)*

Parameters	Vocabulary Breadth	Vocabulary Depth
Interaction Types		
Level 1, Child		
Intercept,	0.514 (0.239)*	-0.838 (1.024)
Pre-test score,	0.435 (0.135)**	1.639 (0.375)**
Level 2, Playgroup		
Definitions	0.001 (0.001)	0.002 (0.005)
Hearing target words used in book text	0.006 (0.003)*	0.033 (0.012)*
Visual supports	-0.001 (0.001)	-0.001 (0.002)
Number of exposures to target words	0.001 (0.001)	0.001 (0.001)
Effect Sizes		
Definitions	0.212	0.086
Hearing target words used in book text	0.362*	0.406*
Visual supports	-0.544	0.111

Note. Standard errors adjusted for interdependency of children nested within playgroups. Effect sizes were calculated by multiplying the predictor's coefficient by its standard deviation, then dividing by the standard deviation of the outcome variable. The resulting effect size is equivalent to Cohen's  $d$ .

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

## Discussion

The purpose of this study was two-fold: 1) to examine the impact of a vocabulary intervention designed to support preschoolers' depth of vocabulary knowledge, and 2) to

investigate specific factors that may contribute to growth in depth. The present intervention had significant positive effects on children's breadth and depth of vocabulary knowledge, with taxonomy words learned more deeply than theme words, and no difference in learning for words taught in two units vs. one. Responsive play interactions and hearing target words read as part of the book text predicted growth in word knowledge, and active processing interactions during play were negatively associated with vocabulary growth. I discuss the implications of these findings in more detail below.

### **Growth in Vocabulary Depth and Breadth**

The present intervention showed a substantial positive impact on children's growth in breadth ( $d = 2.00$ , target vs. control words) and depth ( $d = 1.61$ , target vs. control words) of vocabulary knowledge. On the breadth measure, children on average correctly identified 48% of images corresponding to target words at pretest, and 85% at posttest. This indicates that children gained at least a fast-mapped understanding of most words. However, children also showed substantial growth on the more demanding vocabulary depth measure, which tapped higher lexical quality by asking children to provide semantic and contextual information about words. At pretest, children gave approximately 1 piece of information for every 3 target words, while at posttest, they were able to give more than 4.5 pieces of information for every 3 target words. This demonstrates an increase in the lexical quality of children's knowledge for target words.

The growth in vocabulary knowledge shown here is larger than reported by meta-analyses of preschool vocabulary interventions, with effect sizes of  $d = 0.60$  for shared book-reading interventions (National Early Literacy Panel, 2008) and  $g = 0.85$  for preschool vocabulary interventions in general (Marulis & Neuman, 2010). This study had several features associated with higher effect sizes in meta-analyses: a researcher, rather than teachers or child



care providers, delivered the intervention (Marulis & Neuman, 2010; Mol et al., 2009), author-created rather than standardized measures were used to assess growth (Marulis & Neuman, 2010), and instruction combined both explicit (e.g., giving definitions) and implicit methods (e.g., embedding target words in guided play) (Marulis & Neuman, 2010). The large effect sizes may also be partially driven by our selection of target words, as the concrete nouns taught in the present study are relatively easier to learn than more abstract, less imageable words (Hadley et al., 2016). I highlight other specific features of instruction below that may have also contributed to the substantial growth in vocabulary knowledge.

Fostering depth of vocabulary knowledge, rather than breadth, has sometimes been characterized as a prohibitively time-consuming endeavor, given the large number of words young children need to learn (e.g., Biemiller & Boote, 2006). Our results do indicate that an investment of systematic instructional time helps to support depth: children showed no growth in knowledge for exposure words, which were simply heard in the book text but not explained or used during play. However, the preschoolers in this study showed large gains in word knowledge from a relatively short daily period of instruction (20 min.), and twenty-one words were covered in eight days. This favorably compares to other interventions aimed at supporting extensive word knowledge that have taught a smaller number of words in a similar time frame (e.g., Beck & McKeown, 2007, twelve words in two weeks). The results here suggest that young children are capable of significant improvements in the breadth and depth of their word knowledge in a relatively short amount of time, making depth a reasonable instructional goal for preschool classrooms.

## Teaching Taxonomy Versus Theme Words

Children learned taxonomically-related words more deeply than theme-related words ( $d = 1.12$ ), although there were increases in learning for both word types. These results are consistent with a preferential attachment theory of word-learning (Steyvers & Tenenbaum, 2005), where new words are learned more quickly and deeply when they are semantically related to known words. These results also support a perspective on fostering depth that emphasizes not only supporting lexical quality for individual concepts, but also expanding semantic networks by teaching words in conceptually-related groups. In this view, vocabulary instruction can be considered not just as a one-by-one proposition where a single word is taught and learned, but a systems-level approach where broader networks of related concepts are introduced together to maximize learning.

