Investigating the Efficacy of a Preschool Vocabulary Intervention Designed to Increase Vocabulary Size and Conceptual Knowledge

by

Julie C. Dwyer

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Educational Studies)
in The University of Michigan
2010

Doctoral Committee:

Professor Susan B. Neuman, Chair
Professor Joanne F. Carlisle
Professor Susan A. Gelman
Professor Kevin F. Miller

Acknowledgements

First I would like to extend my deepest thanks to my dissertation committee, including Drs. Susan B. Neuman, Joanne Carlisle, Susan A. Gelman, and Kevin F. Miller. Their efforts, comments, suggestions, questions, and support throughout this process were greatly appreciated. In addition, their great expertise and knowledge in their respective fields contributed a great deal to this dissertation. I would also like to thank Susan B. Neuman for serving as my mentor, advisor, boss, and dissertation chair throughout my years as a doctoral student and during the process of writing this dissertation. Her support has been extremely important and helpful throughout journey. I would also like to thank Susan Neuman's Ready to Learn research team for making this dissertation possible. Specifically, to Susan Neuman, Serene Koh, Tanya Wright, and Fran Lehr who were the other authors of the WOW curriculum. In addition, to Jeanne Friedel, Christina Mendez, and Deb Simon, whose organization and efforts were extremely helpful, and all of the research assistants who participated in data collection for their efforts and professionalism. I would also like to thank my friends Serena Salloum, Sarah Tucker, and Serene Koh for listening to my trials and tribulations during this process and providing sage advice. I'd like to thank my wonderful mom, dad, and sister for their tireless and unending support throughout this process. I would also like to thank the preschoolers who eagerly participated in the WOW curriculum, interacted with our research team, and ultimately have taught me a great deal. Finally, I would like to extend a special thank you

to the preschool teachers who participated in this study for their openness, enthusiasm, and their desire to improve the lives of young children.

Table of Contents

List of Tables

List of Appendices

Abstract

This dissertation study investigated the efficacy of a supplementary preschool embedded multimedia curriculum that was designed to increase one type of conceptual knowledge: taxonomic categories. Named the World of Words (WOW), this curriculum focused on teaching the properties and concepts associated with seven taxonomic categories and providing children with the vocabulary that is paired with these concepts. Participants were 178 low-income preschoolers (89 in the treatment group, 89 in the comparison group) taught by twelve preschool teachers. Six teachers implemented the supplementary curriculum with children in their classrooms four times per week for four months. The other six teachers went about their typical daily routines with for those four months, thus acting as a comparison group. Children in both the treatment and comparison groups were assessed on their conceptual knowledge and vocabulary knowledge before the study began, immediately following instruction around each taxonomic category, and at the conclusion of the study. Hierarchical Linear Modeling (HLM) was used to analyze data from these assessments. Findings indicated that children who experienced the curriculum acquired both the vocabulary and underlying concepts that were taught in the curriculum. In addition, they were able to transfer the conceptual knowledge they acquired to new learning situations, an initial indication that learning vocabulary embedded within the concept of taxonomic categories has the power to foster future learning. However, four months of this type of instruction was not sufficient to increase the growth of general receptive or expressive vocabulary knowledge. Results from this study support the notion

that teaching conceptual knowledge and vocabulary in taxonomic categories is a powerful way to increase vocabulary and conceptual knowledge before children enter formal schooling.

Chapter 1

Introduction

In 1997, Congress convened the National Reading Panel to "assess the status of research-based knowledge, including the effectiveness of various approaches to teaching children to read" (National Reading Panel, 2000). Among other important findings, this report highlighted scientific research demonstrating a critical link between early vocabulary knowledge and successful reading achievement. Both preceding this report and since its publication, a large body of research has highlighted this link (Cunningham & Stanovich, 1997; National Reading Panel, 2000; Scarborough, 1998; Senechal, Oulette, & Rodney, 2006). Longitudinal studies ranging from two to ten years show that early vocabulary knowledge predicts later reading achievement (Cunningham & Stanovich, 1997; Scarborough, 1998; Senechal et al., 2006). For example, in a meta-analysis of 61 studies, Scarborough (1998) found a significant mean correlation of .46 between kindergarten vocabulary and reading achievement one to two years later. In addition, researchers found a significant predictive relationship between vocabulary size in first grade and vocabulary size and reading comprehension in 11th grade, a full *ten years* later (Cunningham & Stanovich, 1997).

Because there is a clear relationship between early vocabulary knowledge and reading comprehension, many researchers have investigated and debated the nature of this relationship. Although it is likely that the nature of this relationship is complex and multifaceted, many seem to agree upon the "general knowledge" hypothesis, which posits that it is not the knowledge of the particular vocabulary words per se that promotes comprehension, but rather the knowledge of the concepts that words represent that fosters comprehension (Anderson & Freebody, 1981; Baumann, 2009; Stahl & Nagy, 2006). In other words, vocabulary represents underlying conceptual knowledge, and is this conceptual knowledge that ultimately contributes to reading comprehension. Importantly, this means that efforts to increase vocabulary (and thereby ultimately increasing reading comprehension and achievement) should begin with building rich stores of conceptual knowledge and teaching children the vocabulary words to communicate that knowledge.

Since vocabulary knowledge and the conceptual knowledge it represents is highly related to later reading achievement and school success, it follows that large vocabulary differences between individuals and groups contribute to achievement gaps. Disturbingly, there is a body of literature indicating that the quantity and quality of language input experienced by children differs according to family socioeconomic background (Hoff-Ginsberg, 1991, 1998; Snow et al., 1976). These differences lead to large gaps in early vocabulary knowledge between children from different socioeconomic backgrounds (Chall, Jacobs, & Baldwin, 1990; Denton, West, & Walston, 2003; Hart & Risley, 1995) and growth trajectory (Hart & Risley, 1995). These vocabulary gaps contribute to the unacceptable achievement gaps we see between children from advantaged and disadvantaged backgrounds in our nation's schools (Jencks & Phillips, 1998). Given these issues, it is heartening that a body of research demonstrates that vocabulary instruction can exert an influence on vocabulary knowledge and later reading ability for young, preliterate children (Beck & McKeown, 2007; Biemiller & Boote, 2006; Mol, Bus, & de Jong, 2009; Dickinson & Smith, 1994; Hargrave & Senechal, 2000; Lonigan &

Whitehurst, 1998; National Early Literacy Panel, 2009; National Reading Panel, 2000; Snow, Burns, & Griffin, 1998; Wasik, Bond, & Hindman, 2006; White, Graves, & Slater, 1990).

In sum, research has demonstrated large, early vocabulary gaps, the importance of early vocabulary and related conceptual knowledge for later school achievement, and the positive influence instruction can exert on vocabulary knowledge. Given these findings, one might expect to see a great deal of attention paid to vocabulary learning and teaching with our nation's youngest children. However, there is little evidence that this issue is treated with urgency in the education of young children. Research has demonstrated that there is a paucity of intentional, rich, explicit vocabulary instruction in primary grade curricular materials or classrooms (Beck, McCaslin, & McKeown, 1980; Blachowicz & Fisher, 2000; National Reading Panel, 2000).

Given the nature of the problem, this lack of attention in early elementary school materials and classrooms is concerning. Of equal concern is the lack of attention, even among researchers, to the vocabulary learning and teaching that actually occurs in preschool classrooms. Research that does exist, however, seems to parallel findings for the primary grades. For example, in their examination of early learning pre-kindergarten standards, Neuman and Roskos (2005) found that few states include specific guidelines regarding vocabulary learning and teaching. In addition, in an analysis of ten widely-used preschool literacy curricula, Neuman and Dwyer (2009) found little explicit guidance for preschool teachers about how to foster vocabulary development in their classroom.

The lack of attention to early vocabulary learning and teaching in preschool and primary grade classrooms and instructional materials is surprising, given what we know about the striking, early differences between economically advantaged and disadvantaged children, the importance of vocabulary and conceptual knowledge for later school success, and the power of instruction to influence vocabulary knowledge. If we are to narrow the vocabulary and concomitant conceptual knowledge gaps between children from different socioeconomic backgrounds prior to kindergarten entry, it is critical that we continue the study of instructional practices and interventions that can and will be implemented in early childhood classrooms.

The World of Words (WOW) Vocabulary Intervention

To address these issues, the Ready to Learn Project (created and overseen by principal investigator Susan B. Neuman) developed a supplementary preschool vocabulary intervention called *The World of Words* (WOW)© (Neuman, Dwyer, Koh, & Wright, 2007). The goal of this supplemental curriculum was to improve vocabulary and conceptual knowledge for low-income preschoolers in Southeastern Michigan. Funded from 2005-2010 by the U.S. Department of Education and the Corporation for Public Broadcasting System, and in collaboration with the University of Michigan, the Ready to Learn Project developed the research-based WOW vocabulary intervention, trained preschool teachers across Southeastern Michigan to implement the intervention, and evaluated the efficacy of the intervention.

Theoretical Background

The WOW vocabulary intervention was designed to teach conceptual knowledge and the vocabulary associated with that knowledge through a specific and targeted representational structure: in taxonomic categories. The rationale for this approach is based on a large body of research on the relationship between vocabulary and conceptual knowledge (Bloom, 2000; Borovsky & Elman, 2006; Carey, 2009; Gopnick & Meltzoff, 1986, 1987, 1992; Murphy, 2002; Waxman, 2004) and the human propensity for taxonomic categorization (Gelman, 1988; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987; Inhelder & Piaget, 1964; Markman, 1989, 1994; Markman & Hutchinson, 1984; Murphy & Lassaline, 1997; Smiley & Brown, 1979). Taken together, this research suggests that teaching children vocabulary and conceptual knowledge in taxonomic categories has the potential to not only increase conceptual knowledge and vocabulary, but provide a foundation that can foster future learning. This section describes the nature of taxonomic categories and summarizes the extant research that supports the approach taken in the WOW curriculum.

It has been well documented that preschool children are capable of, and often drawn to, organizing words and conceptual knowledge thematically (Inhelder & Piaget, 1964; Markman, 1989; Markman & Hutchinson, 1984; Smiley & Brown, 1979). Thematic groupings tend to be based on causal, spatial, or temporal relationships between things in the world rather than on any inherent similarities between items (Markman & Hutchinson, 1984). For example, there is nothing inherent in a cash register and a grocery cart that makes both a part of a grocery store, except that we find both objects there. Although this is a useful and necessary way to organize information in the world, learning new thematic relationships between words and concepts is reliant on actually experiencing or being explicitly taught about the thematic relations. This way of organizing words and concepts does not provide a conceptual structure that has the potential to be generative--facilitating induction, inference, and independence during future learning situations. For this reason, the WOW curriculum is based on teaching conceptual structures that *do* provide children with a self-extending system of learning: taxonomic categorization.

There is a great deal of research supporting the hypothesis that taxonomic categorization has the potential to serve as a self-extending system for future learning. Researchers and theorists have suggested that organizing the world, and the words that name that world, into taxonomic categories allows human beings to not only identify things in their environment and organize words and concepts efficiently in memory, but to draw inductive inferences to extend knowledge beyond the known (Bloom, 2000; Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Gelman & Medin, 1993; Murphy & Lassaline, 1997). The "inductive potential" (Gelman et al., 1998) inherent in this type of representation allows individuals to make inferences when encountering novel objects in the world (Bloom, 2000; Gelman, 1988), which is helpful in future learning situations.

To explain the hypothesis that this type of cognitive organization is efficient, facilitates induction, and supports future learning, consider an example. If a child possesses the concept of "insect", which is a taxonomic category, she will likely know the properties common to members of the category "insect" and several examples of an insect. When encountering a new type of insect, called "cicada", she can easily transfer her existing knowledge about the basic category of insect to the new item, rather than be required to learn each and every property of a cicada through time-consuming experiences with the insect. The inductive potential afforded her through her existing taxonomic representation of "insect" eases the process of learning about the new item. Prior knowledge about the category structured in the mind in a taxonomic way makes the process of learning about this new item more efficient. In this way, taxonomic knowledge easily begets more knowledge.

The WOW intervention is designed to teach children vocabulary associated with taxonomic categories to establish a foundation of conceptual knowledge organized in a way that fosters efficiency and future learning. This is important as children begin to move toward formal schooling and conventional literacy, because research shows that the prior knowledge that children possess has a great influence on what they understand, infer, and remember from texts that they read (Stanovich, 1986). Readers automatically and unconsciously relate what they already know about a topic to ideas in a text, which fosters comprehension of and inferences about new knowledge in that text (Anderson & Pearson, 1984; Pressley, 2000). If children have stores of information organized in ways that promote induction and inference, their prior knowledge will serve as foundation for new learning. In many ways, the WOW curriculum is designed to provide children with conceptual knowledge and vocabulary in a self-extending structure that will provide a foundation for future learning as they move into formal schooling and toward conventional literacy.

Though there is evidence that representing vocabulary in taxonomic categories has the potential to be generative and provide a self-extending system, there is a dearth of research investigating the efficacy of interventions that teach vocabulary in the context of taxonomic categories. This dissertation study is designed to fill this gap by investigating the efficacy of the WOW curriculum on vocabulary, conceptual knowledge, and transfer of that knowledge to new learning situations. This type of intervention may have the potential to narrow the unsettling and early vocabulary, conceptual, and achievement gaps we see between children from different social classes so that all children can begin school prepared for success.

Overview of Present Investigation

The purpose of the present investigation is to examine the potential of the WOW vocabulary intervention to increase vocabulary, conceptual knowledge, and ability to transfer that knowledge to new learning situations for low-income children. Evidence of ability to acquire and transfer knowledge to new learning situations will provide suggestive evidence that teaching vocabulary and conceptual knowledge in taxonomic categories has the potential to foster future learning.

This dissertation study uses data collected during an initial study of the efficacy of the WOW curriculum (under the aegis of principal investigator Susan B. Neuman and the Ready to Learn Project) conducted from January-May 2007 in 12 Head Start and Michigan State Readiness Program (MSRP) classrooms (treatment group = 89; comparison group = 89). The research questions addressed in this study are as follows:

Compared to children in the comparison group:

1. Do children who experience the curriculum learn the words that were taught?
2. Do children who experience the curriculum acquire the concepts that were taught?
3. Do children who experience the curriculum transfer learned concepts to new learning tasks?
4. Do children who experience the curriculum show more growth in general receptive and expressive vocabulary?

This research study empirically tests the hypothesis that the WOW vocabulary intervention has the potential to not only increase vocabulary size and conceptual knowledge but provide a foundation for future learning situations. The goal of this study is to determine if the WOW curriculum is a viable and effective intervention with the power to begin to narrow the vocabulary gap between children from low-income backgrounds and their more advantaged peers so that *all* children begin formal schooling with an excellent foundation for literacy learning.

Chapter 2

Literature Review

Vocabulary knowledge is an important component of reading; a reader must know the meaning of a substantial proportion of words in a text in order to comprehend what they are reading (Biemiller, 2003; Stahl & Nagy, 2006). Not only does an individual's vocabulary knowledge at any given time influence their comprehension at that time, a large body of longitudinal studies ranging from two to ten years in length show that *early* vocabulary knowledge is a strong predictor of later reading achievement (Cunningham & Stanovich, 1997; Scarborough, 1998; Senechal et al., 2006). One popular hypothesis about the nature this strong relationship between vocabulary and reading comprehension is called the "general knowledge" hypothesis which posits that vocabulary knowledge is an outward representation of stores of conceptual knowledge, and that it is the conceptual knowledge that supports and fosters comprehension (Anderson & Freebody, 1981). Importantly, this means that efforts to increase vocabulary must include instruction in the rich conceptual knowledge related to new vocabulary.

Early deficits in vocabulary and associated conceptual knowledge can contribute to achievement gaps and can have a profound influence on school achievement. A body of research demonstrates that there are large, early, and persistent vocabulary gaps between children from high socioeconomic backgrounds and their less advantaged peers (Denton et al., 2003; Hoff-Ginsberg, 1991, 1998; Pan, Rowe, Singer, & Snow, 2005; Snow et al., 1976). These differences contribute to achievement gaps between children from different socioeconomic backgrounds.

Given this issue, it is promising that there are a multitude of intervention studies that have demonstrated the positive and significant influence instruction can have on vocabulary size (Beck & McKeown, 2007; Biemiller & Boote, 2006; Lonigan & Whitehurst, 1998; Wasik et al., 2006). We might expect to see educators and educational publishers making concerted efforts to begin to address and narrow these vocabulary gaps through research-informed instruction, instructional materials, and interventions. Unfortunately, there is little indication that classroom practices or instructional materials include the amount or type of vocabulary instruction necessary to address these problematic vocabulary gaps (Blachowicz & Fisher, 2000; National Reading Panel, 2000; Neuman & Dwyer, 2009; Neuman & Roskos, 2005; Ryder & Graves, 1994). It is imperative that educational researchers continue to create and evaluate instructional techniques and interventions that are effective in increasing vocabulary development and conceptual knowledge and provide adequate support for teachers in actually implementing instruction around conceptual knowledge and associated vocabulary. The WOW vocabulary intervention was created to address these issues by providing an engaging and effective supplementary vocabulary and conceptual knowledge curriculum for preschool teachers to implement with low-income preschoolers.

In this chapter, I explore the complex issues described in the preceding paragraphs. I review the primary causes of early vocabulary gaps and what we know about why vocabulary gaps tend to persist during formal schooling, influencing reading comprehension and school achievement. I then summarize the extant research base on the influence instruction can have on vocabulary size in preliterate children and discuss the dearth of studies that focus on the actual conceptual knowledge that vocabulary represents. Next, I describe in detail the theoretical rationale for teaching vocabulary in the context of taxonomic categories. Finally, I describe how each of these literatures informed the design of the WOW vocabulary curriculum.

Part 1. The Origins and Persistence of Vocabulary Gaps

Individual Differences in Vocabulary Knowledge: Contributing Factors

We know that large individual differences exist in the language development of young children (Hart & Risley, 1995; Snow et al., 1998). Some have posited that these individual differences are attributable to inherited differences in language learning capacity (Scarr & Weinberg, 1978). However, with respect to vocabulary development, there is little evidence to support this claim (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Scarr & Wienberg, 1978). Scarr and Weinberg (1978) found that though there was a statistically significant relationship between maternal and child vocabulary development, the strength of this relationship was the same whether a child was adopted or biological. This finding suggests that for vocabulary development, environmental exposure to language is more important than innate, inherited language capacity (Scarr & Weinberg, 1978).

Since Scarr and Weinberg's (1978) study, researchers have produced ample and converging evidence that the quantity of language input is a critical contributor to vocabulary size and growth trajectory (Hart & Risley, 1995; Huttenlocher et al., 1991; Weizman & Snow, 2001). For example, in a study of 22 middle-class 14- to 26-month-olds, Huttenlocher and her colleagues (1991) found a significant relationship between the quantity of maternal speech directed toward a child and the child's vocabulary acquisition. In addition, in a landmark study, Hart and Risley (1995) observed 42 one- and two-year-old children from disparate socioeconomic backgrounds interacting with their families over the course of two and a half years as they learned to talk. Among other important findings, Hart and Risley (1995) found that the number of words to which a child was exposed was highly related to their vocabulary size at age three.

Research also indicates that the quality of the language input[1] that a child experiences during her preschool years is also a key contributor to vocabulary size and growth trajectory (Beals, 1997; Hart & Risley, 1995; Pan et al., 2005; Weizman & Snow, 2001). Studies have shown that children who were exposed to more sophisticated words and more supportive language interaction and exchanges around those words had significantly higher vocabulary than children who heard less sophisticated words (Beals, 1997; Weizman & Snow, 2001) Thus, it seems that *both* quantity and quality of language input during the preschool years influence vocabulary size and growth.

Considering the importance of quantity and quality of vocabulary input on vocabulary size and growth, it is sobering to consider a body of literature indicating that the quantity and quality of language input children experience differs according to family socioeconomic background (Hart & Risley, 1995; Hoff-Ginsberg, 1991, 1998; Snow et al., 1976). In their seminal study, Hart and Risley (1995) found large social class differences in the quantity and quality of language input experienced. Extrapolating from the data collected, the researchers estimated that by age four, children from advantaged backgrounds would have heard as many as 45 million words, while children from

[1] Though "quality of language input" was characterized differently across studies, it generally referred to the sophistication of adult vocabulary and the extent of adult supportiveness during adult-child language interactions.

disadvantaged backgrounds would have heard as few as 13 million words. The authors suggested that early differences in vocabulary size and growth by social class mean that initial class-based vocabulary differences are likely to increase over time. Several other studies have supported Hart and Risley's (1995) findings, documenting the profound and ever-widening differences in vocabulary knowledge between children from advantaged and disadvantaged social class backgrounds (Denton et al., 2003; Pan et al., 2005).

Persistent Vocabulary Gaps in School

Not only are there large differences between children in vocabulary size before school entry, but these vocabulary gaps persist and perhaps increase throughout formal schooling (Chall et al., 1990, Denton et al., 2003). One theory that may explain this finding is that the more vocabulary knowledge an individual possesses, the more new words they are capable of learning (Nagy & Scott, 2000). Children who have larger vocabularies at a younger age will find it easier to acquire new vocabulary, and will thus learn words more quickly than children who begin school with less vocabulary knowledge. Thus, vocabulary gaps are likely to increase over time. On the other hand, Biemiller and Slonim (2001) found that though there are large differences through grade two between children from different social class backgrounds in vocabulary growth rate and size, after second grade children with the highest and lowest vocabulary knowledge learned words at the same rate. From this perspective, persistent vocabulary gaps in school can be partially attributed to early differences in vocabulary size and growth rate.

Another theory about why vocabulary gaps persist or widen over time relates to the relationship between reading and vocabulary. There seems to be a reciprocal relationship between reading and vocabulary; vocabulary knowledge influences reading comprehension, while wide reading influences vocabulary knowledge (Stanovich, 1986). Consider the child who struggles with reading. This child will tend to have a smaller vocabulary and read less than their peers because the task is so difficult and belabored. The less a child reads, the slower their vocabulary and related conceptual knowledge will develop; slowly developing vocabulary and knowledge will inhibit further growth in reading, and so on. In contrast, children who read well tend to have larger vocabularies (and associated conceptual knowledge), read more, acquire more and more vocabulary and knowledge from reading, become better readers as a result, and so on. This process is known as the "Matthew Effect", based on the proverb "the rich get richer, the poor get poorer" (Merton, 1968; Stanovich, 1986). This phenomenon likely contributes to increases in vocabulary gaps over time.

The Matthew Effect (Merton, 1968; Stanovich, 1986) becomes particularly pronounced as children begin to encounter texts containing more academic and sophisticated vocabulary representing concepts that are beyond their current knowledge of the world (Chall & Jacobs, 2003). According to Chall (1996), at a certain point children begin to use reading "as a tool for learning" about the world; simultaneously, the adults around them begin to expect that they will acquire new information through reading. Importantly, to acquire new knowledge from these more difficult texts, children must already know a substantial proportion of the words and associated concepts in the text (Stahl & Nagy, 2006). For children who have limited vocabulary and conceptual knowledge, comprehending texts replete with highly academic language can be an overwhelming and ultimately an unsuccessful task (Becker, 1977).

Another explanation for why vocabulary gaps persist during formal schooling is related to deficits in instructional materials and instruction. Studies of elementary literacy curricula suggest that there is little attention to vocabulary in these programs (Beck et al., 1980). For example, in an investigation of the extent of vocabulary instruction included in 10 widely-used preschool literacy curricula, Neuman and Dwyer (2009) found little evidence of attention to vocabulary learning and teaching. When investigating the *enacted* curriculum, studies corroborate this finding; there is a paucity of intentional, rich, explicit instruction in vocabulary development in the primary grades (Blachowicz & Fisher, 2000; National Reading Panel, 2000). There has been no investigation of the amount of vocabulary learning and teaching actually occurring in the enacted curricula in preschool classrooms. The dearth of instruction in classrooms or research agendas investigating this issue suggests that there is a lack of urgency in response to what we now know is a critical problem and substantial contributor to achievement gaps**.**

In sum, vocabulary gaps between children from different social class background exist as early as three years of age (Hart & Risley, 1995). Initial class-based differences in vocabulary size and growth trajectory tend to be maintained throughout formal schooling, possibly as a result of one or a combination of the following phenomena: slower vocabulary learning due to Matthews Effects (Stanovich, 1986), similar rates of vocabulary learning from very different beginning points (Biemiller & Slonim, 2001), and lack of exposure to vocabulary instruction (National Reading Panel, 2000). Irrespective of the cause, the persistence of vocabulary gaps that often are fully formed by age three suggests that our schools tend to have little influence on vocabulary size or growth trajectory. Thus, early vocabulary gaps have the potential to further exacerbate already unacceptable achievement gaps between children from advantaged and disadvantaged backgrounds. The WOW curriculum was designed to address this issue by providing at-risk preschoolers with an intervention designed not to increase vocabulary size and conceptual knowledge in a way that will foster future learning.

Part 2. The Influence of Vocabulary Instruction

It is unacceptable that there are large, early, and persistent differences in vocabulary knowledge between advantaged and disadvantaged children. To begin to address these gaps, it is critical that researchers and educators consider how to provide effective vocabulary instruction. This section includes a brief review of the current body of research on the effectiveness of vocabulary instruction in preliterate children, focusing on instructional approaches and techniques that have been deemed effective. It also describes the paucity of instructional approaches and techniques that carefully consider the content of the vocabulary that is being taught and the way that the WOW curriculum fills this gap.

Studies of Vocabulary Instruction with Preliterate Children

Although it is disheartening that left unchecked, early vocabulary gaps often persist or continue to grow wider as children move through our formal schooling system, there is a great deal of evidence that teachers and instruction can exert an influence on vocabulary knowledge and later reading ability for very young, preliterate children (Beck & McKeown, 2007; Biemiller, 2006; Dickinson & Smith, 1994; Mol, Bus, & deJong, 2009; National Early Literacy Panel, 2009; National Reading Panel, 2000; Snow et al., 1998; White et al., 1990). The vast majority of research studying the influence of instruction on vocabulary knowledge has focused on the influence of storybook reading at home and at school because written language is a demonstrable source of sophisticated vocabulary (Hayes & Ahrens, 1988).

In general, these studies have shown that storybook reading, and interactions with parents and teachers around storybook reading, is an effective means of increasing children's vocabulary (Beck & McKeown, 2007; Biemiller & Boote, 2006; Brabham & Lynch-Brown, 2002; Bus, van IJzendoorn, & Pelligrini, 1995; Elley, 1989; Hargrave & Senechal, 2000; National Reading Panel, 2000; Penno, Wilkinson, & Moore, 2002; Scarborough & Dobrich, 1994; Senechal, 1997). These studies, and others, have shown that there are several instructional elements, when coupled with storybook reading, that are effective in increasing the vocabulary of preliterate children. These include: questioning (Blewitt et al. 2009; Senechal, 1997), explanation or definition of target vocabulary during reading (Biemiller & Boote, 2006), repeated exposures to vocabulary in multiple contexts (McKeown et al., 1985), review of target vocabulary (Biemiller & Boote, 2006), actively involving children in the read-aloud (Beck & McKeown, 2007), or some combination of these instructional elements (Beck & McKeown, 2007; Brabham & Lynch-Brown, 2002; Hargrave & Senechal, 2000; Lonigan & Whitehurst, 1998; Mol et al., 2009; Wasik et al., 2006).

Gap in the Vocabulary Instruction Literature

The body of literature on the influence of story book reading, and interaction around story book reading, clearly demonstrates that it is an effective and rich context for improving target vocabulary knowledge of preliterate children. The majority of these studies focused on the efficacy of instructional methods to teach vocabulary during storybook reading, which is certainly important. However, very few of these studies focused equally on the *content* that was taught. In other words, little attention was paid to the concepts associated with the vocabulary chosen or conceptual relationships between the vocabulary words. Rather, words were chosen based primarily on two criteria a) their presence in a storybook (Beck & McKeown, 2007; Biemiller & Boote, 2006) and b) their difficulty level (Beck & McKeown, 2007).