In particular, the results here indicate that teaching words in taxonomies may be of particular benefit for efficient, deep word-learning as compared to teaching words in themes. The extensive support for the higher-level categories (vegetables, flowers, and growing things) taught here may have helped to leverage children's word-learning of the exemplars in this category. For example, once a child knew that an *artichoke* was a vegetable, they could generalize the information they had already learned about vegetables to that exemplar without extensive explicit instruction. To illustrate this point, Table 8 provides the pretest and posttest responses for a single child for a taxonomy and a theme word. For the taxonomy word (*cauliflower*), the child is able to identify the superordinate category and apply one of the category properties to this specific exemplar ("it's a vegetable; you eat it"). These pieces of information could be given for any exemplar in this taxonomy. The child also gives information that pertains more

specifically to the vegetable in question (“It’s white. We don’t eat the green part.”), scoring a total of four points.

In contrast, for the theme word (*soil*), the child gives a synonym (“soil is dirt,”) and gives information about where soil is located, perhaps referring to the plant-growing process (“you put it in the ground,”), scoring a total of two points. However, this child does not have the same knowledge of a higher-level category to draw on to further support her answer. The increased learning shown for taxonomy words as compared to theme words, then, is a signal that the lexical quality not only for individual words improved, but that instruction may have also supported knowledge of, and generalization of, the underlying taxonomic categories.

Table 8  
*One child’s responses on depth measure for a taxonomy and a theme word, at pretest and posttest.*

Prompt	Pretest response	Posttest response
What is a <i>cauliflower</i> ? (taxonomy word)	“I don’t know.”	“It’s a vegetable; you eat it. It’s white. We don’t eat the green part.” (4 points)
What is <i>soil</i> ? (theme word)	“I don’t know.”	“Soil is dirt. You put it in the ground.” (2 points)

It is important to note that the present study used informational texts during book-reading sessions, which were particularly supportive of the concepts underlying the taxonomy words. These results may not generalize to narrative texts, in which thematic, rather than taxonomic, relationships between words are more common. Science and social studies activities could be other activity settings in which teaching words in taxonomies could be productively applied, helping to prepare children for the demands of academic language in those content areas. Other research has shown that preschool teachers provide conceptual information more frequently in content areas such as science, math, and social studies than during book-reading (Bowne,

Yoshikawa, & Snow, 2017), which suggests that these are areas ripe for introducing taxonomic thinking.

### **Teaching Words in One Versus Two Units**

Contrary to my hypothesis, children did not show significant differences in growth for words learned in two units v. one. In fact, there were medium to large, although nonsignificant, effect sizes ( $d = 0.400$  on breadth,  $d = 0.734$  on depth) favoring words taught in one unit. These findings are consistent with those of Horst and colleagues (2011), who found that 3-year-olds learned more about novel words when they were read in a single book context three times than in three different books. In contrast, the kindergartners in McKeown and Beck's 2014 study benefited from discussing words in multiple contexts. These disparate findings may be related to age, as the preschoolers in our study, as compared to the kindergartners in McKeown & Beck (2014), may have needed more time with words in a single context before being challenged to integrate that knowledge into a new context. Hearing a new word used in a different context may have required more attentional resources from children, interfering with the automatic, quick retrieval of word identities. Extending words to new contexts may be beneficial at later stages of the word-learning process for preschoolers, once the lexical representation is more firmly established. From an educational perspective, these findings suggest that the practice of reading the same book multiple times may be as beneficial for supporting depth as the reading of multiple books that have sets of target words in common.

### **Associations Between Adult-Child Interactions and Growth in Vocabulary Knowledge**

I examined a range of interaction types during both book-reading and play and their relation to growth in word-learning. I will examine first the findings from the play activity setting, a small-group activity in which I directed play for the first two days, instructing children

to act out scenes and concepts from the book, and followed children's lead for the second two days of play, building on and expanding their utterances and play ideas. As predicted, responsive interactions during play were the strongest predictor of growth in breadth ( $d = 0.828$ ) and depth ( $d = 0.633$ ) of vocabulary knowledge, with smaller positive effect sizes for responsive interactions during book-reading. These findings indicate the importance of joint engagement in vocabulary learning, as the coding of responsive interactions captured moments in which a child offered a comment, question, play idea, or nonverbal overture (such as silently offering me a toy), and I answered, expanded children's utterances, or extended the play idea while using a target word. This suggests that information about word-meaning is most helpful when instantiated in the context of adult-child interactions that provide contingent, thoughtful scaffolding of children's emergent word knowledge.

An alternate explanation for the responsive interactions finding is that children who were learning more about words also offered more on-topic overtures, creating opportunities for the provision of contingent responses that included target words. This is consistent with transactional models of development in which adults' interactions with children are significantly influenced by the child's own competencies (Sameroff & Mackenzie, 2003). This finding implies that children who are already learning more about words may be more likely to initiate cycles of interaction that lead to deeper word-learning, with children who are learning less making fewer overtures. In a classroom context, this may mean that teachers should be particularly mindful of supporting reticent or children with less developed language skills to initiate more interactions, perhaps by modeling play or recognizing and building on non-verbal overtures in small-group settings (Kaiser, Hancock, & Nietfeld, 2000).

Contrary to my hypothesis, active processing interactions during play were negatively related to growth in word knowledge. While active processing interactions have been productively employed with kindergarten students (Coyne et al., 2010; McKeown & Beck, 2014), the questions used here may have simply been too challenging for preschoolers. For example, I asked questions such as, “Is a *daffodil* a *flower*, or a *vegetable*? How do you know?” and often got no responses from children. It is possible that children’s developing word knowledge was too fragile to accommodate these probing questions. Children may have also been unfamiliar with these types of questions, as analytic talk is relatively uncommon in preschool classrooms (Dickinson & Porche, 2011). Furthermore, play may have not been an appropriate setting for active processing questions, as other studies have used this strategy in post-book-reading discussions (Coyne, McCoach, & Kapp, 2007; McKeown & Beck, 2014). Asking children questions about word meaning had the effect of drawing children away from play, which may have interrupted the absorption in the play world that provided a basis for joint engagement.

Book-reading also took place in small groups of three children. During the reading, semantic and taxonomic information was supplied for new words. Children were encouraged to chime in with information about word meaning during the 3<sup>rd</sup> and 4<sup>th</sup> readings of the book. Contrary to my hypothesis, instructional talk during book-reading was a significant negative predictor of growth in word-learning ( $d = -0.396$  on breadth measure). These results are surprising as instructional talk was a primary vehicle for explicit information about word meaning, a factor which has been shown to support children’s word-learning (Biemiller & Boote, 2006). However, the amount of instructional talk during book-reading (on average, 346 instances over four sessions, see Table 5) may have crowded out opportunities for child participation. In a longitudinal study, children who were in classrooms with lower ratios of teacher-child talk in preschool had higher literacy skills

in kindergarten (Dickinson & Porche, 2011), indicating that creating sufficient space for child talk is an important factor in supporting literacy. Too much extra-textual talk during book-reading could also impair comprehension, which in turn impairs vocabulary knowledge (Rivera & Dickinson, 2013). An alternate explanation for this finding is that children who were learning more about words supplied more definitions and information about words themselves, driving down the amount of instructional talk. In later readings of the book, I actively elicited children's participation, asking them to "fill in the blanks" with target words and definitions as I read. Children who were learning less about words may not have responded to these elicitations, meaning that I supplied more instructional talk in return.

There was also a significant positive association between responsive interactions during book-reading and growth in word knowledge on the breadth measure. While book-reading is not a typical setting for responsive interactions (Gest, 2006), the small group setting here allowed for more opportunities to build on children's responses.

### **Associations Between Word Supports and Growth in Vocabulary Knowledge**

In examining the relationship between growth in vocabulary knowledge and the supports that were provided for word meaning, only hearing target words used in the book text was significantly related to growth ( $d = 0.362$  for breadth,  $d = 0.406$  for depth). This points to the importance of the book contexts as a fertile source of meaning for words. The books used here were informational texts that contained rich conceptual information, and also made explicit connections between related concepts, supporting semantic network growth. For example, *Vegetables in the Garden* listed properties of vegetables and described different kinds of vegetables based on which part of the plant was eaten. This finding should not be interpreted to suggest that *only* hearing words used in the book text is necessary for learning; rather, it indicates

that when controlling for other factors (such as explicit definitions, exposure, and visual supports), hearing words used in the book text can make a unique contribution to growth in word-learning. Furthermore, these findings should be interpreted in light of the negative correlations between pretest scores and the definitions and visual support variables (see Table 5), suggesting that I may have provided less explicit information about word meaning when I perceived that children already knew something about words.

### **Limitations**

This study involves several limitations. First, the sample size is relatively small, and the intervention was implemented by a researcher in small groups of three children. This may limit the generalization of these results to whole-group classroom settings. Secondly, the analyses for hypotheses 4 and 5 were correlational, not causal, in nature. To determine the directionality of the relationships found, experiments would need to be conducted that randomly assigned children to different conditions or used a within-subjects design to teach words with different techniques. Finally, we did not measure children's retention of learned words or transfer to a novel learning context.

### **Conclusions**

The present study suggests that fostering deep vocabulary knowledge involves not only providing information to support the lexical quality of individual words, but also demonstrating how the concepts labeled by words fit together in semantic networks. Furthermore, these findings demonstrate the importance of social interactions between adults and children that build word knowledge on a foundation of joint engagement and communication, supported by the rich content inherent in book-reading and play.