There have been a few notable exceptions to this trend of focusing only on word difficulty when choosing words to teach, rather than focusing also on the underlying concepts associated with the vocabulary and the relationships between words that are taught. One of these exceptions was the well-known intervention by Wasik, Bond, and Hindman (2006), who prepared 22 prop boxes that contained books and objects related to a theme or topic often seen in preschool classrooms (e.g. "the seasons", "gardening", or "welcome to school"). Words from the theme-related books were chosen because they were related to the theme, were thought to be unknown to children, and were necessary for comprehension of the story. Where possible, the prop box contained objects representing the same target vocabulary. For example, a prop box for "gardening" contained a shovel, flowers, a carrot, corn stalk, a small garden hose, insects, seeds, and a rake. Items in the prop boxes were used before and during book reading to introduce and familiarize children with the target vocabulary. Children were given opportunities to hear and use target, theme-related vocabulary through open-ended question during book reading, during extension activities, and through interaction with simple labeling books. Children who experienced this intervention learned the target vocabulary *and* performed significantly better than children in the control group on general vocabulary measures.

In sum, these researchers intentionally chose and taught words that were part of a concept (e.g. “gardening”), and more specifically were related in thematic categories.

Another example of a vocabulary intervention that attended to conceptual knowledge and relationships between vocabulary words was Spycher’s (2009) investigation of a five- week vocabulary intervention focusing on the concept “Insect”, which was a pre-existing curricular unit in the classrooms she studied. The 20 target vocabulary words taught in this intervention were chosen because they were semantically related to the concept of insects (specifically, the words that were taught were thematically related to insects; e.g. escape, hatch, larva, metamorphosis, pupa, pollen, nectar, etc.), were included in the informational books that came with the science curriculum, were related to state standards, and were considered by the author to be “high-utility academic words” (or Tier 2 or 3 words as defined by Beck and her colleagues (2002)). Conducted with monolingual, bilingual, and English-learning kindergarteners, all children experienced the science curriculum and associated read alouds. However, children in the treatment group also received explicit instruction around the 20 related target words. Spycher (2009) found that children in the treatment condition learned more of the target academic vocabulary than children who were simply exposed to the vocabulary during read alouds. She also found that when asked to discuss science concepts, children in the treatment condition used more of the target academic vocabulary than children in the control group. In addition, children who acquired more of the target vocabulary were better able to express their understanding of the target science concepts.

In a departure from the preponderance of vocabulary studies that choose and teach words that are unrelated to one another conceptually, the interventions designed by Wasik et al. (2006) and Spycher (2009) both chose and taught words that were semantically related to one another through a larger concept. Though neither of these studies organized vocabulary in taxonomic categories, as is done in the WOW curriculum, it is promising that interventions that teach children words in semantic clusters (thematically, in both studies mentioned here) have the potential to foster vocabulary development.

In sum, though little is being done in our nation's classrooms to increase vocabulary knowledge (National Reading Panel, 2000), empirical studies of interaction around book reading have demonstrated that specific types of vocabulary instruction can powerfully influence preliterate children's vocabulary development (Biemiller & Boote, 2006; Mol et al., 2009; Senechal, 1997). However, the bulk of these studies have chosen vocabulary based on the level of difficulty rather than considering the conceptual knowledge represented by the words taught (with a few notable exceptions, outlined above). In designing the WOW curriculum, the Ready to Learn research team borrowed the instructional techniques empirically demonstrated as effective from the story book literature, but advances the field by embedding that instruction in a framework that focuses on deep conceptual knowledge development.

Part 3. Rationale for Teaching Vocabulary and Conceptual Knowledge in Taxonomic Categories

The WOW curriculum was designed to address the need for scalable interventions that increase vocabulary and conceptual knowledge in preliterate children. One critical aspect of this intervention that sets it apart from the majority of vocabulary interventions studied to date is a focus on conceptual knowledge development. Vocabulary was taught in this intervention by presenting vocabulary words in the context of richly structured concepts, specifically taxonomic categories. This section focuses on the theory and research supporting this approach. I first define conceptual knowledge, specifically situating taxonomic categories as one type of concept available to human beings. I describe the benefits for future learning that are associated with taxonomic categorization. In addition, I outline research on the important relationships between conceptual knowledge and vocabulary knowledge and more specifically between taxonomic categorization and vocabulary knowledge. Finally, I synthesize the research to make an argument for teaching vocabulary within the context of taxonomic categories, the approach taken in the WOW curriculum.

Conceptual Knowledge

Because one of the primary goals of the WOW curriculum was to increase conceptual knowledge through teaching vocabulary in taxonomic categories, it was critical to consider the nature of conceptual knowledge. This section explores different definitions of conceptual knowledge, how categories relate to conceptual knowledge, and specifically on the benefits and efficiency of taxonomic categories.

Definition of Conceptual Knowledge. Different theorists have defined the term "concept" in different ways. At the most broad level, most agree that concepts are mental representations or ways of representing and storing experiences in memory and that these mental representations help us to interpret our experiences in the world (Gelman & Kalish, 2006; Murphy, 2002). For example, Gelman & Kalish (2006) state that concepts are mental representations that are “efficient…obviating the need to track each and every individual interaction or encounter”. Murphy (2002) states that we use concepts, or mental representations, to understand our experiences of the world. He presents the example of an individual who encounters a particular tomato; He can rely on his existing concept of “tomato” to determine that this tomato is probably like other tomatoes he has eaten before and is therefore edible, even though he has never before seen that particular tomato.

Although both Gelman and Kalish (2006) and Murphy (2002) agree that concepts are mental representations that help us interpret the world, they seem to disagree on the scope of what constitutes a concept. Murphy (2002) posits that concepts are mental representations that correspond to categories, or classes of objects in the world. In many ways, he talks about concepts and categories synonymously. In contrast, Gelman and Kalish (2006) take a somewhat broader view of concepts. They propose that there are many types of concepts available to human beings, including categories, properties, events or states, individuals, and abstract ideas. Though they concede that categories may have a fundamental role in human cognition, unlike Murphy (2002), they do not use the term concept and category interchangeably. Rather, they see categories as one of several types of concepts human beings use to interpret their experiences in the world (Gelman & Kalish, 2006).

In sum, researchers seem to agree that concepts can be defined as mental representations. However, not all researchers agree on the scope of what should be considered a concept. In this dissertation study, I adopt Gelman and Kalish’s (2006) broader stance that there are many different types of concepts that human beings use to understand their experiences. The WOW curriculum and this dissertation study focus on one of the possible type of concepts available to both adults and children: categories. The next section describes what we know the benefits of two types of categories, thematic and taxonomic categories.

Thematic Categories. According the certain theorists (Gelman & Kalish, 2006), thematic categories are one form of conceptual knowledge. Thematic categories are concepts that include items or events that are related causally or by their co-occurrence in space and/or time, rather than any common inherent characteristics (Bauer & Mandler, 1989; Markman & Hutchinson, 1984). Fire truck and fireman are both part of the thematic category "fire station" and are related to one another thematically because they are both found in fire stations and have to do with fighting fires. Thematic relationships are a very real and important part of understanding the world.

There is a body of research on children's classification of objects that suggest that very young children find thematic relationships between objects in their world highly salient (Inhelder & Piaget, 1964; Markman, 1989; Markman & Hutchinson, 1984; Smiley & Brown, 1979). This contention is based on classification and sorting studies that demonstrated that preschoolers (children under six years of age) tended to organize objects based on spatial or proximal relationships, while school-aged children were likely to group objects together based on common characteristics (Inhelder & Piaget, 1964; Smiley & Brown, 1979). This evidence suggests that young children are interested in thematic relationships, are fully capable of understanding them, and have developed concepts that are organized on the basis of thematic relationships.

Taxonomic Categories. Taxonomic categories are another form of conceptual knowledge. Taxonomic categories are concepts that include items that share inherent properties (Markman, 1994). Further, taxonomic categories are linked to one another in a hierarchical way (moving from superordinate, to basic, to subordinate) such that every property shared by members of a category is also true of subordinates of that category (Murphy & Lassaline, 1997). For example, both Golden Retrievers and Bulldogs are types of dog, and all dogs bark, therefore both Golden Retrievers and Bulldogs bark.

Many developmental psychologists believe that human beings naturally organize words and concepts into hierarchical or taxonomic categories at the superordinate, basic, and subordinate levels (Murphy & Lassaline, 1997). However, for many years, some psychologists and researchers believed that organizing conceptual knowledge in such a sophisticated way was beyond the capacity of preschoolers, who were seen as highly reliant on perceptual cues and surface relationships between objects (Bauer & Mandler, 1989; Gelman & Wellman, 1991; Inhelder & Piaget, 1964; Smiley & Brown, 1979). Many believed that there was a developmental shift from a bias in young children toward thematic, relational associations between objects to a focus on taxonomic relationships and categories in older, school-aged children (Bauer & Mandler, 1989; Inhelder & Piaget, 1964; Smiley & Brown, 1979). More recent research has suggested that this is not the case; a body of literature suggests that even very young children are capable of understanding taxonomic relationships between items (Bauer & Mandler, 1989; Gelman, 1988; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987). In a series of studies Gelman and colleagues (Gelman, 1988; Gelman & Coley, 1990; Gelman & Markman, 1986, 1987) demonstrated that children as young as age 2 ½ are able to look beyond perceptual information when thinking and reasoning about categories. Further, Bauer and Mandler (1989) found that children as young as 16-31 months of age are capable of sorting items taxonomically. This evidence suggests that even very young children have developed concepts that are based on taxonomic relationships.

Taxonomic cognitive structures are considered efficient, fostering future learning by promoting inductive inferences (Murphy & Lassaline, 1997). Take the example of the word dog. If a child knows the properties common to dogs (e.g. they bark, they eat dog food, they are pets, they have four legs) and she encounters a new type of dog called a Pomeranian, she can infer that the new dog possess all of the properties that other dogs share. Thus, once she finds out that this new object is a dog, she immediately knows a great deal about this animal (Murphy & Lassaline, 1997). In effect, she can easily transfer her existing knowledge about the basic category of dog to the new item rather than be required to learn each and every property of a Pomeranian through time-consuming experiences with the animal.

In sum, we know that concepts are mental representations that help us to interpret our experiences in the world, that categories are one type of concept available to human beings, and that taxonomic categories offer inductive potential to foster future learning. Because the WOW curriculum and this dissertation study focus on both concepts and vocabulary, at this point I turn to what we know about how concepts and vocabulary are related to one another.

The Relationship Between Conceptual Knowledge and Language

Many researchers stress the importance of the links between conceptual knowledge and vocabulary. Carey (2009) stated "the lexicon stands right in the middle of any theory of conceptual development". Similarly, Waxman (2004) said "Word learning, more than any other development achievement, stands at the very center of the crossroad of human cognition and language". This section describes the relationship between conceptual knowledge and vocabulary.

General Relationship Between Concepts and Vocabulary. The words *concept* and *word* are often used interchangeably; for example, when we say that a child is using the word *habitat* correctly, we take that as evidence that the child has the concept of habitat (Murphy, 2002). However, though the terms *concept* and *word* are closely related to one another, they are not the same (Bloom, 2000; Clark, 1983; Heit, 1997; Kiel, 1989, 1991; Murphy, 2002). Researchers tend to define *concepts* as "nonlinguistic psychological representations", while *words* or *word meanings* are defined as being mapped onto concepts (Bloom, 2000; Murphy, 2002; Waxman, 2004).

Though *concepts* and *words* are different by definition, they are inextricably linked to one another throughout development. We have concepts to represent our knowledge of the world and relations among things in that world. Those concepts provide us with important information as we interact with and experience people, objects, and events in the world. But in order to effectively communicate with others, people must also acquire the words that are associated with those concepts (Murphy, 2002; Waxman, 2004). Though this example presents this relationship as uni-directional, it is also the case that individuals may learn a word and only then learn the concept that underlies that word (Bloom, 2000; Keil, 1989; Murphy, 2002).

In natural word-learning situations, the acquisition of concepts and words are linked to one another in an interactive way and these relationships are refined over time through experience with language and the world. As children's understanding of concepts develop, word meanings have to change and become more sophisticated to account for these changes. On the other hand, as children learn more words and more nuanced information about words, concepts must necessarily change to encompass this new learning (Murphy, 2002).

Specific Relationship Between Conceptual Categories and Language Learning. We know that concepts and words are linked to one another. More specifically, some researchers have theorized that there is a special link between word learning and one kind of concept: taxonomic categories (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1986, 1987, 1992; Markman, 1989; Markman & Hutchinson, 1984). Some researchers (Bloom, 2000) contend that for both very young children and adults, many words function as a means to classify the environment. When children learn the word "dog", they are really using a classification process to categorize all instances of "dog". In order to use this word appropriately, an individual must understand the "conditions underlying category membership" (e.g. all dogs have four legs, bark, and have fur). Thus, as children learn certain words, in many cases they are really learning the general term for concepts that are organized taxonomically. As they come to know more about their world, children learn more precise categories and subcategories of items and actions that are in that world (Bloom, 2000).

There is empirical evidence that suggests that there is, in fact, a bi-directional and mutually beneficial relationship between categorization and language learning (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1986, 1987, 1992). Based on the observation that several important cognitive and linguistic developments occur simultaneously around 18 months of age (including the naming explosion, object permanence, means-end understandings, and relatively sophisticated categorization abilities), Gopnik and Meltzoff (1986, 1987, 1992) conducted a series of studies to investigate if and how these linguistic and cognitive achievements are related during the second year of life. In one study (Gopnick & Meltzoff, 1987), these authors investigated the relationship between sophistication of categorization and the onset of the naming explosion, which they hypothesized are both reliant on the ability to place objects into categories. In a series of categorization tasks over several months, they found evidence that the onset of the naming explosion and ability to engage in sophisticated categorization were highly related (r = .78). In a subsequent study (Gopnick & Meltzoff, 1992), they found that there was a relationship between lexical development and categorization of both identical and basic-level items; children who were better able to categorize either way tended to have larger vocabularies. Taken together, results of their studies suggested that there is an interplay between the conceptual understandings involved in categorization and vocabulary development. However, this series of studies was not able to determine the direction of the relationship or determine if one skill precedes the other; as Gopnik & Meltzoff (1992) stated "It is possible that the conceptual abilities involved in object sorting provide support for the learning of names. However, it is also possible that learning that all things can be named helps children to discover that all objects can be sorted".

More recently, Borovsky and Elman (2006) hypothesized that the ability to learn new words and category structure are related in a "synergistic" way; that language first influences category and conceptual structure, and that this structure in turn influences word learning. Using computer simulations, they investigated how language input mediates the relationship between early taxonomic categorization and vocabulary development. They found that as language input increased, there were positive increases in *both* vocabulary acquisition and category "coherence" (they defined more coherent categories as better specified), which they posited was evidence that category learning and vocabulary acquisition are related. Further, they found that in cases where categories were better specified and more coherent, word learning was improved because the computer-simulated networks were then able to use existing knowledge of category members to generalize to new items.

Importantly for this dissertation study, Borovsky and Elman (2006) suggest that disadvantages related to a paucity of rich, early language input may persist well beyond infancy because deficient input has implications for *both* vocabulary and conceptual development. The authors suggest "Perhaps focused training on category development may boost word learning ability in these children (in impoverished linguistic environments) that could at least partially make up for deficiencies in language experience by aiding them to make the most efficient use of language they do hear".

Taken together, research suggests that even very young children are capable of looking beyond their immediate perceptions to attend to important properties that denote membership in taxonomic categories (Gelman, 1988; Markman, 1989). In addition, it appears that there is a bi-directional and synergistic relationship between children's ability to use their conceptual knowledge to categorize taxonomically and vocabulary acquisition (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1986, 1987, 1992).

Rationale for Focusing Instruction on Taxonomic Categories

Both taxonomic and thematic categories are important and necessary for understanding the world; Waxman and Gelman (1986) note that because “classification is flexible, we can exploit different relations among objects, depending on the task at hand”. Thematic categories are important for some types of relationships and in some instances. Taxonomic categories are equally important for other types of relationships and in other instances.

Traditionally, preschool curricula have been based on thematic concepts or “themes”, such as “My Neighborhood”, “All About Me”, “Friendship”, “Fall”, or “Our Great Big World”. Organizing information in this way is a useful and necessary part of learning and understanding our world. It is useful to help children develop a concept of a *neighborhood* and to understand that neighborhoods often include houses, people that live in those houses, schools, and stores that people need. Items associated with this thematic concept are related to one another through causal, spatial, or temporal relationships rather than any inherent similarity. Though this concept is valuable, learning more about what constitutes the concept *neighborhood* is reliant on actually experiencing it or being explicitly told that the item belonged in a neighborhood.

I am not arguing for the discontinuation of teaching concepts that have a thematic basis in preschools. However, I am arguing that teaching concepts and vocabulary organized in taxonomic categories has the potential to be a more powerful means of improving conceptual knowledge and associated vocabulary for the following three reasons: a) human beings naturally categorize items in the world in taxonomic categories (Bloom, 2000; Murphy & Lassaline, 1997) , b) research indicates that word learning and taxonomic categorization are inextricably linked, related bi-directionally, and are mutually beneficial (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1986, 1987, 1992), and c) because this representational structure offers inductive potential that has the potential to be generative and can foster future learning (Gelman, 1988; Gelman & Markman, 1986, 1987). Therefore, the WOW vocabulary curriculum was intentionally designed to teach children words and concepts in taxonomic categories. Teaching words and concepts this way, in contrast to teaching words and concepts thematically, provides a cognitive architecture that has the potential to be self-extending and generative as children move toward formal schooling and toward conventional literacy--facilitating induction, inference, and independence during future learning situations. This dissertation study is designed to empirically investigate this premise.

Part 4. Synthesizing the Literature to Build the WOW Intervention

This chapter has summarized the theoretical and empirical literature bases for the WOW vocabulary intervention. Large and persistent vocabulary gaps between children from disparate economic backgrounds are currently not being adequately addressed in our nation's classrooms. We know, however, that instructional practices hold a great deal of promise in ameliorating these problems. The Ready to Learn research team created the WOW vocabulary intervention to address these issues.

Focus on Taxonomic Categories. Based on the extant theory and research summarized in this chapter, the WOW curriculum was designed to increase conceptual knowledge, and vocabulary words related to that conceptual knowledge, in taxonomic categories. The rationale for this approach is thoroughly outlined in the previous section, but generally is based on the notion that teaching children concepts that are taxonomic and richly-structured, and teaching the words associated with those concepts, has the potential to foster future learning (Borovksy & Elman, 2006).

Focus on Instructional Techniques. The theoretical approach adopted in the WOW curriculum provides a framework for the inclusion of specific, research-based instructional practices known to increase vocabulary. These practices include interactive book reading (Mol et al., 2009), questioning (Blewitt et al. 2009; Senechal, 1997), explanation or definition of target vocabulary (Biemiller & Boote, 2006; Stahl & Fairbanks, 1986), repeated exposures to vocabulary in multiple contexts (Beck et al., 1985; Stahl & Fairbanks, 1986), review of target vocabulary (Biemiller & Boote, 2006), active involvement of children (Beck & McKeown, 2007; Brabham & Lynch-Brown, 2002; Hargrave & Senechal, 2000; Lonigan & Whitehurst, 1998; Mol et al., 2009; Wasik et al., 2006), and deep thinking about vocabulary (Stahl & Fairbanks, 1986).

Importantly, the inclusion of these instructional techniques in the WOW curriculum contributes to the extant literature in two ways. First, the bulk of studies investigating these instructional techniques used one specific instructional context: interactive book reading. The WOW curriculum extends the literature by embedding these instructional techniques mainly within teacher-child language interactions. Secondly, these studies were focused on the efficacy of specific instructional techniques in increasing children's surface vocabulary knowledge. These studies did not endeavor to investigate if these instructional techniques had an influence on the depth and richness of children's underlying conceptual knowledge. The WOW curriculum diverges from these approaches and contributes to the literature by using proven instructional techniques in the context of teacher-child language interactions and by assessing children's understanding of the concepts that undergird the vocabulary that is taught.

Conclusion

In sum, we know that early vocabulary and conceptual knowledge are highly related to later learning and school success and therefore large differences between children from different social class backgrounds contribute to later achievement gaps. Research suggests that there are promising ways to increase vocabulary in very young children, but there is little evidence that these findings have influenced practice in our nation's preschools. In an effort to investigate ways to begin to close the large vocabulary and concomitant conceptual knowledge gaps that we know exist, the WOW vocabulary intervention is designed increase conceptual knowledge and vocabulary size through a taxonomic structure that has the potential to foster future learning.

Chapter 3

Method

Purpose

The purpose of this study was to conduct an initial[2] investigation of the influence of the WOW curriculum on vocabulary knowledge, conceptual knowledge, and transfer and inference in new learning situations of low-income children preschoolers. If children who learn words and concepts in taxonomic categories are able to apply their knowledge to new items, it suggests that they may be able to use this type of knowledge to foster future learning and comprehension. To investigate if the WOW curriculum was effective, cognitive data were collected on both children in classrooms where the WOW curriculum was administered and children in comparison classrooms. In addition, demographic data were collected to control for differences between the two groups. The research questions addressed in this study are as follows:

Compared to the comparison group:

1. Do children who experience the curriculum learn the words that were taught?

[2] I characterize this study as an initial investigation into the efficacy of this curriculum because it represented the first effort of the research team to implement this curriculum in a classroom setting, but it was much too large to be characterized as a pilot study. After this initial investigation, the curriculum was expanded to cover an entire academic year of instruction and during the following academic year, the expanded curriculum was implemented in more preschool classrooms. The resultant data from that implementation were not used for this dissertation study for three reasons: a) I was deeply involved in the initial implementation of the curriculum described in this dissertation study, but was not involved in the implementation the following year, b) the data from the second implementation were not available to me when I began writing my dissertation, and c) the second implementation did not include the "Picky Peter" tasks, and data from this task is integral to my dissertation study.

2. Do children who experience the curriculum acquire the concepts that were taught?
3. Do children who experience the curriculum transfer learned concepts to new learning tasks?
4. Do children who experience the curriculum show more growth in general receptive and expressive vocabulary?

This chapter reports the methodology used in this dissertation study. First, this section presents detailed information about the teachers and children included in the sample. In addition, it includes a comprehensive description of the World of Words (WOW) vocabulary curriculum, including how taxonomic categories and associated words were chosen and taught. Next, this chapter includes descriptions and psychometric properties of all vocabulary and conceptual knowledge measures used in this study. Then, it includes an account of how missing data was treated and outlines the procedure (including teacher training, implementation of the WOW intervention, and how fidelity of implementation was measured). Finally, this chapter delineates the analytic method.

Sample

To investigate the research questions posed above, under the direction of Dr. Susan Neuman, the Ready to Learn Project recruited Head Start and Michigan School Readiness (MSRP) programs in two southern Michigan cities to participate in the WOW curriculum study from January-May 2007. Head Start and MSRP programs were chosen because they serve economically and otherwise disadvantaged preschoolers, which is the target population for this intervention study. To qualify for Head Start, children must be between three and five years old and live in a family with an income at a proportion below the poverty line. To qualify for MSRP programs, children must be four years old and have at least two of 25 varied risk factors (which include low family income, low birth weight, developmental immaturity, history of physical abuse, chronic illness, high rate of mobility, nutritional deficits, ESL status, etc.). In addition to including children with multiple risk factors, MSRP programs are also required to include 50% low income children.

Teachers. Head Start and MSRP administrators in both cities determined which teachers would participate in the study. At the beginning of the study, seven teachers served as a treatment group and six as a comparison group, for a total of 13 participating teachers.

All sites in both counties included both AM and PM preschool. In the majority of cases, there was one lead teacher that taught both the AM and PM classes. However, in one comparison classroom and one treatment classroom, one teacher taught the AM class and another other taught the PM class. In the case of the treatment classroom, several weeks into the study the PM teacher was unable to continue teaching due to medical issues and the AM teacher took over both morning and afternoon classes. For the remainder of the analyses, this teacher who took over was considered the lead teacher for both AM and PM classes, bringing the number of treatment teachers down to six. In the case of the comparison situation, the two teachers were considered separate, as one teacher taught AM and one taught PM for the entire study. This brought the total number of teachers in the study to six treatment teachers and six comparison teachers.

Table 1

Demographics of Teachers

	Treatment Group (N=6)	Comparison Group (N=6)
Mean years of age	41	34
% White	100	83
Mean years of experience teaching preschool	10	8
% Bachelor's degree or above	100	100
MSRP Teacher	2	0

There were no significant differences between the treatment and comparison teachers in their age, race/ethnicity, years of experience teaching preschool, or whether they had earned a bachelor's degree (see Table 1). Two teachers taught MSRP classrooms, and they were both treatment teachers. The other treatment teachers and all comparison teachers were Head Start teachers.

Children. The original sample of children who participated in the study consisted of 192 children in the treatment condition and 130 in the comparison condition (N=322). However, there were several reasons why the sample included in this dissertation study is smaller than this original sample, including child attrition, absenteeism, children's refusal to participate in assessments, and statistical differences between the groups at pretest on important vocabulary measures. The following sections describe in detail how the final sample, used in all analyses in this study, was determined.

Child Attrition. During the course of the study, there was a 7% attrition rate; 24 children moved or otherwise withdrew from the study. These children were equally distributed within the treatment and comparison conditions (12 treatment children and 12 comparison children). There were no significant differences between children who moved and children who did not move on any pretest vocabulary measures (general expressive vocabulary, general receptive vocabulary, or curricular word knowledge at pretest-each of these measures will be more fully explained in the Measures section), age, gender, whether they attended AM or PM preschool, or race/ethnicity. There were significant differences between movers and non-movers on number of absences, with children who moved absent many more days than children who did not move (see Appendix A, Table A for differences between children who moved and children who didn't). For all further analyses, these 24 movers were removed from the sample.

Absenteeism. When considering the efficacy of a curriculum, it was important to determine that children in the treatment group were actually present on enough study days to adequately experience and benefit from the curriculum. The curriculum was implemented between 62 and 69 days across treatment classrooms. Missing 40% or more of these days was substantial and suggests that a child with this many absences did not experience the intervention as it was intended[3]. Thus, children who missed 28 or more curricular days (or 40% of the 69 curricular days) were removed from the sample. Of the original treatment sample, there were 17 children who missed 40% or more curricular days. Of those 17 treatment children who were chronically absent, 9 were *also* children who moved. Since the 9 movers were removed from the sample because they moved, that left 8 children that had an absentee rate of 40% or more of the days the curriculum was implemented. There were no significant differences between these 8 children and the rest of the treatment group at pretest on expressive vocabulary, receptive vocabulary,

---

[33] It is important to note that because the treatment and comparison samples are similar to one another, it is likely that there were children in the comparison group that were absent more than 40% of the time. Although children in the comparison group who were frequently absent were not missing the WOW intervention, it could be argued that they were missing a great deal of instruction and interaction that occurs typically in preschool. Ideally, to ensure parity between the groups, children in the comparison group who missed more than 40% of preschool days would also have been removed from the sample. Unfortunately, we did not collect information on absenteeism for children in the comparison group so I was unable to do this.

curricular word knowledge, age, gender, whether they attended AM or PM preschool, or race/ethnicity (see Appendix A, Table B for differences between these groups). These eight children were removed from all subsequent analyses.

Refusing to Participate in Assessments. There were 3 children in the treatment group that refused testing on a large number of the assessments across the study. These children were removed from all analyses.

Final Sample. After accounting for children who moved, children in the treatment condition who were absent more than 40% of the curricular days, and children who refused assessment, the overall sample size decreased from 322 to 287 (169 treatment children and 118 comparison children). Importantly, there were significant differences between the treatment and comparison groups on age, general receptive vocabulary at pretest, and curriculum-specific vocabulary at pretest.

Of these differences, the most concerning was the difference in overall receptive vocabulary, favoring the treatment group. Research suggests that children with larger vocabularies more easily acquire new vocabulary (Ewers & Brownson, 1999; Penno et al., 2002; Robbins & Ehri, 1994; Senechal, Thomas, & Monker, 1995). In addition, children with higher receptive vocabulary may have a learning advantage over those with lower scores in other unmeasured ways (e.g. higher general intelligence or superior ability to acquire language). Because of the nature of vocabulary learning, statistically controlling for these initial differences may not be adequate to rule out initial differences favoring the treatment group as an alternative explanation for any relationship between the WOW intervention and learning (Shadish, Cook, & Campbell, 2002).