## References

- Anderson, R. C., & Freebody, P. (1985). Vocabulary knowledge. In H. Singer & R. B. Ruddell (Eds.), *Theoretical Models and Processes of Reading* (3rd ed., pp. 343-371): International Reading Association.
- Beck, I. L., & McKeown, M. G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal*, *107*(3), 251-271. doi:<http://dx.doi.org/10.1086/511706>
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction*. New York: The Guilford Press.
- Biemiller, A. (2010). *Words worth teaching: Closing the vocabulary gap*. Columbus, OH: SRA/McGraw-Hill.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, *98*(1), 44-62. doi:<http://dx.doi.org/10.1037/0022-0663.98.1.44>
- Bloom, P. (2002). *How children learn the meanings of words*. Cambridge, MA: The MIT Press.
- Bolger, D. J., Balass, M., Landen, E., & Perfetti, C. A. (2008). Context variation and definitions in learning the meanings of words: An instance-based learning approach. *Discourse Processes*, *45*(2), 122-159. doi:10.1080/01638530701792826
- Börner, K., Sanyal, S., & Vespignani, A. (2007). Network science. In B. Cronin (Ed.), *Annual Review of Information Science & Technology* (Vol. 41, pp. 537-607). Medford, NJ: Information Today, Inc./American Society for Information Science and Technology.

- Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016a). Lexical leverage: category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science, 19*(6), 918-932. doi:10.1111/desc.12343
- Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016b). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development, 87*(6), 1893-1908. doi:10.1111/cdev.12554
- Bowne, J. B., Yoshikawa, H., & Snow, C. E. (2017). Relationships of teachers' language and explicit vocabulary instruction to students' vocabulary growth in kindergarten. *Reading Research Quarterly, 52*(1), 7-29. doi:10.1002/rrq.151
- Cabell, S. Q., Justice, L. M., Piasta, S. B., Curenton, S. M., Wiggins, A., Turnbull, K. P., & Petscher, Y. (2011). The impact of teacher responsivity education on preschoolers' language and literacy skills. *American Journal of Speech-Language Pathology, 20*(4), 315. Retrieved from doi:10.1044/1058-0360
- Chall, J. S., & Dale, E. (1995). Readability revisited. *The New Dale-Chall Readability Formula. Brookline: Brookline Books.*
- Christakis, D. A., Zimmerman, F. J., & Garrison, M. M. (2007). Effect of block play on language acquisition and attention in toddlers: a pilot randomized controlled trial. *Archives of pediatrics & adolescent medicine, 161*(10), 967.
- Coyne, M. D., McCoach, D. B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly, 74*-88. doi:http://dx.doi.org/10.2307/30035543

- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, *110*(1), 1-18. doi:10.1086/598840
- Coyne, M. D., McCoach, D. B., Loftus, S., Zipoli, R., Ruby, M., Crevecoeur, Y. C., & Kapp, S. (2010). Direct and extended vocabulary instruction in kindergarten: Investigating transfer effects. *Journal of Research on Educational Effectiveness*, *3*(2), 93-120. doi:10.1080/19345741003592410
- Cronin, V. S. (2002). The syntagmatic–paradigmatic shift and reading development. *Journal of Child Language*, *29*(1), 189-204. doi:10.1017/S0305000901004998
- Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, *17*(6), 948-955. doi:10.1111/desc.12164
- Dickinson, D. K. (2001). Large-group and free-play times: Conversational settings supporting language and literacy development. In D. K. Dickinson & P. O. Tabors (Eds.), *Beginning literacy with language: Young children learning at home and school* (pp. 223-255). Baltimore: Brookes.
- Dickinson, D. K. (2011). Teachers' language practices and academic outcomes of preschool children. *Science*, *333*(6045), 964-967. doi:10.1126/science.1204526
- Dickinson, D. K., Griffith, J. A., Golinkoff, R. M., & Hirsh-Pasek, K. (2012). How reading books fosters language development around the world. *Child Development Research*, *2012*.
- Dickinson, D. K., Nesbitt, K. T., Collins, M. F., Hadley, E. B., Newman, K. M., Rivera, B. L., . . . Golinkoff, R. (2016). A depth of processing account of the effects of using book-reading to teach for breadth and depth of knowledge. Manuscript in preparation.

- Dickinson, D. K., & Porche, M. V. (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. doi:10.1111/j.1467-8624.2011.01576.x
- Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science, 18*(3), 254-260. doi:10.1111/j.1467-9280.2007.01885.x
- Fisher, A. V., Godwin, K. E., Matlen, B. J., & Unger, L. (2015). Development of category-based induction and semantic knowledge. *Child Development, 86*(1), 48-62. doi:10.1111/cdev.12277
- Fisher, K. R., Hirsh-Pasek, K., Newcombe, N., & Golinkoff, R. M. (2013). Taking shape: Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child Development, 84*(6), 1872-1878. doi:10.1111/cdev.12091
- Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2011). Predicting robust vocabulary growth from measures of incremental learning. *Scientific Studies of Reading, 15*(1), 71-91.
- Gámez, P. B., & Lesaux, N. K. (2015). Early-adolescents' reading comprehension and the stability of the middle school classroom-language environment. *Developmental Psychology, 51*(4), 447.
- Girolametto, L., & Weitzman, E. (2002). Responsiveness of child care providers in interactions with toddlers and preschoolers. *Language, Speech & Hearing Services in Schools, 33*(4), 268-290.