In order to ensure parity between the treatment and comparison group, children from the remaining sample (after removing children who moved, were absent a great deal, and refused testing) were matched in pairs. Children from the treatment group were matched with children from the comparison group that were within one point on general receptive vocabulary (PPVT) at the start of the study. The matched sample used in all analyses included 89 children from the treatment group and 89 children in the comparison group, for a total of 178 children. See Table 2 for descriptives of the sample.

Table 2

Descriptives of Children in Matched Sample

	Treatment Group (N=89)	Comparison Group (N=89)
Age in Months at Pretest	51.92	50.48
Total Days Absent	7.85	9.42
Pretest Receptive Vocabulary (PPVT)	87.29	87.28
Pretest Curriculum Vocabulary	16.31**	14.11
Pretest Expressive Vocabulary (GGG)	16.28	16.28
% Female	51.7[a]	49.4
% AM Preschool	47.2	50.6
% White	62.9	49.4
% Black	24.7	30.3
% Hispanic	2.2	0.0
% Other	10.1	20.2

*p<.05; **p<.01; ***p<.001

[a] All percentages should be read as the proportion within the group that was in the treatment group or the comparison group. For example, for gender, it should be read as "Within the treatment group, 51.7% were female."

There were no significant differences between the treatment and comparison groups on age in months at pretest, total number of days absent, pretest general receptive or expressive vocabulary, race/ethnicity, proportion female, or proportion attending AM preschool. There was a significant difference between the treatment and comparison groups on curriculum vocabulary at pretest ($t = 2.90$, $p \leq .01$). To statistically control for these differences, this variable was included as a covariate in all subsequent analyses.

Treatment: The WOW Vocabulary Curriculum©

The WOW Vocabulary Curriculum (Neuman, Dwyer, Koh, & Wright, 2007) was developed by Susan Neuman's Ready to Learn research team at University of Michigan, of which I was a part. Because this was a team endeavor, for the purpose of this dissertation study it is important to situate myself within this team and to delineate what my roles and responsibilities were.

The curriculum described in this study was authored by principal investigator Susan Neuman, doctoral students Serene Koh and Tanya Wright, and myself. In addition, Tanya Wright and I created and implemented a day-long training session for preschool teachers implementing the curriculum. Also, I co-developed (with Dr. Neuman) the curriculum-based assessments of vocabulary knowledge, conceptual knowledge, and transfer of conceptual knowledge used in this study (these measures will be described more fully in later sections). During the implementation of the curriculum, I conducted approximately 30% of the classroom observations for fidelity and to provide teacher support. Finally, I independently conducted all of the analyses included in this dissertation study.

The WOW vocabulary curriculum was designed to be a supplement to the existing literacy curriculum. Lasting four months, children were taught words and concepts related to seven taxonomic categories. Lessons related to each category were taught by the classroom teacher in an eight day sequence. Each day of the sequence, the teacher implemented a 12-15 minute lesson.

The WOW curriculum was based on hypothesis that when children learn words and concepts in a well-organized taxonomic representational structure, their current word learning, conceptual knowledge, inferential ability, and learning in new situations may be enhanced. This section describes the rationale behind the taxonomic categories, concepts, and vocabulary that comprised the WOW curriculum.

Choosing Taxonomic Categories and Concepts to Teach. Many educational researchers subscribe to the hypothesis that it not vocabulary per se, but the knowledge of concepts that vocabulary words represent that fosters comprehension (Anderson & Freebody, 1981). Efforts that have as their goal increasing vocabulary, and thus later reading comprehension, must include instruction in the rich conceptual knowledge related to new vocabulary. Therefore, one of the primary goals of the WOW curriculum was to provide children with conceptual knowledge, and vocabulary related to those concepts, that would be necessary for school and literacy success.

Learning new information about a topic is highly dependent on, and influenced by what is known about related topics (Heit, 1997). Conceptual knowledge begets acquisition of new knowledge, and that knowledge serves as the foundation for later successful learning and comprehension. To ensure that the WOW curriculum included concepts and related words that would foster new learning, a research associate on the Ready to Learn team consulted preschool and kindergarten standards in five states to determine the concepts children would most likely be exposed to in their preschools and upon school entry (states included Massachusetts, Michigan, Indiana, California, and Texas). Based on the concepts included in these standards, two superordinate taxonomic categories and seven taxonomic categories related to the superordinate categories were chosen. The superordinate categories, called units, were “Living Things” and “Health”. The Living Things unit included the taxonomic categories Pets, Wild Animals, Animals in Water, and Insects. The Health unit included Parts of the Body, Clothes, and Exercise[4] (See Appendix B for a summary table of preschool and kindergarten standards and how they relate to the WOW curriculum).

Once the taxonomic categories were chosen based on concepts included in the state standards, it was important to determine the specific properties of each concept that would be taught in curriculum. The taxonomic categories served as the framework for learning, and properties were chosen that best defined the taxonomic category. Children were repeatedly taught the properties associated with each taxonomic category during each 8-day lesson sequence through teacher language, teacher questions, video clips, and pages from the informational book. Table 3 depicts the main properties that define and were taught in each taxonomic category.

[4] It should be acknowledged that some of these categories are more rich in nature, in that there are specific and necessary conditions for category membership, while others are somewhat less rich. For example, there are particular properties that all insects share (e.g. six legs, three body segments), which makes it a taxonomic category. In contrast, animals in water can vary greatly and at the most basic level only share one property: they live in water. This "category" was dropped from later versions of the curriculum. Although the topics vary in how well they fit a taxonomic definition, all of the words and concepts in all categories were taught to children around the strictures, conceptual knowledge, and defining properties associated with the category. For example, children were taught that parts of the body are 1) always attached and do not come off and 2) have a function.

Table 3

Summary of Properties and Concepts Taught in the WOW Curriculum

Unit & Category	Properties Taught
Living Things	
Pets	• Pets are animals that live with people. • People have to take care of their pets. • People take care of their pets by giving them food, water, love, play, exercise, and taking them for check-ups.
Wild Animals	• Wild animals are animals that live outside and away from people. • Wild animals cannot be pets because they don't live with people. • Wild animals take care of themselves and find their own food. • Wild animals live in different outdoor places called habitats that are just right for them. They can find the food and weather they like in their habitat.
Animals in Water	• Some animals live in the water all the time. They need to stay in the water to survive. • Animals that live in water eat, sleep, and breathe in the water. • Animals that live in water move around by swimming. • Some animals that live in water must come to the surface to breathe.
Insects	• Insects are very small creatures • Insects have three different sections called segments. • Insects have six legs. • Insects have antennae. • Some insects have wings. • Insects protect themselves from bigger animals in different ways, including stinging, bright colors, and camouflage.
Health	
Parts of the Body	• Our bodies have many parts. • Each body part has a job to do (i.e. some parts help us move, other parts are related to the five senses). • Body parts are attached to our bodies, and they do not come off. • Body parts come in different numbers.
Clothes	• We wear clothes on our bodies. • We put on clothes when we get dressed, and we take them off when we get undressed. • We wear different clothes on different parts of our bodies. • Some clothes keep us warm by covering our bodies. • Some clothes keep us safe by protecting parts of our bodies
Exercise	• Exercise helps us to keep our bodies healthy and strong. • To be healthy means that you do not get sick easily. • When we exercise, we repeat the same movement over and over again. • Exercise helps strengthen our muscles, heart, bones. • We exercise in many different ways. • Exercise can be fun.

Choosing Words to Teach. Having determined the framework of specific taxonomic categories, and properties associated with those categories, the next step was to choose specific vocabulary words to teach. The goal was to first compile a list of words that would help children most effectively learn the taxonomic categories and associated properties taught in the WOW curriculum.

Within that list, attention was paid to established methods of choosing words, as well as the difficulty of the words that would be included in the curriculum. Vocabulary researchers have suggested various ways to choose which words to teach. Beck, McKeown, and Kucan (2002) suggest that vocabulary instruction in the primary grades should focus on words that are unlikely to be learned at home (particularly in disadvantaged homes), called "Tier Two" words. They define Tier Two words as high frequency words for mature language users. Biemiller (2006) proposes a somewhat different approach, suggesting that children in the primary grades with "restricted vocabularies" should be explicitly taught "words typically known by average and advanced children by the end of grade two, but not by children with limited vocabularies." Though both of these approaches to choosing words have merit, both focus on vocabulary learning and teaching in the *primary* grades rather than preschool, which is the focus of the present study.

Adopting Biemiller's (2006) specific approach of choosing words that were known by average children at the end of grade two, or around age 9, seemed inappropriate for choosing words to teach low-income three-, four-, and five-year-olds. However, his general approach of choosing words that were typically known by average and advanced children *at a particular age* seemed a sound starting point. To find such a sample of words that would apply to this study, the MacAurthur-Bates Communicative Development Inventory (CDI) (Fenson et al., 1994) was consulted.

The CDI used parental reports from a study of 1,789 children to compile month-by-month norms for comprehension and production of 680 words from ages 16 to 30 months. The norming sample used to compile this inventory was overwhelmingly white and the majority of parents were college educated (Fenson et al., 1994). Although the authors do not state this directly, the education level of the parents in the sample indicate that the norming sample for the CDI was comprised of children from middle-class families. Because research indicates that social class is highly related to vocabulary size as early as age three (Hart & Risley, 1995), it seems likely that CDI represents age-of-acquisition norms for children with average or above average vocabularies. For this reason, the CDI represented a reasonable starting point for choosing words to teach low-income preschoolers.

One drawback to the CDI age-of-acquisition data is that it only included norms through 30 months of age. There was a mismatch between the CDI norms, which ended with children approximately 2 ½ years of age, and the sample included in the WOW intervention study, which ranged from three to five years old. However, while the norming sample used to compile the CDI was overwhelmingly comprised of middle- to upper-middle class families, the sample in the current study included mainly children coming from families living in poverty. Though the children in the current study are older than the children in the CDI norming sample, the discrepancy in social class, and thereby likely vocabulary size, makes the two samples somewhat comparable.

The goal was to apply the CDI age-of-acquisition norms to the list of words that conveyed the conceptual information to choose a proportion of words that were typically known by average middle class children at 30 months, called “easy” in this study, and a proportion of words that were not typically known by 30 months olds, which were called "more difficult". It was important that children be given opportunities to experience success, but also be challenged by learning previously completely unknown words. This resulted in identifying 183 total words to be taught in the WOW curriculum, 44% were "easy" by CDI standards and 56% of words were considered more difficult. Table 4 includes all of the words taught in the WOW curriculum; words that were considered "easy" by the CDI in each category are in bold and words that were considered more difficult are italicized.

Table 4

Vocabulary Taught in the WOW Curriculum

	Category Members	Words That Help Us Talk About the Category
Living Things		
Pets	**dog, puppy**, **fish**, **cat**, **bird**, *hamster, rabbit*	**feed, play**, **love**, **food, water**, *petting, exercise, shelter, check up, veterinarian*
Wild Animals	**deer, elephant**, **tiger**, **giraffe**, **lion**, **zebra,** *gazelle, coyote, gorilla, hippo, hyena, rhino, polar bear*	**zoo**, **outside**, **ice**, *tame, Arctic, desert, ferocious, habitat, fishing, grassland, hunting, jungle, woods*
Animals in Water	**fish**, *dolphin, goldfish, octopus, seahorse, shark, starfish, stingray, whale*	**swim/swimming**, **eat, food**, **sleep**, *fish tank, ocean, river, sea, aquarium, breathe, coral, fins, gills, snorkeling, survive*
Insects	**ant, bee, butterfly**, *katydid, ladybug, moth*	**outside**, **flowers**, *antennae, segment, creature, wings, anthill, camouflage, cooperate, hive, honey, leaves, protect, sting*
Health		
Parts of the Body	**cheek**, **chin**, **ear**, **eye, face**, **finger, foot/feet, hands**, **head**, **knee, leg, mouth**, **nose**, **shoulder, toe, tummy**, *chest, elbow, back*	**jump**, **run**, **walk**, **throw**, **hold**, **touch**, **taste**, **hear**, **see**, **clap**, *bend, nod, torso, attached, job, move, skip, wave, scratch, feel, smell*
Clothes	**jacket**, **coat**, **glove**, **pants**, **shorts, shoe, sock**, **boots**, **hat**, **sweater, shirt,** *helmet, t-shirt*	**cold**, **cover**, **dress**, **zipper**, **wet**, **dry**, **hurt, buttons/ed**, **outside**, *body, wear, warm, zip, bottom, bare, protect, sleeve*
Exercise	**jumping**, **dancing**, **riding**, **climbing**, **swimming**, **playing,** *jogging, tag, stretching, hopscotch*	**game**, **tricycle/bike**, *working out, heart, bones, muscles, moving, strengthen, active, healthy, strong, pumping, up, down, fun, trampoline*

Ideally, children in the sample (or another comparable sample) would be assessed on all 183 words prior to the study and then exclude words that were known by many of the children. However, because of the sheer number of words included in the curriculum, it was impossible to do this. Although there were limitations to using the CDI to determine which words would be "easy" (and therefore likely known by a decent proportion of children) and which would be "more difficult" (and therefore probably unknown by many children), it was the most reasonable resource available to help choose words to teach.

Careful attention was also paid to the sequence of introducing easy and more difficult words within each unit. Taxonomic categories were grouped in units based on taxonomic relatedness (e.g. the "Living Things" unit included categories of living things: Pets, Wild Animals, Animals that Live in Water, and Insects). Because the unit title denoted a superordinate category that included each of the individual taxonomic categories, there was overlap across the content and properties common to each category within a unit. For example, though pets and wild animals are quite different from one another in many ways, they are both living things, and therefore both need food, water, and shelter to survive. As part of the WOW curriculum, children would be introduced to some of the properties that apply to all categories in the unit in the first category, and this information would be reiterated and reviewed in subsequent categories. Presumably, this would lighten some of the cognitive load of learning the following categories.

For this reason, children were intentionally and gradually introduced to more difficult words as a given unit progressed. For example, in the first category of the "Living Things" unit, Pets, about 41% of words introduced were considered "more difficult". In the next category, Wild Animals, about 69% of words introduced were considered more difficult. In the final two categories in the unit, Animals in Water and Insects, respectively about 79% and 75% of words were more difficult. Then, at the beginning of the next unit, Health, children were again taught easier words, moving toward a greater number of more challenging words (67% of words were more difficult in Parts of the Body and Clothes, followed by 30% easy in Exercise).

Attention was also paid to different types of words introduced in each category. Within each of the categories, there were two "classes" of words that were taught. The first class of words was comprised of labels for category members. Though these words were generally nouns, this was not always the case (e.g. the category members in the Exercise category were generally verbs). There were 77 category member vocabulary words taught across all categories in the curriculum and a mean of 11 category members taught in each 8-day lesson sequence.

The second class of words was comprised of a more varied, and often more sophisticated, group of words that were chosen because knowledge and understanding of these words was imperative for understanding the properties and the category. In the curricular materials, these words were called "Words That Help Us Talk About *name of category*". Children were taught 106 words that enable them to talk about the category, or an average of 15 words per category. Table 4 displays all of the vocabulary taught in the WOW curriculum, both category members and the words that were imperative to understanding the categories.

The research team recognized from the beginning that the number of words to be learned was quite large, with a total of 183 words across categories and word classes.

However, when broken down into words taught per lesson sequence and per day, the number is quite reasonable. An average of 26 words were introduced in each 8-day lesson sequence (183 total words/7 categories=26 words per category), with a range between 17 and 40 words. This means that children were introduced to an average of 3 words per day (26 words per category/8 days of lessons per category=3 words per day). This seems to be an appropriate number, as Biemiller (2004) found an upper limit of three words learned per day in a review of studies using story book reading to teach vocabulary.

Defining Words. To teach the words, the WOW curriculum included several different strategies. Defining the word for children was one of the strategies. All category members were defined by their membership in a particular category (e.g. a giraffe is a kind of wild animal). Children were also taught information about category members that distinguished them from other category members (e.g. both a giraffe and a tiger are wild animals, but a giraffe finds plants and grass to eat and a tiger has to hunt for his food). Children were also taught explicit definitions for “Words That Help Us Talk About the Category”, primarily for the more difficult words. A total of 39 more difficult words were explicitly defined for children using child-friendly definitions across the curriculum, for an average of about six explicitly-defined, more difficult words per category. Children heard each definition three times on average. Examples of explicitly defined words can be found in Table 5.

Table 5

Explicitly Taught Words That Help Children Talk About the Category

Unit 1: Living Things	
Pets	*Shelter*-An animal shelter is a place where animals are taken care of until they find homes of their own. *Check Up*-Pets also need check-ups. A check-up is a way to make sure that a pet is strong and healthy. A checkup is another way to take care of a pet. *Veterinarian*-A veterinarian is an animal doctor who makes sure that pets are healthy.
Wild Animals	*Desert*-It doesn't rain very often in the desert, and it is very hot and dry. *Tame*-Coyotes and polar bears can't live with people because they are not tame, which means they don't like to be played with and petted. *Ferocious*-Sometimes they are ferocious which means they are too dangerous to live with people. *Habitat*-Animals live in places where they can find the food they need and the weather they like. These places are called habitats. Some wild animals live on the grasslands. Some live on the ice. Some animals live in the desert. *Jungle*-Another kind of habitat is the jungle. Jungles are hot and rainy. There are lots of trees and plants. (12,1)
Animals in Water	*Survive*-Survive means to stay alive. *Ocean*-Sharks live in a type of water called the ocean. Another name for the ocean is the sea. Oceans have salty water, and they are very, very big. *Gills*-Fish breathe in the water with something called gills. Gills are on their sides and let them breathe under water. Fish breathe through gills. *Aquarium*-Another way that people can see and learn about animals that live in water is to visit the aquarium. An aquarium is a zoo for animals that live in water.
Insects	*Antennae*-Insects smell and feel using special parts called antennae. *Camouflage*- Because they are very small, insects need to protect themselves from bigger animals. Katydids are the same color as the leaves, so they protect themselves by hiding in the leaves. This is called camouflage. The katydids stay safe because other animals can't find them. *Cooperate*- Like bees, ants live together. They work together to find and carry food back to the anthill. The name for working together is cooperate. Ants cooperate with each other. *Segments*- Insects have bodies with three different sections. These sections are called segments.
Unit 2: Health	
Parts of the Body	*Attached*- Our bodies have many different parts. The parts are all attached to our bodies. That means that they do not come off. *Torso*-The middle part of a person's body is called a torso. Your shoulders, your back, your chest, and your tummy are all part of your torso. Our torsos are important because they help protect the parts that are inside of us like our hearts, our lungs, and our stomachs. Our arms, legs, and head are all attached to our torsos.
Clothes	*Protect*- We wear shoes to protect our feet. If we walk around in the snow without shoes, our feet get cold. If we walk around in the rain without shoes, our feet get wet. If we step on sharp or hard or hot things, we can hurt our feet.
Exercise	*Muscle*-Muscles are important because they help us move our bodies. *Heart*-Exercise helps your heart stay strong and healthy. You can't see your heart, but it is inside your chest. Your heart is your most important muscle because it keeps your blood moving around inside of your body. *Bones*-Exercise keeps your bones healthy and strong. Bones give your body its shape. Bones help you to move. Bones also protect the important body parts that are inside of you, like your heart. *Active*-When you exercise you stay active. Active means you are moving around. *Strengthen*- When we make something stronger, we strengthen it. *Healthy*- We've learned that *exercise* is what we do to help keep our bodies healthy. When we're healthy, we don't get sick easily.

Receptive and Expressive Exposure. In addition to explicitly defining words for children, the curriculum also included many opportunities to hear vocabulary used in meaningful ways (receptive exposures), as well as opportunities to practice using the vocabulary (expressive exposures). This was important because research has shown that although children can begin to "fast map" words after one exposure (Carey, 1978), even four exposures to a word are necessary but not sufficient for increased word learning (Robbins & Ehri, 1994). Multiple exposures to vocabulary are necessary for word learning (Blachowicz & Fisher, 2000; Pressley, Disney, & Anderson, 2007; Schwartz & Terrell, 1983; Stahl & Fairbanks, 1986). Across categories, word classes, and difficulty level of words, on average children heard each word 22 times from their teacher, in the informational book, and from the video. Children were asked to produce each word an average of three times (it should be noted that the estimate of three expressive exposures is likely underestimated, as the last several lessons in each eight-day sequence provide many opportunities for open-ended conversation about the topic). It is also important to note that these averages reflect the exposures a child would have if his or her teacher followed the curriculum very closely.

Although 22 exposures to each word was the average, not all words were taught with the same intensity. There was a considerable range in the number of receptive exposures across the 183 words, from as few as two exposures to as many as 178. There was less range in the expressive exposures, from zero to 17. Generally speaking, somewhat more emphasis and more exposure was focused on category members than on words "that help us talk about a category" (average receptive exposure of category members was 24 and of words that help us talk about the category was 19). However, within the word class “Words That Help Us Talk About The Category”, emphasis was placed on words that were integral to the understanding of the concepts and properties taught in the curriculum. For example, children were exposed to the word “habitat” 41 times during the course of the curriculum because the word and concept it represents was integral to children’s understanding of the category Wild Animals. Similarly, the word “protect” was used quite frequently (43 times during the curriculum) because it was critical for understanding both that Insects protect themselves from larger animals in various ways and that Clothes protect different parts of our body.

Curricular Materials. Lesson materials included a teacher’s manual (see Appendix C, Figure 1 for a sample manual page), related Sesame Street video clips, informational big books, take-home books, and picture cards. The curricular lessons and informational books were co-written by University of Michigan RTL team members and a paid consultant with a great deal of experience in curriculum writing. Team members worked with Sesame Street staff to search through 30 years of video assets to find engaging video clips that included content relevant to our curricular topics. A talented graphic designer was hired to illustrate the informational book, the take-home book, and the picture cards. The use of several different types of materials was an effort to incorporate several media to capture and sustain children’s interest and to complement different learning styles.

Though the teacher language, video clips, and informational book did not contain exactly the same information, these three components were designed to be complimentary to one another. For example, in one lesson the teacher’s manual prompted the teacher to remind children that they have been learning about Insects, and that today they are going to learn about a new kind of insect. The teacher then played the video clip, which depicted several katydids blending in with leaves, but does include the word "camouflage". After the video, the teacher discussed with children why a katydid is an insect, and then read them a few pages from the information book that defined the word "camouflage", using an illustration of katydids camouflaging in leaves as an example.

Contents of Eight Day Lesson Sequence. Across categories, the eight-day lesson sequence contained the same basic principles. The lessons were grouped as follows: Step 1: Get Set Parts A and B, Step 2: Give Meaning Parts A and B, Step 3: Build Bridges Parts A and B, and Step 4: Step Back Parts A and B. Each step included two lessons, Part A and Part B.

Step 1 Get Set was designed to introduce children to the defining concepts and properties associated with the category, to teach children a few items that belong in the category, and to elucidate why they belong in the category. To teach them this, in both Part A and Part B of this step, children viewed a Sesame Street video clip, were read a few pages of the informational book, and engaged in teacher-child language interactions with more emphasis on the teacher providing information and the children repeating this information. Also, in Part B, children began to view picture cards that depict category members.

Step 2 Give Meaning was designed to reiterate the conceptual knowledge and properties that define the category. In addition, children learned more category members, with a focus on why they are category members. Children are asked to consider how different category members are alike, and also how they might differ. As with the previous step, video clips, pages from the informational book, and picture cards were used to teach these words and concepts. Teacher-child language interaction began to move toward individual responses from children.

Step 3 Build Bridges was designed to deepen children's understanding of the topic by reviewing concepts, properties, and members of the category. In addition, in this step children were asked to consider items that are not in the category, with explicit discussion around *why* these items are not in the category (e.g. a cow is not a wild animal because it lives with people on a farm). The goal was to help children to see that there are certain properties and characteristics that are necessary for category membership, and that items that do not possess these characteristics cannot be category members. The same materials were used to convey these concepts, with a focus on comparing picture cards of items that are and are not in the category and engaging children in discussion about *why* certain items are in the category and other items are not (see Appendix C, Figure 2 for an example of picture cards used to illustrate this concept).

Step 4 Step Back was designed to review the concepts, properties, and vocabulary associated with the category. A critical element of this step was presenting children with "challenge" items. These are items that are not clearly in or out of the category (e.g. for insects, challenge words included *centipede* and *spider*). Through discussion, children were challenged to think critically and apply what they learned about the properties common to all category members in order to determine if these items are "in" or "out" of the category (see Appendix C, Figure 3 for an example of suggested teacher language around challenge items). The same materials are used in these lessons, but at this final stage the teacher reads the informational book in its entirety, rather than just a few pages.

In addition, on the last day, children are prompted to write in response to the topic and are given a copy of the informational book to take home and share with their family.

Sections Included in Individual Lessons. Each 12-15 minute lesson was comprised of between 3 and 8 sections. On all days, lesson components included "Tuning In", "Let's Get Started", "Let's Read", "Let's Look at Pictures", and "Let's Wrap Up" (in that order). On some days, lessons also included the following sections: "Time for a Challenge" and "Let's Write About It".

The Tuning In section focused on developing phonological awareness and included a video clip and teacher-child language interactions. This section was designed for two purposes. First, it was meant to engage children during the initial minutes of the lesson so that they would be ready to learn the vocabulary and content included in the main body of the lesson. In addition, it was designed to develop children's early phonological awareness in accord with research suggesting that phonological awareness and vocabulary are inextricably linked (though the directionality of the relationship is still debated) (Metsala & Walley, 1998).

During the Let's Get Started portion of the lesson, the teacher reviewed information learned in the previous days and introduces the video clip. Children watch the video clip, after which the teacher engages children in a conversation about the content of the clip by providing information and asking questions. This conversation focuses on category members, as well as concepts and properties that define the category.

After Let's Get Started, the teacher reads children two to three pages from the informational book. The book pages generally introduced more members of the category and provided additional information about the category. After reading these pages, the teacher talked to children about the content of the book.

In the next section, Let's Look at Pictures, the teacher engaged children in activities with picture cards. Picture cards depicted category members, items that were not in the category, and challenge items (items that were not clearly in or out of the category). Early in the eight day lesson sequence, the picture card activity was used as a labeling activity. As the sequence progressed, picture cards were used as a way to spark conversation around the concepts and properties of the category that children were learning by comparing and contrasting several category members and by sorting cards into items that are and are not members of the category. On some days, this included the "Time for a Challenge" section, where children are asked to use what they have learned to consider and discuss items that are not obviously in or out of the category. Comparing, contrasting, and sorting these items engaged children in applying the conceptual knowledge they were acquiring to think critically about the items and the category.

The Let's Wrap Up Section was originally designed to review the vocabulary and concepts that were learned in a particular lesson through teacher-child dialogue or through additional work with picture cards. However, during the initial weeks of the curriculum, it was observed that children were having difficulty focus on this final piece of the lesson. For that reason, in subsequent lessons, the Let's Wrap Up section was comprised of encouragement and a brief conversation about what children had learned. For example, teachers are prompted to say "Today we began to learn about exercise. Let's talk about what you've learned so far. Exercise keeps us healthy. Say it with me. What is one exercise that can help build strong muscles and bones?"

On the final day of the lesson sequence, lessons included an additional "Let's Write About It" component. In this section, children are given a piece of paper and asked to draw a picture related to the category (e.g. draw a picture of yourself doing your favorite kind of exercise) and to try to write words to go with the picture. Children then share their pictures with one another and pictures are bound together in class book about the category.

Gradual Release of Responsibility. The WOW curriculum was designed to incorporate a gradual release of responsibility for conversation from the teacher to the children. During the initial lessons in the eight-day sequence, the teacher-child language interactions were more teacher-directed. The teacher's role was to heavily scaffold children's use of the new vocabulary and concepts by asking yes or no questions (e.g. "Are a moth and a katydid both kinds of insects?"), asking children to repeat vocabulary, properties, and conceptual knowledge, or asking questions with simple answers. For example, a teacher might have asked "How many segments do insects' bodies have?" to which children would answer "Three". As children learn and gain practice using the vocabulary and concepts associated with the particular category, the teacher provides less language scaffolding by beginning to ask more questions to individual children, more open-ended questions, and to support more child-directed conversation. For example, in Lesson 4A, the teacher is prompted to ask children "How does exercise keep our bodies healthy and strong?".

Literacy Curriculum in Treatment and Comparison Classrooms

The WOW curriculum was designed to be a supplement to the literacy curriculum already in place in preschool classrooms. Three of the treatment teachers implemented

“Creative Curriculum” (Dodge, Colker, & Heroman, 2002). This program includes a literacy component, although it generally takes a broad developmental approach with a focus on interest areas. The remaining three treatment teachers, as well as all six comparison teachers, implemented a modification of the High/Scope Preschool Curriculum called “Building Bridges”. This curriculum was modified from the general High/Scope curriculum to include a more teacher-directed approach and more of a focus on letters, conventions of print, and environmental print.

Child Measures

To investigate the theoretical premise that teaching words and concepts in taxonomic categories was beneficial, children were assessed in each of the following four areas: a) word knowledge and usage, or knowledge and use of words taught in the curriculum, b) conceptual knowledge and usage, or knowledge and use of concepts taught in the curriculum, c) transfer of conceptual knowledge, or ability to transfer conceptual knowledge taught in the curriculum to new learning tasks[5], and d) general word knowledge, including receptive and expressive knowledge of general vocabulary. In addition, demographic information was collected to account for child-level differences. This section describes the tasks that children were asked to perform and the measures that were created from children’s performance on each task (including how the measures were constructed and their psychometric properties).

Tasks were administered at three time points: at the beginning of the study, immediately following the instructional sequence of each taxonomic category, and at the conclusion of the study. For ease of understanding, tasks and measures will be identified

[5] Susan Neuman and I created the curriculum-based assessments designed to assess children’s vocabulary learning, conceptual knowledge, and transfer of conceptual knowledge to new learning tasks. Two widely-used, norm-referenced, standardized measures of general vocabulary were also administered.

as being administered at Time 1, Time 2, or Time 3. Time 1 measures were administered at the beginning of the study, before instruction began. Time 2 measures were administered immediately following the eight-day instructional sequence for each topic. Time 3 measures were administered at the conclusion of the study. Figure 4 depicts the different tasks that were administered, the measures that were constructed based on children's performance on each task, and the time point at which the measure was collected.

Figure 4

Timing of Administration of Word and Conceptual Knowledge Tasks and Associated Measures

Time 1: Before the Study Began	**Time 2: Immediately Following the Instructional Sequence for Each Category**	**Time 3: After the Study Ended**
Task: "What Is It?" *Measure*: Word Knowledge	Task: "Picky Peter" *Measure:* Word Knowledge *Measure:* Conceptual Categorization-Items Taught *Measure:* Conceptual Categorization-Items Not Taught *Measure:* Justification of Conceptual Categorization-Items Not Taught	Task: "What Is It?" *Measure:* Word Knowledge
Task: "Tell Me" *Measure*: More Difficult Word Use *Measure*: Conceptual Knowledge Use		Task: "Tell Me" *Measure:* More Difficult Word Use *Measure:* Conceptual Knowledge Use
Tasks: General Word Knowledge *Measure:* General Receptive Word Knowledge (PPVT) *Measure:* General Expressive Word Knowledge (GGG)		Tasks: General Word Knowledge *Measure:* General Receptive Word Knowledge (PPVT) *Measure:* General Expressive Word Knowledge (GGG)

All child measures were administered by graduate students in education and experienced researchers with college degrees, who were trained in several training sessions prior to their work in the field. In addition, doctoral students who were members of the research team monitored the testing materials to rule out any administration problems.

Demographics

Demographic information collected on children in the study included treatment condition, gender, race/ethnicity, age in months at pretest, whether the child attended AM or PM preschool, and total number of days absent during the study.

Curriculum-Based Vocabulary and Conceptual Knowledge Tasks and Measures

To assess children's knowledge of words and concepts that were taught in the curriculum and ability to transfer that knowledge to new learning tasks, children participated in three distinct tasks: the "What Is It?" Task, the "Tell Me" Task, and the "Picky Peter" Task. From children's performance on each of these tasks, several measures of word and conceptual knowledge were created. This section describes each task in its entirety and details the curriculum-based measures that were created from each task.

The "What Is It?" Task

To measure growth on word knowledge from the beginning to the end of the study, an assessment of expressive word knowledge was developed specifically to assess words taught within the WOW curriculum. Five vocabulary words were randomly selected[6] from each of the seven taxonomic categories taught during the curriculum

[6] Five words were randomly chosen from each topic due to a concern that children would suffer from test fatigue if assessed on all words from all topics.

intervention, for a total of 35 words[7]. Words were depicted on individual cards. For each card, children were asked "What is it?" and each answers were recorded. This task was administered to children in both the treatment and comparison groups at Time 1 (before the study began) and at Time 3 (after the study ended) (see Figure 4 for a graphic that displays the timing of each task). There was one measure created from children's performance on this task, Word Knowledge (Time 1 and Time 3), and it is described next.

Word Knowledge (Time 1 and Time 3). From the "What Is It?" task, a measure of work knowledge was created. This measure captured children's expressive knowledge of curriculum-specific words before the study began and after the study ended. Children's responses on this task were coded at either correct (1) or incorrect (0). A total score was calculated by summing each child's total points (this process was the same for children's performance on this task at Time 1 and Time 3). Possible scores on this measure ranged from 0 to 35. The reliability of this assessment was good (Cronbach's $\alpha = .82$) and both the pretest and the posttest were standardized such that they would have a mean of 0 and a standard deviation of 1 (although this and all subsequent measures were standardized for use in multivariate analyses, the results section presents basic descriptives of each measure *before* standardization). The pretest was used as a continuous predictor in analyses and the posttest was used as an outcome measuring acquisition of curricular word knowledge over the course of the study.

[7] Twenty-three of these words were labels for category members and twelve were "Words That Help Us Talk About" the categories. Nineteen of these words were considered "not known" by middle-class 30 month olds according to the McArthur-Bates Communicative Development Inventory (CDI).

The "Tell Me" Task

This task was designed as an open-ended task to get children talking about the taxonomic categories that were taught in the curriculum. To do this, children were shown one card for each of six categories: Pets, Wild Animals, Animals in Water, Insects, Clothes, and Parts of the Body[8]. On each card, several category members were depicted. For each card, children were told "Tell me everything you know about *name of the category*" (e.g. "Tell me everything you know about Insects") and were given one minute to respond. If the child said nothing, the assessor prompted the child again by saying "What do you know about *name of the category?*" Children's responses were audio-recorded and transcribed. This task was administered to children in both the treatment and comparison groups at Time 1 (before the study began) and Time 3 (after the study ended). There were two measures created from children's performance on the Tell Me Task: More Difficult Word Use (Time 1 and Time 3) and Conceptual Knowledge Use (Time 1 and Time 3). They are described in the following paragraphs.

More Difficult Word Use (Time 1 and Time 3). From the "Tell Me" task, a measure of more difficult word use was created. Many measures of word knowledge are based on showing a child a picture and asking the child to label the picture. However, many of the more difficult vocabulary words taught in the WOW curriculum were difficult to assess using such a simple procedure. For example, it is difficult to adequately depict "habitat" or "survive" in a picture for children to name. One way to assess children's knowledge of these more difficult words was to capture children's spontaneous use of these words during an open-ended task like the "Tell Me" task.

[8] Exercise words were not included because only a few teachers had completed these lessons at posttest.

To determine which words taught in the WOW curriculum were considered more difficult, words had to meet two criteria. First, the word was listed in the front of the teacher's manuals as words that were taught in either of the following two word classes: "Types of *name of category*" and "Words That Help Us Talk About *name of category*". In the former class of words, children were taught the category membership of the word (e.g. "Gorilla is a kind of wild animal"). In the latter case, children were taught the meaning of these words (e.g. "Wild animals live in a *habitat,* which is a place that has the food and weather that they like"), were often asked to say these words (e.g. "When we make something stronger, we *strengthen* it. Say it with me"), or were otherwise exposed to these words (e.g. "Because they are very small, insects need to *protect* themselves from bigger animals" or "What part of our bodies are attached to our *torso*?").

The second criteria for words that were more difficult was that the word was determined by the MacArthur-Bates Communicative Development Inventory (CDI) as not typically known by 30-month-olds. There were 92 words that met these criteria (see Appendix D for a list of these 92 more difficult words). Twenty-eight of these words were category members, or "Types of *name of category",* and the remaining 64 words were "Words That Help Us Talk About *name of the category*.

Using the transcripts of children's responses to the "Tell Me" task, children's responses were searched and coded for spontaneous use of these 92 more difficult words. It is important to note that children's responses were searched for presence of these more difficult words. The coding process *did not* include rules about the child using the word correctly (either semantically or syntactically). Children received one point for each more difficult word that they spontaneously used during this assessment and 0 points for each more difficult word they did not use, and then a total score was derived by summing all points across topics.

Coding was done by a trained, college-educated adult and me. Approximately 20% of responses were independently coded by both coders and compared to determine interrater reliability, which was excellent (97% agreement). Both the Time 1 and Time 3 version of this measure were standardized to have a mean of 0 and a standard deviation of 1 and were used as a predictor and an outcome, respectively.

Conceptual Knowledge Use (Time 1 and Time 3). Also derived from the "Tell Me" task was a measure of children's use of conceptual knowledge. To create this measure, children's responses during the "Tell Me" task were coded for absence or presence of curriculum-specific conceptual knowledge, with a child earning one point for each time they spontaneously used a property or concept that was taught in the curriculum when talking about a topic. Children were only given points for using concepts that were taught in the WOW curriculum[9] (Appendix E depicts the coding scheme used for scoring children's responses). Concepts were any statement that referred to a property that applied to the entire taxonomic category. For example, a children would get one point each for the following statements: "Wild animals live outside and away from people", "Cause they all have six legs" (referring to insects), or "Cause they are all attached to me" (referring to parts of the body). A total score was created by summing all points across all topics.

Two research assistants from the Ready to Learn project coded children's transcripts. Interrater reliability for coding children's responses was good (92%

[9] Children were not given points for repeating the same concept several times (e.g. a child who said "Mittens go on your hands and pants go on your legs and coats go on your arms" would be given credit for using one property "We wear different clothes on different parts of our body").

agreement). Both the Time 1 and Time 3 versions of this measure were standardized to have a mean of 0 and standard deviation of 1. The Time 1 version was used as a predictor and the Time 3 version was used as an outcome.

The "Picky Peter" Task

To assess if children learned the words that were taught and could apply the concepts and properties they were taught, the "Picky Peter" categorization task was administered immediately following instruction for each taxonomic category. This task determined if children learned the words that were taught and could apply the concepts and properties they were taught to accurately categorize items that were taught and were not taught in the curriculum.

Modified from Waxman and Gelman's (1986) Picky Puppet task, after the instructional sequence for each category, assessors showed children a stuffed lion named Picky Peter and explained that Peter was a "picky" animal because he only likes things that are in a certain category (for example, he only likes things that are insects). The assessor asked the child to help find the things that Picky Peter likes. There were several steps to this task.

During the first part of the assessment, children were shown 10 pictures of items that were explicitly taught during the previous eight-day lesson sequence as either in or out of the category (six items were in-category and four were not). Children were asked two things about each of these 10 items: 1) to expressively name the item by answering the question "What is it?", and 2) to categorize the item by answering the question "Is it a *name of category*? Yes or no". Answers were recorded.

During the second part of the assessment, children were shown 10 pictures of items that *were not* explicitly taught during the eight-day lesson sequence. Importantly, this group of "untaught" items included six items that were members of the category that had just been taught (but had not been taught) and four items that were not members of the category. Children were asked three things about each of these 10 items: 1) to expressively name the item by answering the question "What is it?"[10], 2) to categorize the item by answering the question "Is it a *name of category*? Yes or no", and 3) to justify their categorization choice by answering the question "How did you know _____ is/is not *name of category*?" (prompt was dependent on the child's categorization choice). Answers were recorded and verbal justifications were audio-recorded and transcribed.

Six versions of this task were developed, each specific to one taxonomic category[11]. Immediately after the instructional sequence for each taxonomic category, the appropriate version of the task was administered to children in the treatment group. It is important to note that though all children in the treatment group were administered this task immediately following instruction of each category, the timing of administering this task occurred differently in the comparison group. Because children in the comparison group were not experiencing instruction around the categories, it was not imperative that this task be administered to them at a particular point in time. What was important here

[10] Children's word knowledge of items that weren't taught in the curriculum (their response to the question "What is it?") were recorded and scored as correct or incorrect. If the child was unable to name the referent, the assessor provided the label for the child. Children's responses were summed within topics. Totals were averaged across topics. Though this measure was not used as an outcome or a notable predictor, it was included in one multivariate model as a control for prior knowledge of items that were not taught in the curriculum. The reliability was good (Cronbach's $\alpha = .85$).

[11] The six versions of the Picky Peter task assessed children's knowledge of Wild Animals, Animals in Water, Insects, Parts of the Body, Clothes, and Exercise. Because Pets was the first topic taught in the curriculum, the Ready to Learn research team felt that it would be wise to give both teachers and children time and space to settle into the new curriculum before coming in to assess the children. Therefore, children were not administered a Picky Peter task immediately following Pets. For this reason, there were only six Picky Peter tasks administered even though seven taxonomic categories were taught.

was to demonstrate that a similar group of children, during the same four month period, were less knowledgeable about the vocabulary and concepts taught in the curriculum than children in the treatment group. For this reason, Picky Peter tasks were administered to the comparison group during the last month of the intervention period.

In addition, one goal of the research team was to minimize the amount of time that children in the comparison group would be removed from the classroom. For this reason, a random sample of children in the original comparison group (n=50) were assessed on four end of topic vocabulary assessments (Wild Animals, Animals in Water, Insects, and Parts of the Body).

In the implementation of this study, the research team decided that this assessment was to be administered only after instruction (rather than both before and after) because it was meant to serve as a progress monitoring tool that might eventually be streamlined for teachers' use in the classroom. However, after the study was completed, it became clear that this was an important task for the research team to consider more carefully because it effectively captured learning of curricular vocabulary and conceptual knowledge *immediately* after instruction for each category.

Four measures were created from children's performance on the Picky Peter task: Word Knowledge (Time 2), Conceptual Categorization of Items Taught (Time 2), Conceptual Categorization of Items Not Taught (Time 2), and Justification of Conceptual Categorization (Time 2). These measures are described in the following pages.

Word Knowledge (Time 2). To measure word knowledge immediately following instruction, a measure was created from children's responses during each of six topic-specific "Picky Peter" tasks (Wild Animals, Animals in Water, Insects, Parts of the Body, Clothes, and Exercise). As described above, part of the “Picky Peter” task included showing children 10 items that were taught in the curriculum and asking them to label the item by asking “What is it?”. Children’s responses to each item were coded as correct (1) or incorrect (0).

Although children were shown 10 items to *categorize* in each “Picky Peter” task, this specific measure only included children’s *word knowledge* of those items that were category members (in other words, there were non-category members included in the 10 items for sorting purposes, but they were excluded in this measure). In the Wild Animals category, there were four target words assessed. In the remaining five categories, there were six target words per category that were assessed, for a total of 34 words assessed across “Picky Peter” tasks.

Children’s score on each item within a Picky Peter task were summed and divided by the total number of items in that task to get a percentage correct for each Picky Peter task. These percentages were then averaged across tasks[12] to get an average percent correct across topics. It was possible to score from 0 to 100% correct on this measure. This measure had adequate reliability (Cronbach’s $\alpha = .73$) and was standardized to be used as a continuous outcome.

Conceptual Categorization of Items Taught (Time 2). Another measure created from the “Picky Peter” task was children’s ability to use conceptual knowledge to categorize items that were taught in the curriculum. This measure determined if children

[12] The rationale for using an average score across topics, rather than a total sum, was to maximize the number of children with a valid score on this outcome. Across topics, the Time 2 assessments were administered on seven different days. If a sum total was used, children that were absent on one of these days, but present for the remaining six days, would have a lower score than children who were present for all assessment days. Using a mean across topics and allowing for absences, scores of all children are comparable, regardless of the number of Time 2 assessments they actually completed.

learned and could apply the concepts and properties they were taught to accurately categorize items that were taught in the curriculum.

As a reminder, as part of the "Picky Peter" task, children were shown 10 pictures of items that were explicitly taught during the intervention as either in or out of the category and were asked "Is it a *name of category*? Yes or no" (e.g. "Is it a wild animal? Yes or no"). Because this was a sorting task, some items were members of the category and some were not. Six of ten items were considered "in-category" because they were members of the taxonomic category. The other four items were "out-of-category" because they *did not* possess the requisite properties to belong to the category. Both in- and out-of-category items were taught in the curriculum, with a focus on the in-category words.

Children's categorization choices were coded as correct (1) or incorrect (0). Children's score on each item within a Picky Peter task were summed and divided by the total number of items in that task to get a percentage correct for each Picky Peter task. These percentages were then averaged across tasks to get an average percent correct across topics[13]. Possible scores could range from 0 to 100%. The reliability of this assessment was good (Cronbach's $\alpha = .83$). This measure was standardized to be used as a continuous outcome in analyses.

Conceptual Categorization of Items Not Taught (Time 2). Another measure derived from the Picky Peter task was children's ability to demonstrate near-transfer by

[13] As with the Time 2 word knowledge assessment, an average score across topics was calculated for each child rather than a total sum of scores across topics. The rationale for doing this was exactly the same; doing so maximized the number of children with a valid score on this outcome. Across topics, the Time 2 conceptual categorization assessments were administered on seven different days. If a sum total was used, children that were absent on one of these days, but present for the remaining six days, would have a lower score than children who were present for all assessment days.

using conceptual knowledge to categorize items that were not taught in the curriculum. This was considered a measure of near-transfer rather than far-transfer because children were being assessed on their ability to categorize items that were not taught but *were* (or were clearly not) part of the taxonomic category that was taught in the curriculum. The rationale behind this measure was to provide an initial demonstration that children who experienced the WOW curriculum could apply what they learned *outside* of the confines of the curriculum; a first suggestive indication that vocabulary and conceptual knowledge may be used to foster later learning.

As part of the Picky Peter task, children were shown 10 items that *were not* taught in the curriculum. The child was then asked, “Is it *name of category*? Yes or no” (six of these items were “in-category” and four were “out-of-category”). Children’s categorization choices were then recorded as correct (1) or incorrect (0). Responses were summed within each topic and that total was then divided by the total number of items to get a percentage correct. Then an average percent correct across topics was calculated. Possible scores could range from 0 to 100%. The reliability of this assessment was good (Cronbach’s $\alpha = .82$). This measure was standardized to be used as a continuous outcome in analyses.

Justification of Conceptual Categorization of Items Not Taught (Time 2). The final measure that was created from children’s performance on the Picky Peter tasks is children’s verbal justification of how they categorized items that were not taught in the curriculum. This measure provides additional evidence of near-transfer of conceptual knowledge acquired from the WOW curriculum to new learning situations, suggesting that this type of knowledge may augment later learning.

During the Picky Peter task, after children categorized items that were not taught in the curriculum, they were asked to justify their categorization choice by answering the following question: "How did you know _____ is/is not *name of category*?". Children's justifications for each item were audio-recorded and transcribed.

Children's responses were then coded using an adaptation of Langer's (1984) rubric for characterizing the quality of children's prior knowledge. Langer's rubric included three characterizations, as follows: a) Highly Organized, characterized by inclusion of superordinate concepts, definitions, analogies, and links to other concepts, b) Partially Organized, characterized by knowing some examples of the class or category, some attributes, and some defining characteristics, and c) Diffusely Organized, characterized by tangential links, first hand experiences, and possible confusion about words that have similar phonological representations (Langer, 1984). Langer's (1984) rubric assessed the organization of children's prior knowledge by determining if children's responses had characteristics of taxonomic organization. For this reason, there are elements of this rubric that directly apply to this study.

For the purposes of this study, Langer's (1984) rubric was modified to be used to characterize children's verbal justifications of their conceptual categorization choices. Essentially, there was one major difference between Langer's (1984) characterizations and the rubric developed for this study. Children's responses that could be described by Langer's rubric as either "Highly Organized" or "Partially Organized" were collapsed into one category called "Organized". These two categories were collapsed because the characteristics included in both Langer's Highly Organized and Partially Organized categories described evidence of the type of taxonomic organization that was a goal of this study.

After modifying Langer's rubric, the final rubric used to characterize children's conceptual knowledge in this study was as follows: a) Organized (1 point), including at least one of the following: information about properties, attributes, or defining characteristics necessary for category membership (e.g. "A dragonfly is an insect 'cause it has six legs") and b) Diffusely Organized (0 points), characterized by a dearth of information about properties necessary for category membership and often focusing on tangential links or firsthand experiences, (e.g. "I don't know", "Cause a tiger is striped", "My mom told me", or "I saw it on Diego"). Appendix F includes the coding scheme for characterizing children's responses generally and a list of acceptable concepts for one example taxonomic category, Wild Animals[14].

In each Picky Peter task, children were asked to verbally justify 10 categorization choices. Each justification was coded and assigned either 0 (Diffusely Organized) or 1 (Organized) per the coding scheme described above. Children's total points within each Picky Peter task were then summed, for a possible total of 10 points. Total scores from each Picky Peter task were then summed and divided by the total number of Picky Peter tasks to derive a mean score across tasks. It was possible to score between 0 and 10 on this measure. This measure was standardized to be used in analyses.

[14] It is important to note that when coding these responses, children's sorting choice was taken into account. This was important, as this measure was designed to determine if children are able to justify their sorting choices using conceptual information. Take, for example, a child who incorrectly sorted a dragonfly as "not an insect". Then, when the child was asked "How did you know it wasn't an insect?", the child said "Because it has six legs!". Although this child stated a property common to insects writ large, she did not adequately justify *her* sorting choice and would receive 0 points for this answer. By taking the sorting choice into account, this process of coding can accurately be described as coding children's sorting justifications.

To establish inter-rater reliability, 20% of responses were independently coded by two coders (myself and a trained colleague with a master's degree from University of Michigan). Analyses indicated that there was a reasonable degree of inter-rater reliability, with 90% agreement overall.

General Vocabulary Knowledge Tasks

In addition to measuring curriculum specific word and conceptual knowledge, as well as transfer to new learning tasks, it was important to determine if children's exposure to the WOW curriculum had an influence on their general word knowledge. General word knowledge measures were administered at Time 1 (before the study began) and at Time 3 (at the conclusion of the study).

General Receptive Word Knowledge. To measure general receptive word knowledge, all children in both the treatment and comparison groups were individually administered the Peabody Picture Vocabulary Test (PPVT III-R) (Dunn & Dunn, 1997) of receptive vocabulary at the beginning and end of the study. Children were shown a series of plates with four pictures on each plate: the target and three distractors. The children were told to point to the picture that best depicts the target word. As a widely-used, respected, valid, and reliable assessment, this measure serves as a standardized and norm-referenced measure of children's general language ability, taking age into account. Across test forms and ages, median reliability of this measure was .94 (Dunn & Dunn, 1997). The Time 1 version of this measure was used to match children in the treatment and comparison groups and the Time 3 version was standardized and used as an outcome.

General Expressive Word Knowledge. To measure general expressive word knowledge, all children in the treatment and comparison group were individually administered the Get It, Got It, Go! (GGG) (Early Childhood Research Institute on Measuring Growth and Development, 1998), an expressive naming assessment at the beginning and end of the study. The GGG is a standardized, timed assessment of general expressive vocabulary. For one minute, children were shown a series of randomly chosen pictures and asked “What’s that?”. As soon as the child said the name of the item, the assessor moved to the next item. If a child does not respond, the assessor repeated the prompt after three seconds. If the child still didn’t respond, the assessor moved on after an additional two seconds. A child’s score reflects the number of correctly named pictures in one minute. The reliability of this assessment was adequate (Cronbach’s α =.85). This measure was standardized and used as a continuous predictor in analyses and an outcome measuring general vocabulary knowledge.

Summary of Child Tasks and Measures. In sum, the “What Is It?”, “Tell Me”, and General Vocabulary tasks were administered to children at Time 1 (the beginning of the study) and Time 3 (the end of the study). The “Picky Peter” task was administered at Time 2 (immediately following instruction for each category).

From children’s performance on these tasks, nine measures were created to assess curriculum-specific word knowledge, curriculum-specific conceptual knowledge, transfer of conceptual knowledge to new learning tasks, and general word knowledge. The following is a list of the measures that were designed to capture children’s performance in these areas:

1) Curriculum-specific word knowledge

   These measures were designed to assess children’s knowledge and use of words taught in the curriculum.

- Word Knowledge (Time 1 and 3)
- Word Knowledge (Time 2)
- More Difficult Word Use (Time 1 and 3)

2) Curriculum-specific conceptual knowledge

These measures were created to assess children's knowledge and use of conceptual knowledge taught in the curriculum.

- Conceptual Knowledge Use (Time 1 and 3)
- Conceptual Categorization of Items Taught (Time 2)

3) Transfer of conceptual knowledge to new learning tasks

These measures were designed to assess children's ability to apply the concepts they learned in the curriculum to items that were not taught in the curriculum, make inferences about their category membership, and verbally articulate their reasoning.

- Conceptual Categorization of Items Not Taught (Time 2)
- Justification of Conceptual Categorization (Time 2)

4) General vocabulary knowledge

These tasks were designed to assess children's general vocabulary knowledge.

- General Receptive Word Knowledge (Time 1 and 3)
- General Expressive Word Knowledge (Time 1 and 3)

Classroom Measures

Classroom Language and Literacy Environment. The Early Language and Literacy Classroom Observation Tool (ELLCO) (Smith et al., 2002) was administered at the beginning of the study to assess the overall language and literacy environment in each classroom. The ELLCO served as a baseline measure of the classroom language and literacy environment across classrooms and was used to determine if there were significant differences in the literacy environment between the treatment and comparison classrooms.

The instrument includes four sections: a) a literacy environment checklist, including book area, book selection, book use, writing materials, and writing around the room, b) a classroom observation of the general classroom environment and language, literacy, and curriculum environment[15] c) a teacher interview, and d) a literacy activities rating scale, or the number of various book reading and writing activities observed during the observation. For the purposes of this study, the teacher interview was not included.

A total score was derived for each teacher by summing scores from the literacy environment checklist, the observation of the general classroom environment, the observation of language, literacy, and curriculum environment, and the literacy activities rating scale. Possible scores range from 15 to 124.

In this study, each teacher taught both AM and PM preschool classes. The ELLCO was administered in both AM and PM classrooms for each teacher before the study began. Because child analyses will be primarily comparing performance of both AM and PM children of a given teacher to the performance of both AM and PM children of other teachers, a mean score for each teacher was derived by averaging AM and PM

---

[15] Classroom observation sections were scored on a scale between 1 (deficient) and 5 (exemplary). The general classroom environment section was comprised of six scales including organization of the classroom, contents of the classroom, presences and use of technology, opportunities for child choice and initiative, classroom management strategies, and classroom climate. The language, literacy, and curriculum environment section was comprised of eight scales including oral language facilitation, presence of books, approaches to book reading, approaches to children’s writing, approaches to curriculum integration, recognizing diversity in the classroom, facilitating home support for literacy, and approaches to assessment.

ELLCO scores for each teacher. Although ELLCO subscales and total scores were not exactly the same at AM and PM for a given teacher, they were quite similar. The largest score differential between AM and PM for any one teacher was 5 points. Other differentials ranged from .2 points to 2.5 points.

The goal of this dissertation study was to investigate children's learning. Measures of classroom quality are secondary to this goal, but are important for ensuring parity between the treatment and comparison classrooms on the classroom environment outside of the WOW curriculum. For this reason, results comparing ELLCO scores in the treatment and comparison classrooms are presented here, in the Method section, rather than in the Results chapter.