- Girolametto, L., Weitzman, E., & Greenberg, J. (2003). Training day care staff to facilitate children's language. *American Journal of Speech-Language Pathology, 12*(3), 299-311. doi:10.1044/1058-0360(2003/076)
- Goldberg, A. E., Casenhiser, D. M., & Sethuraman, N. (2004). Learning argument structure generalizations. *Cognitive linguistics, 15*(3), 289-316.
- Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind: Predicting fluency with PageRank. *Psychological Science, 18*(12), 1069-1076.
- Hadley, E. B., Dickinson, D. K., Hirsh-Pasek, K., Golinkoff, R. M., & Nesbitt, K. T. (2016). Examining the acquisition of vocabulary knowledge depth among preschool students. *Reading Research Quarterly, 51*(2), 181-198. doi:10.1002/rrq.130
- Han, M., Moore, N., Vukelich, C., & Buell, M. (2010). Does play make a difference? How play intervention affects the vocabulary learning of at-risk preschoolers. *American Journal of Play, 3*(1), 82-105.
- Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Categorical structure among shared features in networks of early-learned nouns. *Cognition, 112*(3), 381-396. doi:http://dx.doi.org/10.1016/j.cognition.2009.06.002
- Hirsh-Pasek, K., & Golinkoff, R. M. (2011). The great balancing act: Optimizing core curricula through playful pedagogy. In E. Zigler, W. S. Gilliam, & W. S. Barnett (Eds.), *The Pre-K Debates: Current Controversies & Issues*. Baltimore, MD: Paul H. Brookes Publishing Company.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). II. The Emergentist Coalition Model. *Monographs of the Society for Research in Child Development, 65*(3), 17-29. doi:10.1111/1540-5834.00092

- Horst, J., Parsons, K., & Bryan, N. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology, 2*(17).  
doi:10.3389/fpsyg.2011.00017
- Kaefer, T., & Neuman, S. B. (2013). A bidirectional relationship between conceptual organization and word learning. *Child Development Research, 2013*.
- Kaiser, A. P., Hancock, T. B., & Nietfeld, J. P. (2000). The effects of parent-implemented enhanced milieu teaching on the social communication of children who have autism. *Early Education & Development, 11*(4), 423-446. doi:10.1207/s15566935eed1104\_4
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Landry, S. H., Smith, K. E., Swank, P. R., Zucker, T., Crawford, A. D., & Solari, E. F. (2012). The effects of a responsive parenting intervention on parent-child interactions during shared book reading. *Developmental Psychology, 48*(4), 969-986.  
doi:http://dx.doi.org/10.1037/a0026400
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, Mass: MIT Press.
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300-335.
- McKeown, M. G., & Beck, I. L. (2014). Effects of vocabulary instruction on measures of language processing: Comparing two approaches. *Early Childhood Research Quarterly, 29*(4), 520-530.

- Mol, S. E., Bus, A. G., & de Jong, M. T. (2009). Interactive book reading in early education: A tool to stimulate print knowledge as well as oral language. *Review of Educational Research, 79*(2), 979-1007. doi:10.3102/0034654309332561
- National Early Literacy Panel. (2008). *Developing early literacy: Report of the National Early Literacy Panel*. Retrieved from Washington, DC: Available at <http://www.nifl.gov/earlychildhood/NELP/NELPreport.html>
- Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literary Research, 43*(103), 103-129. doi:<http://dx.doi.org/10.1177/1086296X11403089>
- Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly, 46*(3), 249-272.
- Neuman, S. B., Pinkham, A., & Kaefer, T. (2015). Supporting vocabulary teaching and learning in prekindergarten: The role of educative curriculum materials. *Early Education and Development, 26*(7), 988-1011.
- Ouellette, G. P. (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. doi:<http://dx.doi.org/10.1037/0022-0663.98.3.554>
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383. doi:<http://dx.doi.org/10.1080/10888430701530730>
- Pollard-Durodola, S. D., Gonzalez, J. E., Simmons, D. C., Kwok, O., Taylor, A. B., Davis, M. J., . . . Simmons, L. (2011). The effects of an intensive shared book-reading intervention for preschool children at risk for vocabulary delay. *Exceptional Children, 77*(2), 161-183.