There were no significant differences between the treatment and comparison groups on overall mean ELLCO score ($t = 1.91$, ns). The treatment classrooms, on average, scored 83 out of 124 possible points, while the comparison group scored on average 70 out of 124 possible points.

Activities in Circle Time. Circle time is a very frequent event in most preschool classrooms. It is generally the time when preschool teachers engage children in instructional exchanges ranging from book reading to "calendar" to numeracy activities. Capturing and analyzing what activities occur during circle times can provide a general sense of what types of instructional interactions children are experiencing.

To learn more about what types of instruction and activities were happening in the preschool classrooms included in this study outside of the WOW curriculum, the Ready to Learn team chose to videotape both treatment and comparison teachers conducting their "circle times" with both their AM and PM classes. For the purposes of this dissertation study, data from videos of circle time instruction were used to illustrate that there was parity between the treatment and comparison classrooms on instruction outside of the WOW curriculum.

Teachers in both treatment and comparison groups were videotaped on three separate observation days over the course of the school year. During the first observation, the WOW curriculum was not present in any of the circle time lessons. During the second and third observations, circle times in treatment classrooms often included the WOW curriculum. Because the purpose of using circle time observations in this dissertation study was to investigate the instruction and activities that were happening in both treatment and comparison classrooms *outside* of the WOW curricular lessons, only data from the first observation are included here.

Teachers in both treatment and comparison groups conducted an average of 6 circle times during each observation and about 3 circle times per class session (AM or PM). Each circle time that a teacher conducted was videotaped and then coded by two independent raters using a coding scheme developed by Susan Neuman and other Ready to Learn team members. The aspects of the coding scheme investigated in this dissertation study focus on the amount of time teachers spent engaging children in activities that have instructional intent (ideally), including time spent on the following: overall circle time, book reading, calendar, weather, attendance, daily news, singing, group writing, alphabet, vocabulary, rhyming, beginning sounds, other phonological awareness activities, text conventions, other language arts skills, colors, numbers/counting, shapes, patterns, or science (see Appendix G for a more detailed description of each of these codes).

Videos were coded by two research assistants (Sarah Tucker and Caitlin Dougherty) using the Noldus Observer Pro system (Noldus Information Technology, 2001). Any activity that lasted 15 seconds or more was coded. Intercoder reliability was computed by the Noldus software by comparing time and activity of a 5-minute segment on a classroom videotape coded by two different researchers for 30% of the tapes selected at random. For each comparison, weighted kappas exceeded 92% (calculations of intercoder reliability were calculated by Susan Neuman and the Ready to Learn research team).

For the same rationale given above regarding presentation of ELLCO scores in the method section, results comparing the treatment and comparison classrooms on instructional activities during circle time are presented here in a narrative and in Appendix H in numerical form. There were no differences between treatment and comparison group circle times on the amount of time spent on any of the instructional activities, with the exception of focusing on colors ($t = 9.36$, $p \leq .05$). It should be noted, however, that simple counts of instructional time in minutes does not address the quality of that instruction. However, the parity between the treatment and comparison groups on the amount of time spent on each instructional activity is one indication that children in each of these groups were exposed to similar types and durations of instructional activities outside of the WOW curricular lessons.

Treatment of Missing Data

Independent Measures. As with most datasets collected in schools, there was some missing data on several measures. Two percent of children were missing general receptive vocabulary (PPVT) data (n = 6), one percent were missing word knowledge at pretest (n = 3), and four percent were missing general expressive vocabulary at pretest (n = 10). There was no missing data on child age, race/ethnicity, gender, or AM/PM status.

Missing data analyses suggested that data were missing at random for general receptive vocabulary pretest, but not missing at random for general expressive vocabulary or word knowledge at pretest[16]. Although the patterns for non-randomness were different across these three independent variables, means were imputed for missing data on all three initial measures.

To do this, rather than simply imputing the mean of the treatment or comparison group on a particular measure, the mean of other children of the same gender, race/ethnicity, age (plus or minus 2 months), and with the same teacher were imputed. Race/ethnicity and teacher were chosen for stratified imputation because there were significant differences between racial/ethnic groups and teachers on each of the continuous pretest measures. In addition, age was chosen because there was a significant correlation between age and each continuous pretest measure. Although there were no statistically significant differences between boys and girls on pretest measures, the trend was girls outperforming boys on each pretest measure so it was also included.

Imputing such a small number of values had no appreciable effect on the overall mean score within each treatment condition or across treatment conditions (e.g. overall pre-imputation PPVT standard score mean was 89.47 and the overall post-imputation mean was 89.72). This was true for all three outcomes. Though imputation addressed a

[16] There were no differences between children general receptive vocabulary and children not missing on word knowledge pretest, age, race/ethnicity, whether attending AM or PM, or gender. Children who were missing word knowledge at pretest were more likely to be a minority and to have a lower general receptive vocabulary at pretest than children who were not missing. Children who were missing general expressive vocabulary at pretest were more likely to attend AM preschool and be in the treatment condition.

relatively small number of missing cases, it was conducted to have the greatest possible statistical power for detecting differences between groups.

Continuous Outcome Measures. Three percent of children were missing general receptive vocabulary at posttest (n = 9), ten percent were missing on word knowledge at posttest (n = 25), and ten percent were missing on general expressive vocabulary (n = 26). Missing data analyses indicated that data were not missing at random[17]. However, rather than imputing means on important outcome variables, children with missing data on outcomes were excluded from analyses. For this reason, sample sizes were slightly different depending on the outcome used.

Procedure

Teacher Training. Two weeks before the start of the study, all treatment teachers received one day of intensive training that provided teachers with an overview of the eight-day sequence they would find related all topics. The training included complete review of curricular materials, modeling of lessons for teachers by the trainers, time to review the materials, and time to read through the teacher manual with one another. In addition to the initial training, project members ran focus groups with treatment teachers in each county twice a month. Focus groups served as a platform to ask teachers about what seemed to be working well as they implemented the curriculum and what they were having trouble with. Researchers strategized with teachers about how to surmount challenges they were having; in this way the focus groups served as ongoing training.

---

[17] Children that were missing receptive vocabulary at posttest scored significantly lower on receptive vocabulary at pretest than children not missing. Children missing on word knowledge at posttest scored significantly lower on word knowledge at pretest and were younger than children not missing. Children missing expressive vocabulary posttest scored significantly lower on receptive vocabulary pretest, word knowledge pretest, and were younger than children not missing.

Implementation of the WOW Intervention. Teachers began implementing the WOW curriculum in January 2007. The curriculum was designed to be a supplement to the existing literacy curriculum. Teachers implemented the lesson in a whole group setting during one of their daily circle times. They began with Unit 1: Living Things. This unit included, in order, Pets, Wild Animals, Animals in Water, and Insects. Next, they moved on to Unit 2: Health. This included Parts of the Body, Clothes, and Exercise. Because Head Start has a four-day week, completing the eight-day sequence for each taxonomic category generally took two weeks (due to holidays, teacher release days, conferences, and other interruptions, there were some cases where teachers took longer than eight days to complete a given sequence).

Fidelity of Implementation. A critical piece of determining internal validity of an intervention study is determining that the independent variable of interest, the intervention, has been implemented as intended (O'Donnell, 2008). In a meta-analysis of how fidelity of implementation was measured in a number of curriculum intervention studies, O'Donnell (2008) concluded that though fidelity of implementation was defined in many ways by different researchers, across studies it seemed to be synonymous with "adherence and integrity".

To ensure that teachers in the treatment group were administering the curriculum similarly, fidelity observations were conducted. Researchers observed each teacher for fidelity an average of two times during each topic (e.g. Teacher A was observed two times during the eight days she taught the Insects instructional sequence), alternating between AM and PM classes. Over the four month study researchers conducted a total of

54 observations in treatment classrooms. Each teacher was observed an average of 11 times.

During each visit, the researcher observed the teacher enact one daily lesson from the WOW curriculum and used an observation protocol to determine fidelity to this lesson. The protocol included the actual pages of the teacher's manual from the appropriate lesson; the observer was expected to take careful note of if, when, and how the teacher deviated from the WOW lesson. In addition, the observation included space for observational notes and for gathering information on the number of children present, start time, and end time. After the observation, the researcher wrote up anecdotal notes in narrative form to augment the observation protocol.

To determine overall fidelity to the curriculum across observations, each observed lesson was first broken into major components. For example, a particular lesson in the Wild Animals instructional sequence included seven major components: a) a discussion to review what they already know about wild animals, b) watching a Sesame Street clip of a girl observing wild animals in the zoo, c) discussion of the animals depicted in the video, d) reading two pages from the informational text, e) discussion after reading the text, f) review and discussion of "challenge" items previously taught, and g) introduction of new challenge items. This is an example from one lesson; the number of major lesson components varied depending on the lesson observed.

Next, each component was assigned one point. Then, using the observation protocol, each teacher received one point for each component enacted (and conversely, zero points if the component was not enacted). To calculate a percentage that represents how well the teacher adhered to the lesson, the total number of components enacted by the teacher was divided by the total number of components in the lesson. These percentages were then averaged across observations for each teacher and across teachers. Using this procedure, average fidelity to the curriculum across teachers was 88%.

Analytic Method

To analyze these data, three steps were used. First, descriptives were analyzed for mean differences between treatment and comparison groups. Next, multivariate analyses were conducted using hierarchical linear modeling. Then effect sizes were calculated for each outcome. This section describes each step in detail.

Descriptive Analyses

Descriptive analyses were conducted to investigate differences in means between the treatment and comparison groups at pretest, immediately following instruction, and at the end of the study on word knowledge, conceptual knowledge, transfer of conceptual knowledge to new learning tasks, and general vocabulary knowledge. T-tests were used to determine if there were significant differences between the two groups.

Multivariate Analyses

Multivariate analyses were used to investigate each of the following research questions. Compared to the comparison group:

1. Do children who experience the curriculum learn the words that were taught?
2. Do children who experience the curriculum acquire the concepts that were taught?
3. Do children who experience the curriculum transfer learned concepts to new learning tasks?

4. Do children who experience the curriculum show more growth in general receptive and expressive vocabulary?

Use of Multi-Level Modeling. In the WOW vocabulary intervention study, children were situated in classes (AM or PM), and classes were nested within teachers (almost all teachers taught both AM and PM preschool). Because of this, it would be erroneous to assume that children were independent from one another; on the contrary, many children shared the same class and/or teacher. To deal with this dependence problem, hierarchical linear modeling (HLM) (Raudenbush & Bryk, 2002) was used. Multi-level modeling isolates the effects of teachers from individual effects and also resolves the problems of codependence that plague ordinary regression techniques by nesting children (Level 1) in a larger unit, either classes of children (in AM or PM) or teachers.

Rationale for Teacher as Level 2 Unit of Analysis. In this study, it is clear that children are the Level 1 unit of analysis. However, there are two approaches to determining the Level 2 unit of analysis. One approach would be to use “class” groupings of children, where children are nested in the AM or PM class that their teacher taught. The other approach would be to use “teacher” groupings of children, where children are nested in their teacher.

There are benefits and drawbacks to both approaches. Using “class” as the Level 2 unit of analysis increases the number of Level 2 units from 12 to 22, thus providing more statistical power for modeling. However, using “class” as the Level 2 unit of analysis does not take into account that 10 of the 12 teachers taught two classes (both AM and PM classes). Thus, these classes, and children within these classes, are not independent from one another. On the other hand, using "teacher" as the Level 2 unit of analysis by grouping together children that attended AM and PM for one teacher does not address the fact that a child in Teacher A's *AM* class was surrounded and influenced by a different groups of peers than another child in Teacher A's *PM* class[18].

Ultimately, there are two reasons that "teacher" was chosen be the Level 2 unit of analysis in this dissertation study. First, the original design specified that teachers, and not AM or PM classes of children, were assigned to treatment or comparison condition. Secondly, the teacher was more likely to have an influence on children's vocabulary and conceptual development than their peer group.

Procedure for HLM Analyses. Several outcomes were used to investigate each research question. To investigate each individual outcome, four steps were taken: a) running a fully unconditional model to partition variance in the outcome between individual and teacher effects, b) running a within-teacher model to verify that the relationship between treatment condition and the outcome varies between classrooms, c) running an intercept-as-outcomes model to investigate the unique effects of treatment condition on each outcome (Raudenbush & Bryk, 2002), and d) conducting a residual analysis.

Fully Unconditional Model. The first step in each analysis was running a fully unconditional model with only the outcome specified. There were two critical pieces of information gleaned from unspecified models. First, the fully unconditional model determined if there was statistically significant variability between teachers on the outcome. Second, the fully unconditional model indicates the proportion of variability

[18]For this reason, whether the child attended AM or PM kindergarten was initially included in all analyses as a way to control for "peer" effects.

that exists between children within teachers and the proportion that exists between teachers. To determine the intraclass correlation (ICC), or the proportion of variability in the outcome that occurs between teachers, the variance components τ (or tau, the between teacher-variance) and σ^2 (or sigma squared, the within-teacher variance) were used in the following equation:

$$\text{Intraclass correlation (ICC)} = (\tau \ / \ (\tau + \sigma^2)) * 100$$

The remaining proportion of variability exists between children, which is derived by simply subtracting the ICC from 1[19]. The goal of HLM models is to try to explain the variability that exists at each level using important predictor variables.

Within-Teacher Model. For each outcome, after running the fully unconditional model, the next step was to specify the Level 1 model. When specifying the child-level models (Level 1) of each HLM analysis, children's word knowledge at the beginning of the study was included as a predictor to control for initial differences between treatment and comparison group. In addition, when available, the pretest version of the outcome was used as a predictor. When initially specifying the child-level models, important demographic variables were included in the model, but were removed if they were not significantly related to the outcome. These demographics were gender, minority status, age in months, whether the child attended AM or PM preschool, and total number of days absent during the study.

---

19 When analyzing data where children are nested in classrooms, such as in this study, HLM is superior to ordinary least squared regression because it addresses and accounts for the nested nature of the data and related lack of independence between children in the sample. By doing so, it allows the researcher to specify models that explain variability both between and within groups. Though HLM can technically be used to explain variability between groups even if the between-group variation is very small (e.g. ICC = 2%), this usually means that there is very little about "groupness" or nesting that has an influence on the outcome. Therefore, the larger the ICC, the more influence teacher-level variables (in this case, treatment condition) had on the outcome.

When including predictors in HLM models (both at Level 1 and 2), a decision must be made regarding how to "center" each variable. Centering decisions influence the meaning of zero for each predictor and therefore the interpretation of the intercept. When specifying models for this study, continuous predictors were grand-mean centered, such that 0 was the mean of the whole sample. Dichotomous predictors were entered into the models uncentered, therefore 0 was the reference group.

After specifying the Level 1 model, the proportion of the variability within classrooms (σ^2) that was explained by the Level 1 model was calculated using the following equation:

$$\text{Proportion of } \sigma^2 \text{ explained} = \frac{\sigma^2 \text{ from the Fully Unconditional Model} - \sigma^2 \text{ from the Level 1 model}}{\sigma^2 \text{ from the Fully Unconditional Model}}$$

Between-Teacher Model. Finally, the fully conditional model was specified for each outcome. In this intervention study, treatment condition was always the Level 2 variable of interest[20]. Comparing performance of children in the treatment classrooms with performance of children in the comparison classrooms bolstered confidence in causal inferences that could be made and ruled out maturation as an explanation for growth and learning occurring in the treatment group. Although it would have been interesting to conduct further investigation of other classroom or teacher characteristics that predicted children's outcomes or the relationship between two variables, the statistical power available to detect relationships was limited by the small number of

[20] Although treatment condition was the main variable of interest, models were also run including classroom literacy environment, or ELLCO score, as a Level 2 predictor. It was not a significant predictor of children's performance on any outcomes.

Level 2 units (teachers). Therefore, Level 2 models were kept simple, with a focus on the effect of treatment condition on children's learning.

The basic, fully conditional HLM model used to address each of the research questions is depicted below (with some minor variation depending on the significance of Level 1 predictors or existence of a relevant pretest):

Level 1 Model

Outcome = β_0 + β_1 (Word Knowledge at Time 1) + β_2 (Pretest Corresponding to Outcome) + r

Level 2 Model

$\beta_0 = \gamma_{00} + \gamma_{01}$ (Treatment Condition) + u_0

$\beta_1 = \gamma_{10}$

$\beta_2 = \gamma_{20}$

β_0 = Mean outcome score for the sample
β_1 = Mean relationship between the outcome and word knowledge at pretest
β_2 = Mean relationship between the outcome and the associated pretest
r = Level 1 error, or how much each child in the sample differs from the overall mean
γ_{00} = Mean outcome score for each teacher
γ_{01} = The influence of treatment condition on the mean for each teacher
u_0 = Level 2 error, or how much the mean of each teacher differs from the overall mean

Finally, the percent of variance between teachers (τ_{00}) that was explained by the Level 2 predictors was calculated. In this case, the only predictor at Level 2 was treatment condition. To determine what proportion of the variance between teachers was explained by the Level 2 model, the following equation was used:

$$\text{Proportion of } \tau_{00} \text{ explained} = \frac{\tau_{00} \text{ from the Level 1 model} - \tau_{00} \text{ from the Level 2 model}}{\tau_{00} \text{ from the Level 1 Model}}$$

Residual Analysis. One goal of regression analyses, including HLM, is to create a fully specified model that will explain all of the systematic variability in the outcome. When this is not possible, the residuals represent the variability left unexplained. If there is systematic variation in the residuals, this indicates that the model was misspecified or that there were issues with the psychometric properties of either predictors or outcomes (Cohen, Cohen, West, & Aiken, 2003). If there is *not* systematic variation in the residuals, the following assumptions will be upheld: a) Level 1 error (r) is independent, normally distributed, has a mean of 0, and constant variance, b) all Level 1 predictors are independent of the error, c) Level 2 error (u) is independent, normal, has a mean of 0, and constant variance, and d) all Level 2 predictors are independent of the error (Raudenbush & Bryk, 2002).

To investigate if there was systematic variation in the residuals, and therefore if the assumptions of the HLM models are upheld, a residual analysis was conducted for each fully specified model at both Level 1 and Level 2. To evaluate assumption a, a histogram fitted with a normal curve determined if the Level 1 errors were normally distributed above and below the mean. To determine if Level 1 errors had a mean of 0 and constant variance, three strategies were used: creating a boxplot representing the residuals of each teacher around 0, a scatterplot of the fitted values for each child and the residuals, and statistically testing for homogeneity of Level 1 variance. To evaluate assumption b, that all Level 1 predictors are independent of the errors, scatterplots were created for the residuals and each individual predictor.

The residual analysis at Level 2 entailed determining if assumptions c and d were upheld. To ascertain if assumption c was valid (that u was independent, normally distributed, has a mean of 0, and constant variance) two steps were taken: a QQ plot was created with observed values on the x-axis and expected values on the y-axis and a scatterplot was created of fitted values for the intercept and the residuals for the intercept. Lastly, to examine assumption d, that all Level 2 predictors are independent of the error, a scatterplot was created for treatment by residuals from the intercept.

Calculating Effect Sizes

Using posttest means and standard deviations for each outcome, effect sizes were calculated. To calculate Cohen's *d* effect sizes, the following equation was used:

$$d = (\text{mean}_T - \text{mean}_C)/s_p$$

T is the treatment group, C is the comparison group, and s_p is the pooled standard deviation of the two groups. To calculate the pooled standard deviation, the following equation was used:

$$s_p = \sqrt{(((n_T-1)*s_T^2 + (n_C-1)*s_C^2)/(n_T+n_C-2))}$$

N is the sample size of a particular group and s is the standard deviation of a particular group. In general, an effect size of .2 is considered small, .5 medium, and .8 large (Cohen, 1988).

Outcomes and Predictors Used to Address Each Research Question

The next section describes in detail the outcomes used to address each question, covariates included in each model, and any additional analytic procedures used. Table 6 summarizes the outcomes used and the covariates used in each model at both Level 1 and Level 2.

Table 6

HLM Outcomes and Associated Level 1 and Level 2 Predictors

Outcome	Level 1 Predictors	Level 2 Predictors
Word Knowledge		
Word Knowledge Time 2	• Word Knowledge Time 1	• Treatment Condition
Word Knowledge Time 3	• Word Knowledge Time 1	• Treatment Condition
More Difficult Word Usage Time 3	• Word Knowledge Time 1 • More Difficult Word Usage Time 1	• Treatment Condition
Conceptual Knowledge		
Conceptual Categorization of Items Taught Time 2	• Word Knowledge Time 1 • Word Knowledge Time 2 (items taught) • Age in months • Attends PM preschool	• Treatment Condition
Conceptual Knowledge Usage Time 3	• Word Knowledge Time 1 • Conceptual Knowledge Usage Time 1	• Treatment Condition
Transfer of Conceptual Knowledge		
Conceptual Categorization of Items Not Taught Time 2	• Word Knowledge Time 1 • Word Knowledge Time 2 (items not taught) • Age in months • Days absent	• Treatment Condition
Justification of Conceptual Categorization of Items Not Taught Time 2	• Word knowledge Time 1 • Word knowledge Time 2 (items not taught) • Age in months • Days absent	• Treatment Condition
General Word Knowledge		
General Receptive Word Knowledge Time 3	• Word Knowledge Time 1 • General Receptive Word Knowledge Time 1	• Treatment Condition
General Expressive Word Knowledge Time 3	• Word Knowledge Time 1 • General Expressive Word Knowledge Time 1	• Treatment Condition

It is important to note that for each HLM analysis, each of the following child level predictors were initially entered into the Level 1 equation: age, gender, attending

AM or PM preschool, days absent, and race/ethnicity. However, if the predictor was not significant in the model, it was removed. Table 6 and the following paragraphs describe the final models, after this process was completed.

Research Question 1: Do children who experience the curriculum learn the words that were taught? This research question investigates the hypothesis that the children in the treatment group demonstrated significantly more curricular word knowledge and usage than the children in the comparison condition. Three outcomes were used to investigate this research question: a) word knowledge at Time 2 (immediately following each instructional sequence), b) word knowledge at Time 3 (at the end of the study), and c) more difficult word usage at Time 3 (see Table 6 for the predictors included in HLM models at both the Level 1 (child level) and Level 2 (teacher level)). In each of these models, the appropriate pretest and word knowledge at Time 1 were included as controls.

Research Question 2: Do children who experience the curriculum learn the concepts that were taught? This research question investigates the hypothesis that children in the treatment condition learned the concepts that were taught. Two outcomes were used to investigate these research questions. The first outcome was children's conceptual categorization of items taught in the curriculum (measured at Time 2); this measure assessed children's ability to use what they learned in the curriculum to categorize items that were explicitly taught in the curriculum. To isolate conceptual knowledge from word knowledge, children's word knowledge of items they were asked to categorize (Word knowledge at Time 2) was included as a control variable. Word knowledge at Time 1 was included to control for initial group differences. In addition, age in months and whether a child attended AM or PM preschool were significant predictors, so were kept in the model.

The second outcome used to assess children's conceptual knowledge was children's conceptual knowledge usage at Time 2. Conceptual knowledge usage at Time 1 and word knowledge at Time 1 were included as controls. In addition, because this outcome was created from children's verbal responses, specific examples of children's conceptual knowledge usage are presented to illustrate different types of responses.

Research Question 3: Do children who experience the curriculum transfer conceptual knowledge to new learning tasks? This research question investigates the hypothesis that children who experienced the WOW curriculum were able to transfer and apply the conceptual knowledge they learned in the curriculum to make inferences in new learning situations. Evidence of this type of application suggests that children may be able to use the type of knowledge acquired through this intervention to foster future learning and comprehension.

Two outcomes were used to investigate this research question. The first outcome was children's conceptual categorization of items not taught in the curriculum (measured at Time 2, or immediately following each instructional sequence). Children's word knowledge of these items was included as a covariate to control for prior knowledge, and word knowledge at Time 1 was included to account for group differences. In addition, age in months and days absent were included because they were significant in the model.

The second outcome used to investigate transfer of conceptual knowledge was children's verbal justification of their inferences and resulting categorization choices. As described earlier, children were asked *why* they made the inference that they did (e.g.

"How did you know that a raccoon is a wild animal?"); responses were coded as Organized (1 point) or Diffusely Organized (0 points), summed, then averaged across all topics. In short, this outcome assesses how children were able to use, apply, and express their conceptual knowledge when confronted with new learning tasks. To account for differences in word knowledge of items categorized, children's word knowledge of these items was included as a covariate. Word knowledge at Time 1, days absent, and age at pretest were also included in the model (the former to account for pretest group differences and the latter two because they were significantly related to the outcome).

An additional analysis was conducted to illustrate the nature of children's justifications. The outcome described in the previous paragraph, designed to investigate children's ability to use conceptual knowledge to justify their categorization choices, represented a total score. To look more closely at the quality of children's justifications, the average proportion of children in each treatment condition who provided Organized or Diffusely Organized responses was calculated. In addition, examples of each type of response were randomly chosen to illustrate the quality of each type of justification.

Research Question 4: Do children who experience the WOW curriculum show more growth on general vocabulary measures than children in the comparison group? This research question investigates if children in the treatment group showed more growth on general receptive and expressive word knowledge measures than children in the comparison condition. For general receptive word knowledge, the standard Peabody Picture Vocabulary Test (PPVT) score served as the outcome measure. For general expressive word knowledge, the Get It, Got It, Go! (GGG) Rapid Naming assessment was used as the outcome measure. Although these are different assessments, both in the administration procedures and in the constructs they attempt to capture (expressive vs. receptive vocabulary; rapid naming vs. untimed), results from both of these measures was used to determine whether experiencing the WOW curriculum had a measurable influence on general expressive or receptive word knowledge.

Limitations to Methods

There are two potential limitations related to my methodological approaches. First, in order to ensure parity between the treatment and comparison groups, I matched students in both groups on their general receptive vocabulary scores (PPVT) at pretest. This substantially decreased the sample size, thereby decreasing statistical power. This could potentially have the effect of creating a Type II error, or finding no effect when an effect actually exists (due to limited statistical power) (Shavelson, 1996). However, finding statistically significant effects of the intervention *despite* limited statistical power only bolsters claims about the efficacy of the intervention.

To address this issue, these analyses were run using four possible approaches: 1) using OLS regression and the full, non-matched sample with PPVT as a covariate, 2) using OLS regression and the sample matched on PPVT, 3) using HLM and the full, non-matched sample with PPVT as a covariate, and 4) using HLM and the sample matched on PPVT. Results were robust across each approach, both in significance and direction of the effects. Therefore, I chose to adopt the most conservative approach in this dissertation study: using HLM and the sample matched on PPVT. See Appendix I, Tables A-E for a summary of these different analyses.

A second issue regarding statistical analyses is related to decisions I made when preparing data for HLM analyses. Determining the Level 2 unit of analysis was a complex endeavor. On the one hand, children were nested in their particular class of children attending either AM or PM preschool. Further, AM and PM classes of children were nested in teachers (almost all teachers taught both AM and PM classes). For children, both their peers in their particular AM or PM group *and* their teacher might potentially influence their learning. One approach, given the structure of these data, might have been to conduct 3-level HLM's, with children nested in AM or PM classes, and AM/PM classes nested in teachers. However, given how small the sample sizes would be at Level 2 (2; one AM and one PM class per teacher) and at Level 3 (12 teachers), it is unlikely that this approach would garner the necessary statistical power to detect treatment effects. Given this problem, another approach would have been to nest children within their AM and PM class. However, doing so would ignore the fact that AM and PM classes were not independent from one another-there are always two classes, or Level 2 units, that were taught by the same teacher.

The alternative approach, which I ultimately chose, was to nest children within their teacher. The strength of this approach is that treatment condition was originally assigned to teachers, not to AM or PM classes. The limitation to this tactic is that by grouping children in AM and PM classes together, any influence of peer group on learning was masked. However, my logic was that the teacher likely has a stronger influence on children than their peers. Further research on this curriculum would be enhanced by any of the following improvements: increasing the sample size such that a three-level HLM is possible, including only full-day preschool, or increasing the sample size such that the curriculum would be implemented and data collected in only one class (AM or PM) for each teacher.

Conclusion

The supplementary WOW vocabulary curriculum was designed to foster word knowledge, conceptual knowledge, and transfer of conceptual knowledge in low-income preschoolers in an effort to begin to narrow the vocabulary gaps that we know exist between children from disparate backgrounds. It is based on the hypothesis that teaching conceptual knowledge and associated vocabulary to children in well-organized taxonomic categories can provide a foundation and structure that fosters future learning. Taken together, results from the analyses described in this chapter address if the WOW curriculum, and the approach of teaching words and concepts in taxonomic categories, had a measurable and powerful effect on children's curriculum-specific word knowledge, conceptual knowledge, ability to transfer and apply conceptual knowledge to new learning situations, and general vocabulary knowledge.

Chapter 4

Results

The primary goal of this study was to investigate the overall efficacy of the WOW curriculum in increasing vocabulary and conceptual knowledge and fostering the ability to make inferences in new learning situations. This dissertation study investigates several research questions. Compared to the comparison group:

1. Do children who experience the curriculum learn the words that were taught?
2. Do children who experience the curriculum acquire the concepts that were taught?
3. Do children who experience the curriculum transfer learned concepts to new learning tasks?
4. Do children who experience the curriculum show more growth in general receptive and expressive vocabulary?

This section reports descriptive and multivariate findings investigating if children who experienced the WOW curriculum acquired the vocabulary and conceptual knowledge taught in the curriculum. In addition, this section presents results addressing if children who experienced the intervention were able to transfer and apply knowledge acquired from the curriculum when making inferences in new learning situations. It also includes findings from analyses investigating if the WOW curriculum exerts an influence on general receptive and expressive vocabulary. In addition to presenting descriptive and multivariate results, this section includes anecdotal examples of children's actual responses in the conceptual knowledge and conceptual knowledge transfer tasks.

Descriptive and Multivariate Results

In this section, results are presented based on relevance to each individual research question. Therefore, there are several different results tables. For the reader that would like to see one table that depicts all of the results in one place, see Appendix I, Table D.

Research Question 1: Do children who experience the curriculum learn the words that were taught? To address this question, descriptive analyses were conducted and hierarchical linear models were created using the following three outcomes: a) word knowledge at Time 2 (immediately following instruction), b) word knowledge at Time 3 (at the end of the study), and c) more difficult word usage at Time 3. Results are presented in the order listed above.

Word Knowledge Time 2. This measure captures if children were able to expressively identify topic-specific words immediately following each 8-day lesson sequence. As a reminder, children were asked "What is it?" for 4-6 words following immediately following each instructional sequence, a percentage correct was calculated at each testing point, and finally an overall average percent correct was calculated across topics. The ability of children in the treatment group to outperform similar children in the comparison condition on this measure would suggest that the instruction and activities included in the WOW curriculum fostered immediate word knowledge.

Initial investigation into mean differences between the treatment and comparison groups indicated that following instruction in each topic, there were significant differences between the two groups on word knowledge at Time 2 ($t = 7.83$, $p \leq .001$) (see Table 7 for group means). Children in the treatment group, on average, correctly identified 72% of curricular words assessed immediately after instruction, while children in the comparison only correctly identified 45% of words, on average.

Table 7

Means and Standard Deviations for Word Knowledge at Time 2

	Word Knowledge Time 2 Mean % Correct[21]	SD
Treatment Group	72***	18
Comparison Group	45	20

*p<.05; **p<.01; ***p<.001

When word knowledge at Time 2 was used as an outcome in an HLM model, the fully unconditional model indicated that there was significant ($\chi = 86.08$; $p \leq .001$) and sufficient variability between teachers (ICC = 36%) on word knowledge at Time 2 (immediately following instruction) (see Table 8). Word knowledge at Time 1 significantly predicted performance on word knowledge at Time 2 ($\beta = .56$, $p \leq .001$) (see Table 9)[22]; it explained 42% of the variability between children. Holding children's word knowledge at Time 1 constant, children in treatment teachers' classrooms significantly outperformed children in comparison teachers' classrooms on word knowledge at Time 2 ($\gamma = .84$, $p \leq .001$); treatment condition explained 87% of the variability between

[21] As described in Chapter 3, children were administered this assessment after experiencing instruction (or not) in each taxonomic category. This means that at six different time points, children were assessed on the words that relate to what they had just been taught. For five of the categories, children were assessed on six words, but for one, children were assessed on four. To get a total score, a percentage correct was calculated at each time point and then percentage correct was averaged across all topics for the score you see in this table.

[22] It is also important to remind the reader that in each analysis, all of the following covariates were initially entered as Level 1 predictors: minority status, age, gender, attends PM preschool, and days absent during the study. Those predictors that were not significant were removed from the final models.

teachers. Though there was still significant variability between teachers after accounting for treatment condition, inclusion of this variable greatly reduced the amount of variation between teachers from the fully unconditional model (from $\chi = 86.08$; $p \leq .001$ in the fully unconditional model to $\chi = 18.79$, $p \leq .05$ in the fully conditional model). This was a large effect; the Cohen's *d* effect size was 1.37.

Table 8

Fully Unconditional Hierarchical Linear Modeling (HLM) Models: Psychometric Characteristics of the Word Knowledge Outcome Variables

	Word Knowledge Time 2 (n=139 children in 12 teachers' classes)	Word Knowledge Time 3 (n=164 children in 12 teachers' classes)	More Difficult Word Usage Time 3 (n=171 children in 12 teachers' classes)
Intercept	-.16	-.00	.01
Between-teacher variance (tau)	.38	.11	.20
Within-teacher variance (sigma squared)	.66	.90	.82
Intraclass correlation (ICC) (%)	36	11	20
Reliability	.85	.60	.76

Note: The intraclass correlation (ICC) is the percentage of total variance in the outcome that lies systematically between teachers. It is calculated as follows: ICC (%) = (tau/tau + sigma squared) * 100.

Table 9

Fully Conditional HLM Models Estimating the Effects of WOW on Word Knowledge

Variable	Word Knowledge Time 2	Word Knowledge Time 3	More Difficult Word Usage Time 3
	β / γ[a] (S.E.)	β / γ (S.E.)	β / γ (S.E.)
Within-Teacher Variables			
More Difficult Word Use Time 1			.33*** (.07)
Word Knowledge Time 1	.56*** (.06)	.72*** (.05)	.29*** (.07)
Between-Teacher Variables			
Treatment	.84*** (.15)	.34** (.10)	.71*** (.14)
Between-teacher variance (tau)	.03	.00	.01
Within-teacher variance (sigma squared)	.39	.43	.57

*p<.05; **p<.01; ***p<.001

a. β coefficients correspond to within-teacher variance and γ coefficients correspond to between-teacher variance

Word Knowledge Time 3. This measure was an expressive measure of curriculum-specific word knowledge at Time 3, or at the end of the intervention. This measure was created by summing children's correct responses when asked "What is it?" to 35 separate curriculum-specific words. The ability of children in the treatment group to outperform their comparison group peers would be an indication that word knowledge that was learned during the study was maintained over the course of the study.

Prior to treatment (at Time 1), mean scores on curriculum based word knowledge were significantly different in treatment and comparison classrooms ($t = 2.9$, $p \leq .01$) (see Table 10 for group means). Children in the treatment group knew about 16 of the 35 words on the assessment, while children in the comparison group knew 14. Following the treatment, there were also significant differences between groups on word knowledge (and t = 4.08, $p \leq .001$). After the treatment, children in the treatment group knew, on average, between 20 and 21 curricular words and children in the control group knew about 17 words. Importantly, the magnitude of the difference between groups was larger after the intervention than before it began, which is an initial indication of somewhat more learning over the course of the study in treatment classrooms.

Table 10

Means and Standard Deviations for Word Knowledge at Time 1 and Time 3

	Word Knowledge Time 1 (total)	SD	Word Knowledge Time 3 (total)	SD
Treatment Group	16.31**	4.92	20.47***	5.27
Comparison Group	14.11	5.20	16.98	5.68

*p<.05; **p<.01; ***p<.001

To investigate this supposition more rigorously, an HLM model was constructed using word knowledge at Time 3 as the outcome. The fully unconditional model demonstrated that there was significant ($\chi = 29.22$, $p \leq .01$) and sufficient variability between teachers (ICC = 11%) on word knowledge at Time 3 (see Table 8). As expected, word knowledge at Time 1 significantly predicted word knowledge at Time 3 ($\beta = .72$, $p \leq .001$) (see Table 9); it explained 51% of the variability between children. After controlling for word knowledge at Time 1, children in treatment teachers' classrooms significantly outperformed children in comparison teachers' classrooms on word knowledge at Time 3 ($\gamma = .34$, $p \leq .01$); treatment condition explained 99.6% of the variability between teachers. In fact, treatment condition explained *all* of the variation between teachers on word knowledge at Time 3, such that it was no longer significant (χ

= 6.51, ns). The effect of the intervention on word knowledge at Time 3 was moderate (Cohen's *d* = .63).

More Difficult Word Usage Time 3. This outcome measured if children who experienced the intervention learned and were able to use the more difficult words that were taught in the curriculum. Children's responses when told "Tell me everything you know about *name of the topic*" were searched for use of 92 more difficult words that were taught in the curriculum and this measure is the total number of more difficult words used.

Prior to the intervention, there were no significant differences between the treatment and comparison groups in scores on more difficult word usage ($t = .30$, ns). Students in both groups used about four "more difficult" curricular words. Following the intervention, there were significant differences between groups ($t = 5.49$, $p \leq .001$) (see Table 11 for group means), with treatment children using almost seven more difficult words and children in the comparison group using less than four. This is an initial indication that experiencing the curriculum contributed to children's growth in the ability to use more difficult words.

Table 11

Means and Standard Deviations for More Difficult Word Usage at Time 1 and Time 3

	More Difficult Word Usage Time 1 (total)	SD	More Difficult Word Usage Time 3 (total)	SD
Treatment Group	4.73	3.42	6.75***	3.65
Comparison Group	4.56	3.57	3.94	3.01

*p<.05; **p<.01; ***p<.001

When investigating more difficult word usage using an HLM model, I found that a sizable proportion of variation in this outcome did lie between teachers (ICC =20%), and this was significant ($\chi = 48.44$, $p \leq .000$) (see Table 8). Children's more difficult word usage at Time 1 significantly predicted more difficult word usage at Time 3 ($\beta = .33$, $p \leq .001$) (see Table 9). Word knowledge at Time 1 was also a significant predictor ($\beta = .29$, $p \leq .001$). Together, more difficult word usage at Time 1 and word knowledge at Time 1 explained 31% of the variability between children on the outcome. After controlling for these important variables, there was a significant difference between children in treatment classrooms and children in comparison classrooms on more difficult word use at Time 3, with children in treatment classrooms outperforming children in comparison classrooms ($\gamma = .71$, $p \leq .001$). This was a large effect (Cohen's *d* = .84). Treatment condition explained 92% of the variability between teachers on more difficult word use at Time 3, which decreased it to non-significance ($\chi = 14.30$, ns).

Taken together, findings from these three analyses demonstrate that children who experienced the WOW curriculum had significantly more curriculum-specific word knowledge and ability to use curricular words than children in the comparison group, both immediately following instruction and at the end of the study. Because the two groups were statistically equivalent on general word knowledge (both on the general expressive and receptive word knowledge measures) at Time 1, and models controlled for the small differences between the groups on curriculum word knowledge at Time 1, these analyses are a first step in demonstrating a link between the WOW curriculum and children's growth in curriculum-specific word knowledge and use.

Research Question 2: Do children who experience the curriculum acquire the concepts that were taught? To investigate if children who experienced the curriculum learned the concepts that were taught and displayed knowledge of those concepts significantly more than children in the comparison group, HLM was used to analyze two outcomes: a) children's conceptual categorization of items that were taught in the curriculum at Time 2 (immediately following instruction), and b) children's conceptual knowledge usage at Time 3.

Conceptual Categorization of Items Taught Time 2. This measure investigated if children were able to demonstrate their knowledge of the concepts taught in the curriculum in a categorization task. Immediately following instruction of each curricular topic, children were asked to categorize items that were explicitly taught in the curriculum as either in or out of the category. Their correct responses were summed, and divided by the total number of items to get a percent correct. Finally, an average percent correct across topics was calculated. Treatment children's ability to outperform children in the comparison group when asked to categorize items that were taught in the curriculum would demonstrate that children learned the concepts that were taught (in this case, the concepts underlying category membership of each taxonomic category).

Immediately following instruction, scores on conceptual categorization of items taught in the curriculum were significantly different in the treatment group than the comparison group ($t = 7.71$, $p \leq .001$) (see Table 12 for group means). Following instruction, on average, children in the treatment group were able to correctly categorize almost 80% of the items. On average, children in the comparison group correctly categorized about 60% of items.

Table 12

Means and Standard Deviations of Conceptual Categorization-Items Taught

	Conceptual Categorization of Items Taught Time 2 Mean % correct	SD
Treatment Group	78***	12
Comparison Group	62	12

*p<.05; **p<.01; ***p<.001

A more robust analysis of children's ability to categorize items taught in the curriculum was conducted using an HLM model. The fully unconditional model demonstrated that there was significant (χ = 83.65, p $\leq$ .001) and sufficient (ICC=37%) variability between teachers on children's ability to conceptually categorize items that were taught (see Table 13). Word knowledge at Time 2 (β = .36, p $\leq$ .001), age in months (β = .01, p $\leq$ .01), and attending PM preschool (β = -.31, p $\leq$ .01) were all significant predictors of conceptual categorization of items taught in the curriculum (see Table 14); together these variables explained 32% of the variability between children (word knowledge at Time 1 was entered in the model but was not a significant predictor). Even after controlling for these significant Level 1 variables, children in treatment teachers' classrooms significantly outperformed children in comparison teachers' classrooms on conceptual categorization of items taught in the curriculum (γ = .62, p $\leq$ .01). This was a large effect (Cohen's *d* = 1.35). Treatment condition explained 83% of the variability between teachers on this outcome, diminishing the unexplained variability between teachers (variability in the fully unconditional model χ = 83.65, p $\leq$ .001 vs. variability in the fully conditional model χ = 15.90, ns) to non-significance.

Table 13

Fully Unconditional Hierarchical Linear Models (HLM): Psychometric Characteristics of the Conceptual Categorization of Items Taught and Conceptual Knowledge Usage Outcome Variables

	Conceptual Categorization of Items Taught Time 2 (n=139 children in 12 teachers' classes)	Conceptual Knowledge Use Time 3 (n=152 children in 12 teachers' classes)
Intercept	-.16	-.03
Between-teacher variance (tau)	.40	.20
Within-teacher variance (sigma squared)	.67	.81
Intraclass correlation (ICC) (%)	37	20
Reliability	.85	.73

Note: The intraclass correlation (ICC) is the percentage of total variance in the outcome that lies systematically between teachers. It is calculated as follows: ICC (%) = (tau/tau + sigma squared) * 100.

Table 14

HLM Investigating the Effect of Treatment on Conceptual Categorization of Items Taught in the Curriculum and Conceptual Knowledge Use

Variable	Conceptual Categorization of Items Taught at Time 2	Conceptual Knowledge Use at Time 3
	β / γ^a (S.E.)	β / γ^a (S.E.)
Within-Teacher Variables		
Conceptual Knowledge Time 1		.37*** (.11)
Word Knowledge Time 2	.36*** (.09)	
Word Knowledge Time 1	.12 (.08)	.22* (.09)
Age in Months	.03** (.01)	
Attends PM Preschool	-.31** (.12)	
Between-Teacher Variables		
Treatment	.62** (.17)	.55* (.19)
Between-teacher variance (tau)	.02	.05
Within-teacher variance (sigma squared)	.45	.64

*p<.05; **p<.01; ***p<.001

Conceptual Knowledge Use Time 3. Another way children's conceptual knowledge was measured was to record children's use of concepts and properties in an open-ended task. Children were told "Tell me everything you know about *name of category*", their responses were coded for use of concepts taught in the curriculum, and the number of concepts used was summed. Children's spontaneous use of concepts and properties central to the topic and were taught in the curriculum would be an indication of deep, rich, and flexible conceptual knowledge.

Prior to treatment, scores on conceptual knowledge use were not significantly different between the treatment and comparison groups ($t = .98$, ns). Children in both groups were able to use, on average, one concept when talking about curricular topics. However, following the intervention there significant differences between the groups ($t = 5.14$, $p \leq .001$), with treatment children significantly outperforming children in the comparison group (see Table 15 for group means). After experiencing the curriculum, children in the treatment group used, on average 3 concepts when talking about curricular topic (while children in the comparison group used about 1 concept). These findings are an initial suggestion that, on average, children in the treatment group acquired more conceptual knowledge over the course of the study than children in the comparison group.

Table 15

Means and Standard Deviations of Conceptual Knowledge Use at Time 1 and Time 3

	Conceptual Knowledge Use Time 1 (total concepts used)	SD	Conceptual Knowledge Use Time 3 (total concepts used)	SD
Treatment Group	1.33	2.03	3.01***	3.01
Comparison Group	1.05	1.49	1.29	1.65

*p<.05; **p<.01; ***p<.001

To investigate this more thoroughly, an HLM model was fitted for this outcome. There was significant ($\chi = 47.38$, $p \leq .001$) and sufficient (ICC=20%) variability between teachers on conceptual knowledge use at Time 3 (see Table 13). Word knowledge at Time 1 ($\beta = .22$, $p \leq .05$) and conceptual knowledge use at Time 1 ($\beta = .37$, $p \leq .001$) were both significant predictors of conceptual knowledge use at Time 3 (see Table 14), explaining 21% of the variability between children. After accounting for important child-level variables, there was still a significant difference between children in treatment teachers' classrooms and children in comparison teachers' classrooms on conceptual knowledge use at Time 3, favoring treatment classrooms ($\gamma = .56$, $p \leq .05$); including treatment explained 58% of the variability and there was no longer significant variability between teachers after accounting for it ($\chi = 16.31$, ns). The intervention exerted a large effect on this outcome (Cohen's $d = .83$).

In sum, the analyses presented above clearly indicate that children in the treatment condition acquired the concepts taught in the WOW vocabulary intervention and were able to demonstrate that knowledge in both a categorization tasks and an open-ended task. Coupled with the clear increase in word knowledge and usage demonstrated in the previous section, it is clear that the WOW curriculum was effective in increasing word and conceptual knowledge in low-income preschoolers.

Research Question 3: Do children who experience the curriculum transfer learned concepts to new learning tasks? Children's ability to transfer conceptual knowledge to new learning tasks is a first indication that well-organized vocabulary and conceptual knowledge can foster future learning. To investigate if children who experienced the WOW curriculum were able to transfer conceptual knowledge to new learning tasks, HLM's were conducted using two outcomes: a) children's conceptual categorization of items that were *not* taught in the curriculum, and b) children's ability to justify those categorization choices with conceptually organized responses.

Conceptual Categorization of Items Not Taught Time 2. This measure assessed children's ability to transfer and apply the conceptual knowledge acquired in the curriculum to new learning tasks. Children were asked to make inferences about the category membership of items that were not taught in the curriculum. To successfully engage in this task, a child was required to transfer their conceptual knowledge to this new learning situation. Evidence of this near transfer is a powerful indication that the WOW curriculum has the potential to provide a foundation for further learning. As a reminder, immediately following instruction in each category, children were shown items that *were not* taught in the curriculum, were asked "Is it a *name of category*? Yes or no", and their correct answers were summed and divided by the total number of items to get a percent correct. An average percent correct across topics was then calculated.

Immediately following instruction, there were significant differences between the treatment and comparison groups on conceptually categorizing items that were not taught in the curriculum ($t = 5.51$, $p \leq .001$) (see Table 16 for group means). Even though none of the items had been taught in the curriculum, children who experienced the curriculum were able to correctly categorize, on average, about 75% of the items. Their peers in the comparison group, on average, were able to categorize significantly fewer items, with an average of 63% correct. These findings provide the initial suggestion that, not only did children who experienced the WOW curriculum learn the concepts that were taught in the curriculum, but they were able to transfer conceptual knowledge to make inferences when categorizing new items.

Table 16

Means and Standard Deviations for Conceptual Categorization of Items that Were Not Taught at Time 2

	Conceptual Categorization of Items Not Taught % Correct Across Topics	SD
Treatment Group	75***	11
Comparison Group	63	12

*p<.05; **p<.01; ***p<.001

To investigate this more rigorously, an HLM model was fitted using this outcome. The fully unconditional model demonstrated that there was significant (χ = 53.85, p $\leq$ .001) and sufficient variability between teachers (ICC = 26%) on children's ability to categorizing items that were not taught (see Table 17). Word knowledge of not taught items (β = .32, p $\leq$ .001), age in months (β = .03, p $\leq$ .01), and word knowledge at Time 1 (β = .18, p $\leq$ .05) significantly predicted performance on categorizing not taught items (see Table 18); together these variables explained 35% of the variation between children. After controlling for these important predictors, children in treatment teachers' classrooms significantly outperformed children in comparison teachers' classrooms on ability to categorize items not taught in the curriculum (γ = .55, p $\leq$ .05); treatment condition explained 59% of the variability between teachers. Treatment condition alone explained two-thirds of the variation between teachers in this outcome and did diminish the variability between teachers on conceptual categorization of new items, though there was still significant variability between teachers to be explained (from χ = 53.85, p $\leq$ .001 in the fully unconditional model to χ = 21.56, p $\leq$ .05 in the fully conditional model). Treatment condition had a large effect on this outcome (Cohen's *d* = .97).

Table 17

Fully Unconditional Hierarchical Linear Models (HLM): Psychometric Characteristics of the Conceptual Categorization and Justification of Conceptual Categorization Outcome Variables

	Conceptual Categorization of Items Not Taught Time 2 (n=139 children in 12 teachers' classes)	Justification of Conceptual Categorization of Items Not Taught Time 2 (n=140 children in 12 teachers' classes)
Intercept	-.12	-.11
Between-teacher variance (tau)	.26	.39
Within-teacher variance (sigma squared)	.77	.68
Intraclass correlation (ICC) (%)	26	36
Reliability	.77	.85

Note: The intraclass correlation (ICC) is the percentage of total variance in the outcome that lies systematically between teachers. It is calculated as follows: ICC (%) = (tau/tau + sigma squared) * 100.

Table 18

HLM Investigating the Effect of Treatment on Ability to Categorize and Conceptually Justify Categorization Choices of Items Not Taught in the Curriculum

Variable	Conceptual Categorization of Items Not Taught Time 2	Justification of Conceptual Categorization of Items Not Taught Time 2
	β / γ^a (S.E.)	β / γ^a (S.E.)
Within-Teacher Variables		
Word Knowledge of Not Taught Items	.32*** (.09)	.31*** (.09)
Word Knowledge Time 1	.18* (.09)	.12 (.09)
Age in Months	.03** (.01)	.03* (.01)
Days Absent		-.02* (.01)
Between-Teacher Variables		
Treatment	.55* (.19)	.74** (.19)
Between-teacher variance	.05	.06
Within-teacher variance	.50	.48

*p<.05; **p<.01; ***p<.001

Justification of Conceptual Categorization of Items Not Taught Time 2. This measure provides additional support for the investigation of children's ability to transfer conceptual knowledge to new learning tasks. Children who were able to verbally justify their categorization choices in an organized way demonstrated unequivocally that they were using their conceptual knowledge to make an inference about category membership. Children's justifications were coded for how taxonomically organized they were; using a modification of Langer's (1984) approach to assessing children's knowledge, children were given a 1 for an organized justification and a 0 for a diffusely organized justification and responses were summed then averaged across topics.

Immediately following instruction across topics, children in the treatment group were significantly better at providing organized justifications for their categorization choices ($t = 7.86$, $p \leq .001$) (see Table 19 for group means). Children in the treatment group were able to provide an organized justification an average more than three times out of ten (which is an average across topics). Children in the comparison group, in contrast, were only able to provide slightly more than 1 organized justification on average out of 10 categorizations.

Table 19

Means and Standard Deviations for Justification of Conceptual Categorizations at Time 2

	Justification of Conceptual Categorization Time 2 (mean across topics, high score of 10)	SD
Treatment Group	3.17***	1.87
Comparison Group	1.14	1.18

*p<.05; **p<.01; ***p<.001

To investigate further, an HLM model was fitted with children's justification score as the outcome. The fully unconditional model indicated that there was significant ($\chi = 77.28$, $p \leq .001$) and sufficient variability between teachers (ICC=36%) on children's ability to justifying their conceptual categorization choices (see Table 17). Word knowledge of not taught items ($\beta = .32$, $p \leq .001$) and days absent ($\beta = -.02$, $p \leq .05$) both significantly predicted performance on justifying conceptual categorization, while age in months ($\beta = .02$, $p \leq .10$) was a marginally significant predictor (see Table 18). Together these child characteristics explained 30% of the variability between children. After controlling for these significant predictors, children in treatment teachers' classrooms significantly outperformed children in comparison teachers' classrooms on ability to conceptually justify sorting choices ($\gamma = .74$, $p \leq .01$). Treatment condition explained 76% of the variability between teachers (though after accounting for treatment there was diminished, but still significant variation between teachers on this outcome ($\chi = 21.47$, $p \leq .05$)). Treatment condition had a large effect on this outcome (Cohen's *d* = 1.23).

Taken together, these analyses demonstrate that children in the treatment group significantly outperformed similar children in the comparison group on transferring and applying concepts to make inferences about items that were not taught in the curriculum. These findings suggest that when children learn words and concepts in taxonomic categories, they are able to transfer the knowledge they learn to learning situations outside of the curriculum.

Research Question 4: Do children who experience the WOW curriculum show more growth on general vocabulary measures than children in the comparison group?

It was important to determine if the curriculum influenced children's knowledge and use of the target words and concepts that were taught in the curriculum. However, it was also important to determine if the curriculum had an influence on more general vocabulary measures of receptive (Peabody Picture Vocabulary Test) and expressive (Get It, Got It, Go!) word knowledge.

General Receptive Word Knowledge. Prior to the treatment, scores on general receptive word knowledge were not significantly different in the treatment and control groups ($t = .01$, ns). This is unsurprising, given that the groups were matched on initial general receptive word knowledge. What is notable, however, is that a mean standard score of 87 in both groups on this measure indicates that this sample began the study with below average receptive vocabulary knowledge. An average standard score on this measure is 100, with a standard deviation of 15. However, children in this sample on average were almost one full standard deviation below the mean on general receptive vocabulary knowledge. Following the intervention, there were also no significant differences between groups ($t = -.23$, ns) (see Table 20 for group means). However, children in both groups increased about 3 points during this four month study.

Table 20

Means and Standard Deviations for General Receptive Word Knowledge at Time 1 and Time 3

	General Receptive Word Knowledge Time 1 (standard score)	SD	General Receptive Word Knowledge Time 3 (standard score)	SD
Treatment Group	87.29	13.23	90.27	14.95
Comparison Group	87.28	13.26	90.83	16.88

*p<.05; **p<.01; ***p<.001

To investigate this more rigorously, HLM models were fitted with general receptive vocabulary as the outcome. The fully unconditional model indicated that there was significant ($\chi = 26.50$; $p \leq .01$) and sufficient variability between teachers (ICC = 9%) on children's general receptive vocabulary (PPVT) (see Table 21). Performance on general receptive vocabulary at Time 1 ($\beta = .57$, $p \leq .001$) and word knowledge at Time 1 ($\beta = .21$, $p \leq .01$) both significantly predicted general receptive vocabulary at Time 3 (see Table 22), explaining 50% of the variability between children. There was no difference between children in treatment teachers' classrooms and children in comparison teachers' classrooms on general receptive vocabulary knowledge at the end of the study.