- Richardson Bruna, K., Vann, R., & Perales Escudero, M. (2007). What's language got to do with it?: A case study of academic language instruction in a high school "English Learner Science" class. *Journal of English for Academic Purposes*, 6(1), 36-54.  
doi:<http://dx.doi.org/10.1016/j.jeap.2006.11.006>
- Rivera, B.L. & Dickinson D.K., (April, 2013). Exploring the relationship between vocabulary learning, play, and narrative retell. In D.K. Dickinson (Chair), *The Mechanisms and Impact of Adult Support for Language Learning in Early Childhood: A Detailed Approach to Assessing Preschool Children's Depth of Word Knowledge*. The biennial conference of the Society for Research in Child Development, Seattle, WA.
- Roskos, K., & Burstein, K. (2011). Assessment of the design efficacy of a preschool vocabulary instruction technique. *Journal of Research in Childhood Education*, 25(3), 268-287.
- Roth, F. P., Speece, D. L., & Cooper, D. H. (2002). A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research*, 95(5), 259-272. doi:10.1080/00220670209596600
- Rowe, D. W. (1998). The literate potentials of book-related dramatic play. *Reading Research Quarterly*, 33(1), 10-35.
- Sameroff, A. J., & Mackenzie, M. J. (2003). Research strategies for capturing transactional models of development: The limits of the possible. *Development and psychopathology*, 15(03), 613-640.
- Schleppegrell, M. J. (2012). Academic language in teaching and learning: Introduction to the special issue. *The Elementary School Journal*, 112(3), 409-418.



- Seston, R., Golinkoff, R. M., Ma, W., & Hirsh-Pasek, K. (2009). Vacuuming with my mouth?: Children's ability to comprehend novel extensions of familiar verbs. *Cognitive Development, 24*(2), 113-124. doi:10.1016/j.cogdev.2008.12.001
- Shine Early Learning (2016). *Head Start curriculum*. New York.
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge Handbook of Literacy* (pp. 112-133). New York: Cambridge University Press.
- Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science, 29*(1), 41-78. doi:10.1207/s15516709cog2901\_3
- Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126*, 395-411. doi:http://dx.doi.org/10.1016/j.jecp.2014.06.003
- Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology, 43*(2), 454-464. doi:10.1037/0012-1649.43.2.454
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*(4), 381-398. doi:10.1207/s1532799xssr1004\_3
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development, 57*, 1454-1463.
- van Kleeck, A. (2014). Distinguishing between casual talk and academic talk beginning in the preschool years: An important consideration for speech-language pathologists. *American*

*Journal of Speech-Language Pathology*, 23(4), 724-741. doi:10.1044/2014\_AJSLP-14-0032

Vitevitch, M. S. (2008). What can graph theory tell us about word learning and lexical retrieval? *Journal of Speech, Language, and Hearing Research*, 51(2), 408-422.

Wasik, B. A., & Bond, M. A. (2001). Beyond the pages of a book: Interactive book reading and language development in preschool classrooms. *Journal of Education Psychology*, 93(2), 243-250.

Weisberg, D. S., Ilgaz, H., Hirsh-Pasek, K., Golinkoff, R., Nicolopoulou, A., & Dickinson, D. K. (2015). Shovels and swords: How realistic and fantastical themes affect children's word learning. *Cognitive Development*, 35, 1-14.

Wignell, P., Martin, J. R., & Eggins, S. (1989). The discourse of geography: Ordering and explaining the experiential world. *Linguistics and Education*, 1(4), 359-391.  
doi:http://dx.doi.org/10.1016/S0898-5898(89)80007-5

Yuan, S., & Fisher, C. (2009). "Really? She Blicked the Baby?": Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychological Science*, 20(5), 619-626.  
doi:10.1111/j.1467-9280.2009.02341.x

## Appendix A

### *Examples of Child Responses and Codes Assigned*

Target Word	Child Response	Information Unit(s) Coded For
Basket ( <i>n.</i> )	“You carry stuff with it.”	Function
Nostrils ( <i>n.</i> )	[Points to nostrils and sniffs]	Gesture
Handkerchief ( <i>n.</i> )	“Wipe your nose with handkerchief.”	Function
Shield ( <i>n.</i> )	“[A] shield protects you / when you get in a fight with a dragon and he blows fire at you.”	Function Meaningful Example
Charging ( <i>v.</i> )	“Run towards each other / when the knight charged at the dragon in the fight.”	Synonym Meaningful Example
Chuckling ( <i>v.</i> )	“A quiet laugh.”	Synonym
Fetching ( <i>v.</i> )	“I throw the ball to my dog and he fetches it and gives it to me.”	Meaningful Example
Returning ( <i>v.</i> )	“Run away and go back / in the story the farmer ran away and he never returned.”	Synonym Meaningful Example
Sobbing ( <i>v.</i> )	“You crying.”	Synonym

## Appendix B

*Form Class, Difficulty, and Means for Target, Exposure, and Control Words Tested on NWDT-M Measure*

Word	Form Class	Abstract or Concrete	Level of Instruction	Theme	Words Worth Teaching Difficulty Rating (Biemiller, 2010)	On Common Word List? (Dale & Chall, 1995)	Tier Two? (Beck & McKeown, 2007)	Pretest <i>M</i> (SD)	Posttest <i>M</i> (SD)
charging	verb	concrete	target	dragon	E	Y	Y	0.04 (0.23)	0.43 (0.71)
chuckling	verb	concrete	target	dragon	E	N	Y	0.05 (0.24)	0.62 (0.72)
emerging	verb	concrete	target	dragon	T2	N	Y	0.01 (0.10)	0.49 (0.81)
galloping	verb	concrete	target	dragon	E	Y	Y	0.25 (0.56)	0.64 (0.77)
handkerchief	noun	concrete	target	dragon		Y	Y	0.02 (0.20)	0.86 (1.07)
heel	noun	concrete	control	dragon	L2	Y	Y	0.00 (0.00)	0.00 (0.00)
nostrils	noun	concrete	target	dragon		N	Y	0.05 (0.31)	1.15 (1.21)
pliers	noun	concrete	control	dragon	D	N	Y	0.08 (0.33)	0.16 (0.49)
pond	noun	concrete	exposure	dragon	T2	Y	Y	0.83 (1.03)	1.00 (1.02)
quilt	noun	concrete	exposure	dragon	E	Y	Y	0.13 (0.41)	0.29 (0.74)
scales	noun	concrete	target	dragon		Y	Y	0.05 (0.20)	0.52 (0.88)
scowling	verb	concrete	control	dragon	T6	N	Y	0.02 (0.13)	0.03 (0.16)
servants	noun	concrete	target	dragon		Y	Y	0.24 (0.58)	0.75 (1.07)
shield	noun	concrete	exposure	dragon		N	Y	0.45 (0.93)	0.60 (0.93)
throne	noun	concrete	target	dragon	L2	Y	Y	0.06 (0.37)	0.77 (1.10)
valley	noun	concrete	target	dragon	E	Y	Y	0.02 (0.11)	0.57 (0.95)
accidentally	adjective	abstract	exposure	dragon	T2	Y	Y	0.32 (0.51)	0.30 (0.43)
celebration	noun	abstract	control	dragon	E	N	Y	0.64 (0.76)	0.81 (0.81)
enemies	noun	abstract	target	dragon	E	Y	Y	0.14 (0.39)	0.44 (0.62)
foolishness	noun	abstract	target	dragon		N	Y	0.08 (0.28)	0.21 (0.45)
intelligent	adjective	abstract	target	dragon	T2	N	Y	0.02 (0.20)	0.21 (0.46)
chimney	noun	concrete	target	farm	E	Y	Y	0.33 (0.67)	0.87 (0.82)
curtain	noun	concrete	control	farm		Y	Y	0.58 (0.71)	0.67 (0.80)
fetching	verb	concrete	target	farm		N	Y	0.39 (0.63)	0.67 (0.75)
field	noun	concrete	exposure	farm	E	Y	Y	0.46 (0.81)	0.58 (0.94)

Word	Form Class	Abstract or Concrete	Level of Instruction	Theme	<i>Words Worth Teaching</i> Difficulty Rating (Biemiller, 2010)	On Common Word List? (Dale & Chall, 1995)	Tier Two? (Beck & McKeown, 2007)	Pretest <i>M</i> (SD)	Posttest <i>M</i> (SD)
hedge	noun	concrete	control	farm	T6	N	Y	0.17 (0.42)	0.30 (0.64)
lane	noun	concrete	target	farm	E	Y	Y	0.07 (0.34)	0.48 (0.86)
plummeting	verb	concrete	control	farm		N	Y	0.01 (0.04)	0.00 (0.00)
sobbing	verb	concrete	exposure	farm	T2	Y	Y	0.01 (0.09)	0.11 (0.40)
stool	noun	concrete	exposure	farm	E	Y	Y	0.69 (0.98)	0.78 (0.90)
tip	noun	concrete	exposure	farm	T2	Y	Y	0.09 (0.34)	0.13 (0.31)
weeds	noun	concrete	target	farm	L2	Y	Y	0.22 (0.63)	0.76 (1.02)
wheelbarrow	noun	concrete	exposure	farm		N	Y	0.47 (0.88)	0.69 (0.98)
peaceful	adjective	abstract	target	farm	T2	Y	Y	0.17 (0.47)	0.46 (0.66)
plan	noun	abstract	exposure	farm	E	Y	Y	0.23 (0.49)	0.25 (0.49)
quarrel	noun	abstract	target	farm	T6	N	Y	0.00 (0.00)	0.27 (0.55)
recent	adjective	abstract	control	farm	T6	N	Y	0.00 (0.00)	0.07 (0.25)
returning	verb	abstract	target	farm	E	Y	Y	0.34 (0.51)	0.52 (0.63)
wearily	adjective	abstract	target	farm	T2	Y	Y	0.02 (0.18)	0.08 (0.32)

*Note.* E = words known by most children at the end of Grade Two; T2/L2 = high-priority (T2) and low-priority (L2) words known by 40 to 80 percent of children by the end of grade two; T6 = words known by 40 to 80 percent of children by the end of grade six; D = words known by fewer than 40 percent of children by end of grade six.