Table 21

Fully Unconditional Hierarchical Linear Models (HLM): Psychometric Characteristics of General Expressive and Receptive Word Knowledge Outcome Variables

	General Receptive Word Knowledge (n=171 children in 12 teachers' classes)	General Expressive Word Knowledge (n=140 children in 12 teachers' classes)
Intercept	-.00	.02
Between-teacher variance (tau)	.09	.10
Within-teacher variance (sigma squared)	.91	.91
Intraclass correlation (ICC) (%)	9	10
Reliability	.57	.57

Note: The intraclass correlation (ICC) is the percentage of total variance in the outcome that lies systematically between teachers. It is calculated as follows: ICC (%) = (tau/tau + sigma squared) * 100.

Table 22

HLM Investigating the Effect of Treatment on General Receptive and Expressive Word Knowledge at Posttest

Variable	General Receptive Word Knowledge	General Expressive Word Knowledge
	β / γ^a (S.E.)	β / γ (S.E.)
Between-Teacher Variables		
General Receptive/Expressive Word Knowledge Time 1	.57*** (.07)	.25** (.08)
Word Knowledge Time 1	.21** (.07)	.35*** (.08)
Within-Teacher Variables		
Treatment	-.11 (.15)	.28 (.18)
Between-teacher variance	.03	.04
Between-child variance	.46	.67

*p<.05; **p<.01; ***p<.001

General Expressive Word Knowledge. Prior to treatment, scores on general expressive vocabulary were not significantly different ($t = .00$, ns). On average, children in both groups correctly named about 16 items on this timed measure. After the intervention, there were significant mean differences between the treatment and comparison groups ($t = 2.60$, $p \leq .01$) (see Table 23 for group means). Children in the treatment group were able to correctly name, an average of about 19 items while children in the comparison group named an average of about 16.6 items. This *initial* analysis indicates that children in the treatment group may have made more growth on general expressive vocabulary knowledge over the course of the study than children in the control group.

Table 23

Means and Standard Deviations on General Expressive Word Knowledge at Time 1 and Time 3

	General Expressive Word Knowledge Time 1 (total)	SD	General Expressive Word Knowledge Time 3 (total)	SD
Treatment Group	16.28	6.74	19.16**	5.87
Comparison Group	16.28	7.04	16.59	6.66

*p<.05; **p<.01; ***p<.001

To investigate these mean differences more fully, an HLM model was fitted. The fully unconditional model indicated that there was significant ($p \leq .01$) and sufficient variability between teachers (10%) on children's performance on general expressive word knowledge (GGG Rapid Naming) (see Table 21). Word knowledge at Time 1 ($\beta = .35$, $p \leq .001$) and general expressive word knowledge at Time 1 ($\beta = .25$, $p \leq .01$) together explained 27% of the variability between children and were both significant predictors of general expressive word knowledge at Time 3 (see Table 22). After controlling for these predictors, there were no differences between children in treatment teachers' classrooms and children in comparison teachers' classrooms on general expressive word knowledge at Time 3. This is in contrast to the results from comparing means between groups, which indicated that that there may have been more growth in the treatment group than the comparison group on this general expressive measure; it seems that parsing the variance between children and classrooms and/or accounting for children's initial performance explained accounted for these differences.

Children who experienced the WOW curriculum did not demonstrate significantly more growth than their comparison group peers on general expressive or receptive word knowledge. Perhaps only four months of vocabulary and conceptual knowledge instruction was insufficient to influence general word knowledge, or it may be that the general word knowledge measures are not sensitive enough to capture more fine grained increases in word knowledge.

Summary of Quantitative Analyses

In sum, descriptive and multivariate analyses indicated that children who experienced the WOW curriculum acquired the target words and concepts, outperforming children in the comparison group (see Table 24 for a chart summarizing the influence of treatment on all outcomes). In addition, children who experienced the WOW curriculum were able to apply conceptual knowledge acquired in the curriculum beyond the curricular lessons to make inferences and justify those inferences when presented with new learning situations. However, there was no evidence that this four month intervention increased children's growth on general word knowledge measures. These results provide ample support for the efficacy of the WOW curriculum in increasing target vocabulary knowledge, target conceptual knowledge, and ability to transfer conceptual knowledge beyond the confines of the curriculum.

Table 24

Summary of Treatment Effect on All Outcomes

	Significant Treatment Effect
Research Question 1: Do children who experience the curriculum learn the words that were taught?	
Word Knowledge Time 2 Curriculum-specific word knowledge immediately following instruction	+[a]
Word Knowledge Time 3 Curriculum-specific word knowledge at the end of the study	+
More Difficult Word Use Time 3 Use of more difficult words at the end of the study	+
Research Question 2: Do children who experience the curriculum learn the concepts that were taught?	
Conceptual Knowledge Use Time 3 Use of curriculum-specific concepts at the end of the study	+
Conceptual Categorization of Items Taught Time 2 Ability to categorize items that were taught in the curriculum immediately following instruction	+
Research Question 3: Do children who experience the curriculum transfer learned concepts to new learning tasks?	
Conceptual Categorization of Items ***Not Taught*** *Time 2* Ability to categorize items that were not taught in the curriculum immediately following instruction	+
Justification of Conceptual Categorization-Items Not Taught Ability to verbally justify categorization of items that were not taught in the curriculum immediately following instruction	+
Research Question 4: Do children who experience the curriculum show more growth in general receptive and expressive vocabulary?	
General Receptive Vocabulary (PPVT) General receptive vocabulary knowledge	-
General Expressive Vocabulary (GGG) General expressive vocabulary knowledge	-

a. A + sign means that treatment had a positive and significant effect on the outcome.
b. A – sign means that the treatment had no significant effect on the outcome.

Anecdotal Examples of Children's Language

A variety of different assessments were administered throughout the course of this study. Some of these assessments were more receptive in nature and only required the child to make a choice (e.g. conceptual categorization), and others only required children to provide a one-word response (expressive word knowledge). However, there were two important assessments that prompted children to provide more complex verbal responses: conceptual knowledge use and justification of conceptual categorizations. This section is designed to illustrate children's actual verbal responses during these assessments.

During these two assessments, children's responses were always audiorecorded and transcribed. To review, to prepare these data for quantitative, multivariate analyses, children's verbal responses were coded, assigned numeric values, and summed or averaged to create a quantitative outcome. Though that process was imperative, it is also illustrative to look more closely at children's actual verbal responses. This section includes examples of children's verbal responses on two assessments: a) conceptual knowledge use and b) justification of conceptual categorizations.

Samples of Children's Language on the Conceptual Knowledge Use Task. HLM analyses were conducted using children's conceptual knowledge use as the outcome. Those analyses indicated that that children in the treatment classrooms significantly outperformed children in the comparison classrooms on conceptual knowledge use at Time 3.

The quantitative HLM analysis of children's conceptual knowledge use was very important and empirically rigorous; however, it is important to remember that scores for conceptual knowledge use were created by assigning numeric values to children's actual verbal responses[23]. This section provides a richer tableau that illustrates the actual distribution of these responses and examples of the content and quality of different responses.

First, to contextualize examples of different types of responses, it is helpful to demonstrate the unequal distribution of scores across the treatment and comparison groups (although the HLM analyses demonstrated this, the data are presented in a slightly different way to foster understanding of different types of responses)[24]. Table 25 shows that a much larger proportion of children in the treatment group than the comparison group were able to use three or more concepts when asked to talk about the categories at Time 3 (56% and 18%, respectively). Further, a much *smaller* proportion of children in the treatment group than the comparison group were unable to use any concepts (22% and 44%, respectively).

Table 25

Children's Use of Curricular Concepts at Time 3

Score	% of the Treatment Group	% of the Comparison Group
Low (Score = 0)	22	44
Medium (Score = 1 or 2)	23	37
High (Score = 3 or above)	56	18

[23] As described in Chapter 3, for each topic children were shown a card depicting several members of the category and told "Tell me everything you know about *name of category*". Children's verbal responses were coded for use of curricular concepts when talking about the category. Children were assigned one point for each concept included and points were summed for a total conceptual knowledge use score. Though there was no ceiling to the total amount of conceptual knowledge a child could potentially use during this task, the actual range was from 0 to 10 concepts used.

[24] To illustrate this, scores were grouped into high (3 or more concepts used), medium (1 or 2 concepts used), and low (0 concepts used) scores.

Children who had a low score were unable to use any concepts about topics (total score = 0) were much more prevalent in the comparison group. Children who were unable to use any concepts on this task often simply labeled the items on the card or gave information specific to one of the category members. However, they were unable to articulate information that summarized the concept as a whole. For example, when asked what they know about the concept Pets, one child who scored 0 on this measure gave the following response ("A" means Assessor and "C" means child):

Pets

A: All of these are pets. Tell me everything you know about pets.

C: He barks. She says meow. He hops. He swims.

The same child, when asked about what she knew about Parts of the Body, said:

Parts of the Body

A: All of these are parts of the body. Tell me everything you know about parts of the body.

C: This is a muscle. Hand, leg, foot, head, ear, muscle.

A: How do you know that all of these are parts of the body?

C: Because.

As children's scores increased, their responses to some of the topics became more conceptually dense. They began to look beyond listing the individual items to make comments that summarized the concept as a whole, rather than commenting on individual items. Some children were only been able to state concepts for one or two specific categories, resulting in a total score of 1 or 2. For example, one child who had an overall score of 2 was able to use concepts to talk about the taxonomic category of Pets, but Pets only. The following is an excerpt from his transcript:

Pets

A: All of these are pets. Tell me everything you know about pets.

C: Feed them, put their water for them, pet the bunny, pet the toad.

The child articulated that pets need to be fed, given water, and given affection in the form of petting-all concepts that apply to Pets as a category. This child was able to use concepts that applied to the category as a whole, rather than simply listing types of Pets.

Children who were able to use even more concepts when talking about the categories demonstrated even more conceptual knowledge of the categories. Children who were able to use three or more concepts when talking about the categories, which was much more common among children in the treatment group (56% of children in the treatment group used three or more concepts when talking about the categories), were often able to use overarching concepts for several categories. For example, one child who scored a three gave the following responses when asked separately about Pets, Parts of the Body, and Animals in Water:

Pets

A: All of these are pets. Tell me everything you know about pets.

C: These take care of them. We feed them. We give them water. We let them outside.

Parts of the Body

A: All of these are parts of the body. Tell me everything you know about parts of the body.

C: You can't take them off.

Animals that Live in Water

A: All of these are animals that live in the water. Tell me everything you know about animals that live in the water.

C: They sharks. Dolphins, octopuses, seahorses, stingrays.

A: How do you know they live in the water?

C: Because they can't come out of the water. Can't come out.

This child understood and was able to verbalize his understanding critical information about the concepts of Pets (that we need to take care of them by feeding them and giving them water), Parts of the Body (that they are attached and do not come off of our body), and Animals that Live in Water (that they must stay in the water to survive). This child demonstrated that he knew and was able to use conceptual knowledge about these categories as a whole to talk about the concepts.

Children who were able to use many concepts to talk about the topic, and thus had a total conceptual knowledge use score at the higher end of the scale, were able to demonstrate conceptual knowledge about all (or almost all) topics. Children who scored highly on this measure were almost exclusively in the treatment group. One child in the treatment group, who scored an eight on the posttest, gave the responses listed below. It should be noted that this particular child was unable to use any concepts to talk about the categories at Time 1:

Pets

A: All of these are pets. Tell me everything you know about pets.

C: They play, eat, they get stuff.

A: How do you know that all of these are pets?

C: Because you get them from the pet store.

Parts of the Body

A: All of these are parts of the body. Tell me everything you know about parts of the body.

C: We grab, we walk, we run, jump. Put your pinky up.

A: How do you know that all of these are parts of the body?

C: Cause to help you. They're attached.

Animals in Water

A: All of these are animals that live in the water. Tell me everything you know about animals that live in the water.

C: Those swim, they eat, they (inaudible), they swim around.

A: How do you know that all of these are animals that live in the water?

C: Cause when they come out they go back under the water.

Clothes

A: All of these are clothes. Tell me everything you know about clothes.

C: You put them on to protect us.

A: How do you know that all of these are clothes?

C: Because we put them on.

Insects

A: All of these are insects. Tell me everything you know about insects.

C: They eat, fly around, walk around, eat, drink.

A: How do you know that all of these are insects?

C: Cause they fly and walk.

Wild Animals

A: All of these are wild animals. Tell me everything you know about wild animals.

C: They eat, they drink.

A: How do you know that all of these are animals?

C: Because they live in the grassy field.

Given that this child had only been able to use one relevant concept at Time 1, it seems that not only did she acquired a great deal of conceptual knowledge during the four months of the study, but she was flexibly able to use that knowledge to talk about the topic.

The examples of children's language provided in this section augment the HLM analyses and illustrate the qualitative differences in way children in the treatment classrooms were able to use their newly acquired conceptual knowledge to talk about the topic. This ability demonstrates a flexible, and potentially deep conceptual knowledge about the taxonomic categories included in the curriculum.

Examples of Children's Language When Justifying Conceptual Categorization Choices. HLM analyses were conducted to quantitatively investigate mean differences in children's ability to verbally justify their categorization choices when presented with items that were not taught in the curriculum. Those analyses indicated that children in treatment classrooms gave significantly more organized justification for their categorization choices than children in the comparison group.

The quantitative HLM analysis of this outcome was critical. However, it is important to consider that this outcome that was created by assigning numeric values to children's verbal responses[25]. This section provides a closer look at the distribution and quality of children's actual verbal responses. Before providing examples of Organized and Diffusely Organized justifications, it is important to demonstrate that children in the treatment group made far more Organized responses than children in the control group (though HLM analyses demonstrated this, proportions are an additional way to display the data). Table 26 demonstrates that, on average across items, children in the treatment condition were much more likely than children in the comparison group to make Organized justifications of their sorting choices (33% compared to 11%, respectively). Further, children in the treatment group were much *less* likely than children in the comparison group to make Diffusely Organized justifications (67% compared to 89%, respectively).

Table 26

Proportion of Children Who Provided Organized and Diffusely Organized Justifications

	Diffusely Organized	Organized
% in the Treatment Group	67	33[a]
% in the Comparison Group	89	11

a. This number should be interpreted as the average proportion of children in the treatment condition (across all items) who gave organized justifications of their categorization choice. In contrast, on average, 11% of children in the comparison condition gave organized justifications. All other numbers should be interpreted similarly.

Organized responses, which were much more common in treatment classrooms than comparison classrooms (33% compared to 11%, respectively), included those

[25] As described in Chapter 3, children's verbal justifications of their sorting choice on each individual item was coded as Organized (1 point) or Diffusely Organized (0 points).

responses that indicated a solid conceptual knowledge of the category. This was demonstrated by correctly categorizing the item and then providing an organized verbal justification for that choice that included properties, attributes, or defining characteristics that the item had (or did not have) in common with the superordinate category. For children who experienced the curriculum, Organized responses also illustrated an ability to transfer and apply conceptual knowledge from the curriculum when faced with a new task.

An example of an Organized response was as follows:

A: Is a beetle an insect?

C: Yes.

A: How did you know a beetle is an insect?

C: 1, 2, 3, 4, 5, 6, cause it got 6 legs.

In another example, when asked if a snake was an insect, a child correctly said no. The following is the remainder of the exchange:

A: How did you know snake is not an insect?

C: Because he don't have body parts, because he don't have antennas or wings or not 6 legs.

These responses indicate an ability to apply conceptual knowledge about the category as a whole when considering the item. Though some children in the comparison condition were able to do this, HLM analyses indicated that children in the treatment group were significantly more likely to provide Organized responses like those illustrated above (see Table 27 for more examples of Organized responses).

Justifications coded as Organized also included another type of response. These were instances where a child categorized an item *incorrectly*, but then gave a justification for that sorting choice that indicated an organized and accurate knowledge of the concept. This type of response was relatively uncommon in both the treatment and comparison conditions (3% of treatment group justifications and 1% of comparison group responses). The following is an example of this type of response:

A: Is playing basketball exercise?

C: No.

A: How did you know a playing basketball is not exercise?

C: Because it doesn't make your heart beat.

Children were taught in the WOW curriculum that exercise strengthens your heart by making it beat faster. Though this child was incorrect about the specific attributes of playing basketball, he did justify his categorization choice using organized knowledge about the concept of exercise (see Table 27 for more examples of this type of response).

Children who provided justifications that were coded as Diffusely Organized, which were more common in comparison classrooms, were unable to give an organized justification for their sorting choice that included any conceptual properties, attributes, or defining characteristics of the category. In short, Diffusely Organized justifications lacked information that linked the item to the larger category. In addition, these responses were often characterized by tangential links or information about the specific item. An example of a Diffusely Organized response of this type is as follows:

A: Is an alligator a wild animal?

C: Yes.

A: How did you know an alligator is a wild animal?

C: Because they can bite, they got sharp teeth.

Another example was:

A: Is a crab an insect?

C: No.

A: How did you know a crab is not an insect?

C: Cause it something that go at the beach.

Other children provided Diffusely Organized justifications that referred to a personal experience they may have had with the item, but lacked any reference to the concepts that underlie the category. For example, when asked how the child knew an item was a member of the category, children responded in the following ways: “Cause I saw it at the zoo”, “Cause I learned it at school”, “Cause I see it at school”, or “My mom told me”. Still other responses included absolutely no information, such as no response or a simple “I don’t know”. See Table 27 for more examples of Diffusely Organized justifications.

Table 27

Examples of Organized and Diffusely Organized Responses

	Organized	Diffusely Organized
WILD ANIMALS		
Black Bear	Is a wild animal ***"cause he lives out in the woods in the wild"***	Is a wild animal because ***"grrr!"***
Bird (pictured in a cage)	Is not a wild animal ***"because he lives with people, wild animals don't live with people"***	Is not a wild animal ***"cause it's not"***
INSECTS		
Mosquito	Is an insect ***"cause it got antennas and wings and 3 body parts, 1, 2, 3.. and 6 legs"***	Is an insect ***"cause I just told myself"***
Mouse	Is not an insect because ***"it only has one body segment"***	Is not an insect ***"cause they live in your house sometime"***
ANIMALS IN WATER		
Blowfish	Is an animal that lives in water ***"because if a puffer fish, because that kind of fish can get dead if it's on land"***	Is an animal that lives in water ***"because they will poke you"***
Tiger	Is not an animal that lives in water ***"because he lives in the grassland"***	Is not an animal that lives in water ***"because it's so mean, so are sharks"***
PARTS OF THE BODY		
Ankle	Is a part of the body ***"because it's attached to our bodies"***	Is part of the body ***"cause my brain told me"***
Glasses	Are not a part of the body ***"because you can take them off"***	Are not part of the body ***"because my mom got some new ones"***
CLOTHES		
Overalls	Are clothes ***"cause you can put them on"***	Are clothes ***"cause you like them. My mom told us the flower ones are ugly, and they're not"***
Pillow	Is not clothes ***"because you can't put it on. You can just lay on it"***	Is not clothes ***"because you lay on it like Picky Peter laying on me"***
EXERCISE		
Ice Skating	Is exercise ***"because it helps your bones (get) strong"***	Is exercise ***"because my sister ice skates, my sister big, all way up to clouds, my sister is big like a tiger"***
Driving	Is not exercise ***"because it's not making muscles"***	Is not exercise ***"cause I know stuff like that"***

The ability of children in the treatment group to provide qualitatively better verbal justifications to explain their categorization choices is evidence providing further support for the HLM results; children in the treatment group not only acquired the concepts taught in the curriculum, but were able to transfer, apply, and demonstrate that conceptual knowledge in an organized way when engaged in a new learning task.

The examples of children's verbal responses presented in this section illustrate the stark differences between the treatment and comparison groups in the quality of and conceptual content included in children's language. These examples augment the quantitative findings around conceptual knowledge and demonstrate the deep and flexible conceptual knowledge acquired by children who experienced the WOW curriculum.

Conclusion

In sum, children who experienced the WOW curriculum acquired the target words and concepts, outperforming children in the comparison group. In addition, children who experienced the WOW curriculum were able to apply conceptual knowledge acquired in the curriculum beyond the curricular lessons to make inferences and justify those inferences when presented with new learning situations. Further, their verbal justifications were qualitatively different than children in the comparison group. The ability of children in the treatment group to outperform the comparison group on inferring and making organized and sophisticated conceptual justifications in these new learning situations provides support for the potential of the approach taken in the WOW vocabulary curriculum. These findings suggest that teaching words and concepts in taxonomic categories has the potential to foster future learning. However, despite the impressive vocabulary and conceptual knowledge accrued by children who experienced the WOW curriculum, experiencing this type of instruction for four months was not sufficient to influence general receptive and expressive word knowledge. The next chapter expounds on each of these findings more fully and offers directions for further research.

Chapter 5

Discussion

The WOW vocabulary curriculum was based on the theory that teaching words and concepts in taxonomic categories is a robust and viable means of enhancing the vocabulary and conceptual knowledge of preschoolers. Further, it was based on the hypothesis that learning words and conceptual knowledge in this well-organized fashion would be transferrable and beneficial in new learning situations, with potential to foster later learning. This dissertation study was designed as an initial investigation of these hypotheses with a sample of low-income preschoolers, comparing the performance of children who experienced a trial version of the WOW curriculum to a group of children matched on general receptive and expressive vocabulary.

Low-Income Sample

The children that comprised the sample used in this study were growing up in families living in poverty. Research suggests that the influences of poverty are wide-ranging, including less exposure to large quantities and high-quality language interactions (Hart & Risley, 1995; Weizman & Snow, 2001). Because we know that the quantity and quality of the language input experienced by children is a major determinant in vocabulary size (Pan et al., 2005; Weizman & Snow, 2001) and growth trajectory (Hart & Risley, 1995), perhaps it is unsurprising that, in this study, the general receptive vocabulary of children in both the treatment and comparison groups was far below average. The mean standard score of this sample on the widely-used and well-respected

Peabody Picture Vocabulary Test (PPVT) was almost one standard deviation below the mean of the norming sample for the measure. Because the WOW intervention was designed to increase the vocabulary and conceptual knowledge of low-income preschoolers in an effort to narrow vocabulary gaps before children enter kindergarten, the sample used in this study is highly appropriate for this initial test of the efficacy of the curriculum. However, because the literature indicates that children with more vocabulary knowledge more easily acquire new vocabulary knowledge (Nagy & Scott, 2000), it is important to highlight the low general vocabulary knowledge of this sample when considering the results of the study. The effects of the WOW curriculum on vocabulary, conceptual knowledge, and transfer of conceptual knowledge on this low-income, low-vocabulary sample were impressive and achieved the stated goals of the intervention.

Primary Findings

Curriculum-Specific Word Knowledge. There were several key findings supported by the results from this study. First, results demonstrated that children who experienced the WOW curriculum learned the words that were taught. There were statistically significant differences between the treatment and comparison groups on word knowledge immediately following instruction. In the WOW curriculum, children were taught an average of three words per day. The finding that children in this study did seem to demonstrate immediate learning of words that were taught at this rate supports Biemiller's (2004) contention that learning three words per day may be an appropriate expectation for word learning in preschool.

There were also statistically significant differences in word knowledge between the treatment and comparison groups at the end of the study. However, it is important to note that standardized coefficients from HLM analyses and effect sizes[26] indicated that there were larger differences between treatment and comparison groups immediately following instruction than there were at the end of the study. This indicates that children who experienced the curriculum demonstrated immediate recall of the words they were taught, but that all of this learning was not retained after some time had elapsed. Though children in the treatment group were exposed to each word an average of 22 times and asked to produce each word at least 3 times, the majority of these exposures were during the 8-day lesson sequence for the corresponding taxonomic category. We know that multiple exposures to vocabulary are necessary for word learning (Blachowicz & Fisher, 2000; Pressley, Disney, & Anderson, 2007) and the WOW curriculum seemed to achieve immediate word learning, perhaps partially due to the number of exposures in each 8-day lesson sequence. However, we also know that adequate review is necessary for continued word knowledge and learning (Biemiller & Boote, 2006). Perhaps the WOW curriculum did not include enough review of words beyond the eight-day lesson sequence to maintain the all word learning until the end of the study. Future versions of the WOW curriculum may need to build in more review throughout the curriculum in order to maintain the level of word knowledge demonstrated immediately following instruction.

Another important finding was that children who experienced the WOW intervention were significantly more likely than children in the comparison group to use the more difficult words that were taught in the curriculum in an open-ended task (e.g. “Tell me everything you know about Insects”). There was a large effect of the curriculum

[26] Standardized HLM coefficients indicated that the treatment group outperformed comparison group by .86 standard deviation units immediately following instruction, compared to .34 standard deviation units at the end of the study. The Cohen’s *d* effect size of the intervention was 1.54 immediately following instruction, but diminished to .63 at the end of the study.

on use of more difficult words (Cohen's $d = .84$), and HLM analyses indicated children in the treatment group used significantly more difficult words at the end of the study than children in the comparison group[27]. The ability of children in the treatment group to begin to produce and use the more difficult words that were taught is a promising finding; using a word in a natural and spontaneous way in a sentence is likely an indication of deep knowledge (Nagy & Scott, 2000) and a harbinger of the ability to transfer word knowledge beyond the confines of the curriculum.

Curriculum-Specific Conceptual Knowledge. It is important that children "learned" the words that were taught in the curriculum. However, because we know that vocabulary words are essentially mapped onto the mental representations called concepts (Bloom, 2000; Murphy, 2002; Waxman, 2004), it is important to investigate beyond children's surface ability to name an item to their deeper conceptual knowledge. One way that this study went beyond the majority of studies of children's vocabulary knowledge is that in addition to assessing vocabulary knowledge in a relatively simple way, one goal of the study was also to thoroughly assess children's understanding of the concepts to which that vocabulary is attached.

An important finding was that in addition to acquiring knowledge of words that were taught in the curriculum, children who experienced the WOW curriculum also accrued knowledge of the concepts taught. Children who experienced the curriculum demonstrated their conceptual knowledge by significantly outperforming children in the comparison group in two ways. First, children in the treatment group were significantly

[27] It is important to note that on an open-ended task like "Tell me everything you know about Wild Animals", just because a child *did not* spontaneously use more difficult words does not necessarily mean the child does not know the word. For this reason, it is likely that this measure underestimated both treatment and comparison children's knowledge of these words.

more adept at categorizing items that were taught in the curriculum, indicating an understanding of the taxonomic categories, and thus the overall concepts, that were taught in the curriculum. Second, when asked to spontaneously talk about what they knew about the taxonomic categories, children in the treatment group were significantly more likely than children in the comparison group to talk at a higher level of abstraction about the properties and overarching concepts related to the taxonomic categories rather than focusing on individual exemplars of the category. In some ways, these findings support Spycher's (2009) finding that children who were taught and learned vocabulary related to insects (albeit not necessarily members of the taxonomic category of insects) were better able to express their understanding of target concepts.

Both of these findings suggest that children who experienced the WOW curriculum had a rich conceptual structure for the taxonomic categories that were taught. They were able to think about and talk about the taxonomic categories at a higher level of organization and abstraction than children who didn't experience the curriculum; they could categorize items and talk about the category in ways that demonstrated rich understanding of the entire concept and how individual kinds are related (or not related to) this concept. This is an initial indication that children appropriated the deep and well-organized taxonomic conceptual structure and the vocabulary associated with those concepts. The inherent inductive potential of this type of representation (Gelman et al., 1998) may help children extend knowledge beyond the known to foster future learning.

Children's Success Using Curricular Vocabulary and Concepts. In both measures of vocabulary knowledge and measures of conceptual knowledge, children were presented with two kinds of tasks. Some assessments represented closed tasks, such as the expressive vocabulary and conceptual categorization task, which were more passive in nature. In these types of tasks, children were presented with stimuli and asked to provide a one word answer, which is an indication of a certain level of understanding. Other tasks were more open-ended, such as when children were asked to talk about the taxonomic categories taught in the curriculum and their responses were coded for use of both concepts and more difficult vocabulary taught in the curriculum. Not only were children who experienced the curriculum more successful than their peers in the comparison group at demonstrating their knowledge of words and concepts taught in the curriculum on closed tasks like expressive vocabulary assessments and conceptual categorization, but they were also able to *use* the words and concepts acquired in the curriculum when asked to talk about the taxonomic categories in open-ended tasks. Success using vocabulary and conceptual knowledge in more open-ended tasks is evidence of a deep conceptual knowledge that can be applied and used flexibly. This deep conceptual knowledge will likely foster future learning, because we know that taxonomic representation fosters inductive inferences (Murphy & Lassaline, 1997), has great inductive potential (Gelman, 1988; Gelman & Markman, 1986, 1987), and because well-specified and coherent knowledge of categories foster word learning (Borovsky & Elman, 2006).