## Appendix C

### *Words Used in Intervention*

Word	Level of Instruction	Taxonomy or Theme	Units Taught In	Measure	Biemiller (2010) rating	On Common Word List? (Dale & Chall, 1995)
stem	target	theme	flower & vegetable	D/B	E	Y
bulb	target	theme	flower & vegetable	D/B	L	Y
seeds	target	theme	flower & vegetable	D/B	E	Y
root	target	theme	flower & vegetable	D/B	E	Y
soil	target	theme	flower & vegetable	D/B	E	Y
flowers	target	taxonomy	flower	D	E	Y
tulip	target	taxonomy	flower	B	D	Y
tiger lily	target	taxonomy	flower	D/B	L	N
daffodil	target	taxonomy	flower	B	L	N
iris	target	taxonomy	flower	D/B	N/A	N
hyacinth	target	taxonomy	flower	B	N/A	N
petals	target	theme	flower	D/B	N/A	N
sprouting	target	theme	flower	D/B	E	N
vegetables	target	taxonomy	vegetable	D	E	Y
radishes	target	taxonomy	vegetable	B	E	Y
leeks	target	taxonomy	vegetable	B	N/A	N
cauliflower	target	taxonomy	vegetable	D/B	L	N
eggplant	target	taxonomy	vegetable	D/B	N/A	N
artichoke	target	taxonomy	vegetable	B	N/A	N
raw	target	theme	vegetable	D	E	Y
vines	target	theme	vegetable	D/B	N/A	Y
blooms	exposure		flower	D	E	Y
summer	exposure		flower	B	E	Y
catalogs	exposure		flower	B	N/A	N
vitamins	exposure		vegetable	D/B	E	N
pod	exposure		vegetable	D/B	L	N
spade	exposure		vegetable	D/B	L	Y
frock	control			D/B	N/A	N
vase	control			D/B	E	Y
valley	control			D/B	E	Y
platter	control			B	N/A	Y
cavern	control			B	T	N
spectacles	control			B	N/A	N
scales	control			B	T	Y
throne	control			D/B	D	Y

*Note.* D=depth measure, B=breadth measure. E = known by most children at the end of grade 2; L/T = high priority (T) and low priority (T) words known by 40-80% of children by end of grade 2; D = known by fewer than 40% of children by the end of grade 6; N/A = does not appear in word list.

## Appendix D

### *Selected Transcripts from Play Sessions*

Language Support Strategy	Transcript
<i>Encourage use of new words</i>	Adult: What kind of <u>flower</u> do you want to grow? Child 1: This! [uses farmer toy to pick up yellow flower] Adult: What kind of <u>flower</u> do you think that is? Child 1: <u>Daffodil</u> .
<i>Expand child's utterance</i>	Child 2: We're missing something! We're missing something! Adult: What are we missing? Child 2: Water! Adult: Yeah, our <u>plants</u> need water to grow.
<i>Ask open-ended questions</i>	Adult: How do you know that's a <u>vegetable</u> and not a fruit? Child 3: Fruit is cold and <u>vegetables</u> is warm.

*Note.* Underlined words are those that were taught in the intervention.

## Appendix E

### *Teacher-Child Interaction Codes and Examples*

Code	Description	Examples
Didactic	Statements/questions intended to teach or transmit knowledge rather than respond to children's cues	Giving definition of a word Reading book Giving a definition and asking child to supply target word Asking children to repeat target word Labeling object during play or book-reading Using "teacher voice" to direct play
Responsive	Any use of a target word that responds to child cues (something child does or says)	Answering child's question (by using target word) Extending play scenario Expanding or recasting child utterance Repeating target word after a child During play, using character voice Narrating children's play actions Modeling play – playing "out loud"
Active processing	Asking children to synthesize or analyze word knowledge	Asking children for a definition of a word Asking about nuances of word meaning Asking open-ended questions about words



## Appendix F

### *Examples of Student Responses and Codes Assigned*

Target word	Student Response	Information Unit Coded for
Tiger lily	“Kind of flower.	Category information
	They’re orange.	Perceptual information
	Have spots on them /	Perceptual information
	and leaves.	Part
	They grow.”	Function
Eggplant	“It’s a vegetable, but it’s really a fruit.”	Category information
Vegetable	“You eat them.	Function
	Eggplant.”	Category information
Soil	“It’s dirt.	Synonym
	You can dig in it.”	Function
Roots	“Grow under the ground	Perceptual
	to help the flower.”	Function