Transfer of Conceptual Knowledge. One of the most critical findings in this dissertation study was the evidence that children who experienced the WOW curriculum were able to transfer and apply conceptual knowledge from the curriculum to new learning tasks. When presented with items that were not taught to any children in either group, children in the treatment group were significantly better at making inferences about category membership of these new items. This is initial evidence of near-transfer. Just based on treatment children's superior ability to categorize items that weren't taught in the curriculum, it could be *inferred* that children in the treatment group were transferring and applying the concepts they learned about the taxonomic categories to make better inferences. However, without delving more deeply, this closed, receptive task (e.g. "Is it a wild animal? Yes or no") could only provide speculative and suggestive evidence of near-transfer and application of conceptual knowledge to new learning tasks.

To investigate near transfer and application of conceptual knowledge more fully, children were asked to verbally justify their inferences. Children in the treatment group were significantly better than children in the comparison group at providing verbal justifications of their categorization choices that were well-organized and referred to a higher level of abstraction (e.g. named a property common to all category members). These findings corroborated the hypothesis that children in the treatment group were, in fact, drawing on their conceptual knowledge and transferring and applying it when faced with the task of categorizing new items. Coupling children's categorization choices with their verbal justifications of these choices provides robust evidence that children who experienced the WOW curriculum were in fact transferring and applying the conceptual knowledge they acquired in the curriculum to make inferences and aide them when encountering a new learning task. This finding is provocative-it appears that not only have children acquired the vocabulary and concepts in a taxonomic structure, but they are able to apply it outside of the curriculum. This is an initial indication that this type of instruction leads to conceptual knowledge that is organized in ways that promote induction and inference such that it serves as a foundation for new learning (Bloom, 2000; Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Gelman & Medin, 1993; Murphy & Lassaline, 1997). The WOW curriculum was designed to provide children with conceptual knowledge and associated vocabulary in a self-extending, taxonomic structure that will provide a foundation for future learning as they move into formal schooling. These results provide initial evidence that learning words and concepts in taxonomic categories can provide a foundation that will foster future learning.

General Vocabulary Knowledge. There was no evidence that experiencing this 16 week intervention influenced general receptive or expressive word knowledge. Ideally, a vocabulary intervention would have such a large influence that children's general vocabulary knowledge would increase. It is possible that four months of instruction was simply not enough time to engender change in general vocabulary knowledge. However, others have suggested that standardized measures of vocabulary knowledge (such as the Peabody Picture Vocabulary Test used to measure general word knowledge in this study) do not seem to be "sufficiently sensitive to vocabulary changes" to accurately capture vocabulary growth (National Reading Panel, 2000). In their meta-analysis of 50 studies of vocabulary instruction, the National Reading Panel found that "there may be a need for the development of standardized measures that are much more sensitive to the nuances and complexities involved in vocabulary acquisition" (National Reading Panel, 2000).

Limitations

This study produced robust findings regarding the effect of the WOW curriculum on curriculum-specific word and conceptual knowledge and children's ability to transfer and apply their knowledge to new learning tasks. However, there were several important limitations to this study involving the design, measurement schedule, and applicability of taxonomic categories.

Design. The research design of this study compared a very specific type of instruction, teaching words and concepts in taxonomic categories, to “business as usual”. Given the structure of this research design, in many ways it is unsurprising that there were large differences between children in the treatment classrooms and children in the comparison classrooms-children learn what they are taught. Further research is necessary to determine if teaching children words and concepts in taxonomic categories is superior to teaching the same words and concepts in a different way (e.g. thematically or in unrelated groups).

A second limitation to the design of this study was that it only evaluated the efficacy of the curriculum as a whole, rather than investigating if certain aspects of the WOW curriculum were more effective than others. For example, the study was not designed to allow an examination of the contribution and power of the Sesame Street video clips versus the informational texts, or comparison of either to the utility of the picture cards or teacher-child language interactions. Further research is necessary in order to fine tune this initial version of the WOW curriculum such that the most critical elements of the instructional sequence are isolated and highlighted.

Measurement Schedule. One limitation regarding measurement schedule was that children’s transfer of their conceptual knowledge to new learning tasks was only measured immediately following instruction. Findings at that time were robust and indicated that children in the treatment condition significantly outperformed their comparison group peers on this task, an initial indication that this type of instruction would foster future learning. However, a delayed posttest measure was not administered to determine if this ability was retained over time. It is important that further research investigate if children who experience the WOW curriculum retain their conceptual knowledge over time *and* retain their ability to transfer this conceptual knowledge to new tasks like learning new content or comprehending text.

Limitations to Taxonomic Categories. There is a great deal of evidence that organizing information in the world into taxonomic categories is natural to humans (Bloom, 2000; Murphy & Lassaline, 1997), offers inductive potential that can foster future learning (Gelman, 1988; Gelman & Markman, 1986, 1987), and is highly related to and beneficial for word learning (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1986, 1987, 1992). However, there are many important concepts (and words that are associated with those concepts) that exist in the world and in our minds that *are not* organized taxonomically. These include, but are not limited to, events or states, individuals, and abstract ideas, and thematic categories (Gelman & Kalish, 2006). It may be that teaching vocabulary in taxonomic categories is highly effective and relevant for increasing children's conceptual knowledge about specific content areas, like science. For example, though the type of instruction presented in the WOW curriculum is highly appropriate and effective for teaching the concepts "insect" or "reptile". However, it may be less than appropriate for teaching children the concept "hibernation", "post office", or "friendship". For this reason, the applicability of this type of instruction may likely be limited to certain types of concepts. However, the findings in this dissertation study suggest that it is a highly effective means of increasing conceptual knowledge of concepts related to the scientific categories "Living Things" and "Health".

Contributions to Literature Base

This study adds to the current literature on vocabulary interventions for preliterate children in two important ways. First, the majority of vocabulary studies for children not yet reading conventionally use storybook reading as the main template for vocabulary instruction. This takes advantage of one way that we know that pre-readers acquire vocabulary-through shared book reading (National Early Literacy Panel, 2009). Though the WOW vocabulary intervention uses shared book reading as one element of instruction (along with Sesame Street video clips and picture cards), the primary source of instruction occurs within teacher-child language interaction. Embedding vocabulary instruction in this important context capitalizes on the other way that young children learn language, which is through oral language input and language interaction with adults (Beck & McKeown, 2007). By including several different instructional components, the WOW curriculum exposes children to new vocabulary in both decontextualized settings (in informational texts) and contextualized settings (in extensive teacher-child language interactions) settings, thus providing them with several means by which to learn vocabulary.

The second way that this study contributes to the extant research base on early vocabulary instruction is that very few early vocabulary interventions have endeavored to teach children words in semantic categories. The majority of existing vocabulary studies focus instruction around vocabulary words that are a) present in a particular book and b) are determined to be at a particular level of difficulty and/or particularly useful to know (Beck & McKeown, 2007; Biemiller & Boote, 2006). Though it is often the case that the target vocabulary taught in these interventions are useful and important words for children to know, only a very few studies in the existing literature base chose to teach children vocabulary in conceptually related groups (Wasik et al., 2006; Spycher, 2009). Though these few studies did teach words in conceptual groups (Wasik et al., 2006; Spycher, 2009), the WOW curriculum represents the first study (to my knowledge) to adopt an approach that organizes vocabulary instruction around taxonomic categorization. This approach not only takes advantage of the human capacity to represent and organize information taxonomically (Murphy & Lassaline, 1997), but presents information in a structure that has inductive potential (Gelman et al., 1998) and can therefore likely facilitate and ease future learning and comprehension.

Areas for Future Research

There are several areas of research necessary to further investigate the findings of this dissertation study. First, though the current dissertation study does indicate that learning words and concepts in taxonomic categories does result in increases in vocabulary, conceptual knowledge, and transfer of conceptual knowledge, it does not provide evidence that this type of instruction is superior to *other* methods of instruction. Subsequent experimental tests of this curriculum should be designed in a way that allows comparison of teaching words and concepts in taxonomic categories to some other approach (or approaches) to teaching vocabulary and conceptual knowledge (e.g. thematically, through read alouds).

Second, all vocabulary and conceptual knowledge measures were collected either immediately following instruction or at the end of the study. Though very informative, these measures only provide insights into short-term learning. Given that one premise of the WOW curriculum is to increase vocabulary and conceptual knowledge in preschoolers to provide a foundation for *later* learning, it is important for future research to follow children longitudinally to determine if the effects of the curriculum are retained. Further, following children into kindergarten and first grade, as children begin to read conventionally, would help establish the hypothesized link between this type of learning and reading comprehension. In part, this is already being done. During the year subsequent to collection of the data used in this dissertation study, The Ready to Learn Project (under the direction of Susan B. Neuman) engaged in a year-long implementation of the WOW curriculum and then followed children who experienced the WOW curriculum into their homes and kindergarten classrooms. The team (of which I am no longer a part) is currently analyzing these longitudinal data for long term retention of words and concepts. I eagerly await these findings; if results support long term retention, the findings reported in this dissertation study will be replicated and extended.

Conclusion

This study represents an initial investigation into the power of the unique instructional approach of teaching vocabulary and conceptual knowledge in taxonomic categories. Overall, there is strong evidence that teaching words and concepts in a taxonomic structure was effective in increasing target vocabulary and target conceptual knowledge. Most provocatively, there was also robust evidence that this type of instruction fostered transfer and application of conceptual knowledge to new learning tasks. The positive effects of this unique approach on knowledge of curricular concepts and vocabulary, coupled with the clear capacity for children to flexibly use this knowledge beyond the confines of the curricular lessons, suggest that this type of vocabulary and conceptual knowledge instruction has the potential to foster future vocabulary learning, concept acquisition, and possibly comprehension of text. It is the potential of this type of instruction to provide a foundation of knowledge and a structure that is facilitative of future learning that may begin to narrow the stark vocabulary gaps we know exist between children from disparate socioeconomic backgrounds.

Appendix A

Descriptive Tables Comparing Children Who Moved and Who Were Chronically Absent With All Other Children in the Study

Table A

Descriptives of Children Who Moved and Who Did Not Move

	Moved (N=24)	Did Not Move (N=298)
Age in Months at Pretest	51.38	51.49
Total Days Absent	38.52***	9.28
Pretest Receptive Vocabulary (PPVT)	89.45	89.60
Pretest Curriculum Vocabulary	14.58	15.73
Pretest Expressive Vocabulary (GGG)	15.78	16.11
% Female	46[a]	51
% AM Preschool	67	49
% White	75	56
% Black	17	26
% Hispanic	0	2
% Other	8	16

*p<.05; **p<.01; ***p<.001

[a] All percentages should be read as the proportion within the group that moved or did not move. For example, for gender, it should be read as "Within the group of children who moved, 46% were female."

Table B

Descriptives of Treatment Children Who Were Absent More Than 40% of Study Days

	Absent > 40% of Days (N=8)	Absent < 40% of Days (N=169)
Age in Months at Pretest	49.63	52.34
Pretest Receptive Vocabulary (PPVT)	87.50	92.99
Pretest Curriculum Vocabulary	14.88	17.13
Pretest Expressive Vocabulary (GGG)	11.00	16.68
% Female	38[a]	52
% AM Preschool	62	50
% White	50	60
% Black	25	24
% Hispanic	13	2
% Other	13	14

*p<.05; **p<.01; ***p<.001

[a] All percentages should be read as the proportion within the group that was absent > than 40% of days or absent < 40% of the days. For example, for gender, it should be read as "Within the group of children who were absent > 40% of the days, 38% were female."

Appendix B

How Preschool and Kindergarten State Standards Related to Taxonomic Categories Included in the WOW Curriculum

	California*	Massachussetts	Michigan	Texas	Indiana**
UNIT 1: LIVING THINGS					
<u>PETS</u>					
Preschool Standard		Observe and identify characteristics and needs of living things: animals, plants, and humans.	Children use observation skills to build awareness of plants and animals, their life cycles, and basic needs.	Child identifies and describes the characteristics of organisms; Child recognizes, observes, and discusses the relationship of organisms to their environments.	
Kindergarten Standard		Students will recognize that animals (including humans) and plants are living things that grow, reproduce, and need food, air, and water.			Give examples of plants and animals. Observe plants and animals, describing how they are alike and how they are different in the way they look and in the things they do.
<u>WILD ANIMALS</u>					
Preschool Standard		Observe and identify characteristics and needs of living things: animals, plants, and humans.	Children use observation skills to build awareness of plants and animals, their life cycles, and basic needs.	Child identifies and describes the characteristics of organisms; Child recognizes, observes, and discusses the relationship of organisms to their environments.	

Kindergarten Standard	Students know how to observe and describe similarities and differences in the appearance of plants and animals (e.g. seed-bearing plants, birds, fish, insects). Students know how to identify major structures of common plants and animals (e.g., stems, roots, arms, wings, legs).	Students will recognize that animals (including humans) and plants are living things that grow, reproduce, and need food, air, and water.	Explain characteristics and fuctions of a variety of observable body parts in animals (vertebrate and invertebrate animals such as humans, cows, sparrows, goldfish, spiders, crayfish, insects). Compare and contrast familiar organisms on the basis of observable physical characteristics.		Give examples of plants and animals. Observe plants and animals, describing how they are alike and how they are different in the way they look and in the things they do.
<u>ANIMALS IN WATER</u>					
Preschool Standard		Observe and identify characteristics and needs of living things: animals, plants, and humans.	Children use observation skills to build awareness of plants and animals, their life cycles, and basic needs.	Child identifies and describes the characteristics of organisms; Child recognizes, observes, and discusses the relationship of organisms to their environments.	
Kindergarten Standard	Students know how to observe and describe similarities and differences in the appearance of plants and animals (e.g. seed-bearing plants, birds, fish, insects). Students know how to identify major structures of common plants and animals (e.g., stems, roots, arms, wings, legs).	Students will recognize that animals (including humans) and plants are living things that grow, reproduce, and need food, air, and water.	Explain characteristics and fuctions of a variety of observable body parts in animals (vertebrate and invertebrate animals such as humans, cows, sparrows, goldfish, spiders, crayfish, insects). Compare and contrast familiar organisms on the basis of observable physical characteristics.		

<u>INSECTS</u>					
Preschool Standard		Observe and identify characteristics and needs of living things: animals, plants, and humans.	Children begin to describe relationships among familiar plants and animals (e.g. catepillars eat leaves).	Child identifies and describes the characteristics of organisms; Child recognizes, observes, and discusses the relationship of organisms to their environments.	
Kindergarten Standard	Students know how to observe and describe similarities and differences in the appearance of plants and animals (e.g. seed-bearing plants, birds, fish, insects). Students know how to identify major structures of common plants and animals (e.g., stems, roots, arms, wings, legs).		Compare and contrast familiar organisms on the basis of observable physical characteristics (animals that looks similar-- snakes, worms, millipedes, flowering and non flowering plants; pine tree, oak tree, rose, algea).	The student knows that systems have parts and are composed of organisms and objects. The student is expected to sort organisms and objects into groups according to their parts and describe how groups are formed. The student is expected to record observations about parts of animals including wings, feet, heads, and tails.	Give examples of plants and animals. Observe plants and animals, describing how they are alike and how they are different in the way they look and in the things they do.

UNIT 2: HEALTH	**California**	**Massachusetts**	**Michigan**	**Texas**	**Indiana**
<u>PARTS OF THE BODY</u>					
Preschool Standard		Listen to and use the appropriate language describing the names and functions of parts of the human body.	Children begin to recognize and name parts of the body and their locations.	Child identifies similarities and differences in characteristics of people.	

Kindergarten Standard	Students identify and describe parts of the body: the head, shoulders, neck, back, chest, waist, hips, arms, elbows, wrists, hands, fingers, legs, knees, ankles, feet, and toes. Students identify the body part involved when stretching.	Students will name the external and internal parts of the body and body systems.		The student is expected to identify selected body parts such as head*, back*, chest*, waist, hips, arms*, elbows*, wrists, hands*, fingers*, legs*, knees*, ankles, feet*, and toes*. The student knows the basic structures and functions of the human body and how they relate to personal health. The student is expected to name the five senses, and name the major parts of the body and their functions.	Describe the basic structure and functions of the human body systems.
<u>CLOTHES</u>					
Preschool Standard			Begin to identify what families need to thrive (e.g., food, shelter, clothing, love	Child demonstrates that all people need food, clothing, and shelter.	
Kindergarten Standard			Describe ways people use the environment to meet human needs and wants (e.g., food, shelter, clothing).		

<u>**EXERCISE**</u>					
Preschool Standard		Build body awareness, strength, and mobility through various locomotor activities.	Begin or continue to develop traveling movements such as walking, jogging, running, climbing, jumping, hopping, skipping, marching, and galloping.	Child identifies good habits of nutrition and exercise.	
Kindergarten Standard	Students identify the locomotor skills of walk, jog, run, hop, jump, slide, and gallop. Students identify physical activities that are enjoyable and challenging. Students describe the benefits of being physically active. Students identify the location of the heart and explain that it is a muscle. Students explain that physical activity increases the heart rate	Students will explain the benefits of physical fitness to good health and increased active lifestyle.	Generate examples of physical activity that are personally enjoyable. Movement concepts including the following locomotor skills: walk, run, leap, jump, skip, hop, gallop, slide, chase, flee, and dodge.	The student is expected to identify types of exercise and active play good for the body. The student is expected to describe and select physical activities that provide opportunities for enjoyment and challenge. The student is expected to describe the benefits from involvement in daily physical activity such as feel better and sleep better.	Perform locomotor and non locomotor skills at the beginning level.

*California has a standards document for preschool called "Preschool Learning Standards". However, the only content knowledge included in those standards has to do with math.

**Indiana does not have preschool standards.

Appendix C

Examples of WOW Curriculum Materials

Figure 1

Suggested Teacher Language from Insects Teacher Manual

Let's Get Started!

Say: **Who can tell me one kind of insect that we've learned about? Raise your hand and wait for me to call on you.**

Call on individual children to name katydids, moths, and bees.

Say: **Katydids, moths, and bees are all insects. They all have the same number of legs. Who can tell me how many legs an insects has? Raise your hand and wait for me to call on you.**

Call on a child to give the answer, then have the class count to six with you as you hold up fingers.

Say: **All insects have the same number of body segments. Who can tell me how many body segments an insect has? Raise your hand and wait for me to call on you.**

Call on a child to give the answer, then have the class count to three with you as you hold up fingers.

Say: **What else can you tell me about insects? Raise your hand and wait for me to call on you.**

Call on several children to tell what they know about insects. Prompt them by asking questions such as the following:

What body part do insects use to feel?
Where do insects live most of the time?
How do insects protect themselves?

Say: **You know so much about insects! Congratulations for being such good learners. Let's watch a new video. This one is about another kind of insect. Watch and pay close attention!**

Play Insects 3: Chris Columbus and Ants

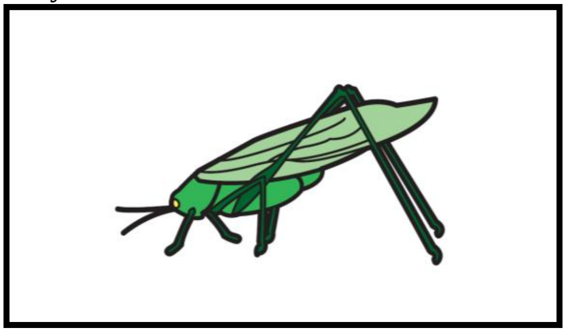
Say: **Let's talk about the video.**

22

Figure 2

Sample Picture Cards from Insects Lessons

Katydid-Insect



Centipede-Not an Insect

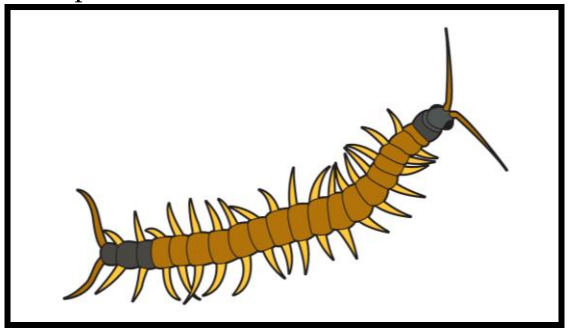


Figure 3

Sample Teacher Support for Discussion Challenge Items

WHY ISN'T IT AN INSECT?
Although the INSECTS book points out that spiders and centipedes aren't insects, children may still believe that they are. Indeed, spiders and centipedes seem to meet most of the criteria for insects and are often called bugs. *The points to stress with the children is that although spiders are small creatures that live in the same habitat as insects, they are not insects because they do not have six legs or three body parts. Spiders have eight legs and two body parts. Centipedes have many legs and many body parts.*

Time for a Challenge!

Display all of the Picture Cards in a row and have children name each small creature at you point to it. For each picture, have children tell whether the creature is an insect and how they can tell whether it is or is not an insect.

Say: **I have a challenge for you. Remember that a challenge means that you have to think very hard to figure something out. Listen closely.**

Say: **I'm going to show you a picture of another small creature. And I will challenge you to tell me whether it is or is not an insect. Remember what you've learned about what is and is not a insect. Remember that not all small creatures are insects, even though we sometimes call them bugs. Think hard now! Let's go.**

Teacher: [Show Spider Picture Card.] This is a spider. Remember that we've seen spiders in our videos and read about them in our book. Who thinks a spider is an insect? Raise your hand if you do.

Call on each child who raises a hand to give a reason for her/his answer. As necessary, prompt them to use information they have learned in this unit. Then point to the spider's legs and have the children count them aloud with you as you point to each leg. Do the same for the body parts. Help children to see that a spider is NOT an insect.

Follow the same procedure for the Centipede Picture Card.

Let's Wrap-up!

Say**: What terrific work! You thought very hard and figured out my challenge. Good job everybody!**

Appendix D

List of 92 More Difficult Words Taught in the WOW Curriculum

Pets

- **Types of Pets**: goldfish, rabbit, hamster
- **Words That Help Us Talk About Pets**: tame, pet, check-up, shelter, veterinarian

Wild Animals

- **Types of Wild Animals:** coyote, deer, gazelle, gorilla, hippo, hyena, rhino, polar bear
- **Words That Help Us Talk About Wild Animals:** arctic, desert, ferocious, fishing, hunt, grassland, jungle, woods, habitat

Animals in Water

- **Types of Animals in Water:** whale, octopus, shark, stingray, starfish, dolphin, seahorse
- **Words That Help Us Talk About Animals in Water:** fish tank, ocean, river, sea, aquarium, breathe, coral, fin, gill, snorkel, survive, wild

Insects

- **Types of Insects:** katydid, ladybug, moth
- **Words That Help Us Talk About Insects:** antennae, segment, creature, wing, anthill, camouflage, cooperate, hive, honey, leaf, protect, sting, warn, worm

Parts of the Body

- **Types of Parts of the Body:** back, chest, elbow, torso
- **Words That Help Us Talk About Parts of the Body:** attached, job, move, sense, five, skip, tie, wave, scratch, grab, feel, body, smell, bend, nod, snap

Clothes

- **Types of Clothes:** helmet, t-shirt, bathing suit
- **Words That Help Us Talk About Clothes:** take off, undress, warm, zip, bottom, bare, lace, sleeve

Appendix E

Coding Scheme for Conceptual Justification of Categorization Choices

Score	Characteristics	Examples
1	SORTED CORRECTLY ORGANIZED	IN CATEGORY “A dragonfly is an insect because it has six legs” OUT OF CATEGORY “A snake is not an insect because it doesn’t have 3 body parts”
1	SORTED INCORRECTLY ORGANIZED[a]	IN CATEGORY “A dragonfly is NOT an insect because it Has 1,2,3,4,5 legs” OUT OF CATEGORY “A crab is an insect because it has six legs”
0	SORTED CORRECTLY DIFFUSELY ORGANIZED	IN CATEGORY “A dragonfly is an insect because it has five legs” OUT OF CATEGORY “A mouse is not an insect because my mom told me”

a. NOTE: Here, the child sorted the item incorrectly but provided a justification that was consonant with their sorting choice and used correct conceptual information for their choice **(although the information may be inaccurate for that item, it is true of the category).**

WILD ANIMALS CONCEPTS		
Organized Response for In-Category Items	Organized Response for Out-of-Category Items	Diffusely Organized Responses
• It lives outside and away from people/It lives away from people/It lives outside. • It takes care of itself. • It finds it’s own food. • It is not tame/It doesn’t like to be petted/It doesn’t like to be played with. • It is too big. • It is ferocious. • It is not a pet. • It lives in the water/woods/jungle/forest.	• It lives with people. • It doesn’t take care of itself/people take care of it. • It doesn’t find its own food/People feed it/People give it food. • It is tame/It likes to be petted/It likes to be played with. • It is small. • It is not ferocious. • It is a pet/It is a farm animal.	• Because. • Don’t know. • Cause my mom told me. • Cause Picky Peter likes it. • Cause I know. • Cause I saw it.

Appendix F

Coding Scheme for Conceptual Knowledge Use Assessment

ACCEPTABLE CONCEPTS

PETS

- They are animals that live with people.
- They are tame.
- You have to feed them by giving them food and water.
- You take care of them (ie. you clean them), take them to the vet.(not doctor, pet doctor is ok)
- You give them exercise/they run around.
- You have to play with them, they play.
- You have to love them.
- They can be purchased in a pet store.
- Live in a cage.
- You pet them.

WILD ANIMALS

- They are animals that live outside and away from people.
- They can live in a zoo.
- They live in different habitats (if the child lists different types of habitats, the child earns a point for each one, including grassland, jungle, water, desert).
- They live in the wild.
- They can't be pets.
- They are often big.
- Can't pet wild animals.

INSECTS

- They are very small creatures/animals.
- They mostly live outside.
- They have three body parts called segments.
- They have six legs.
- They have special ways to protect themselves from bigger animals.
- Most have wings/most fly.

ANIMALS THAT LIVE IN WATER

- Some animals live in water all the time.
- They swim.
- Some have gills which they use to help them breathe under water (just saying sharks or fish have gills is sufficient).
- Some can breathe out of the water/come to the surface (or top) of the water to breathe

PARTS OF THE BODY

- Our bodies have many parts.
- Each body part has a job to do (i.e. some parts help us move, other parts are related to the five senses).
- Body parts are attached to our bodies and go together.
- Body parts do not come off.
- Body parts come in different numbers.
- Bodies need to be healthy and strong.

CLOTHES

- We wear them on our bodies.
- We put on them on when we get dressed.
- We take them off when we get undressed.
- We wear different clothes on different parts of our bodies.
- Some clothes keep us warm by covering our bodies.
- Some clothes keep us safe by protecting parts of our bodies.

Appendix G

Description of Codes for Analysis of Circle Time Videos

Overall Circle Time	Circle time is defined as any time children sit together as one group and the teacher engages them in a whole-group activity. Circle time begins when the teacher begins the first whole-group activity. Circle time ends when the teacher dismisses the first student or whole group to another activity. To create this variable, the total time spent in all circle times observed during Observation 1 was summed.
Book reading	This could mean any kind of book, whether fiction, informational, or the WOW book. Book reading begins when the teacher begins talking about the cover of the book or begins reading.
Calendar	Activity discussing date, day, month, year.
Weather	Discusses condition of the weather outside.
Attendance	Time spent discussing how many children are in school today, or who is/are absent.
Daily News	Time spent on finding out and sharing information about particular children.
Singing	Anything that is a song, regardless of its purpose.
Group-writing	Teacher-led writing activity, with the children contributing to the writing process.
Alphabet activity	Identifying or reciting letters of the alphabet.
Vocabulary	Emphasizing new terms; providing definition of term or giving example to count as vocabulary.
Rhyming	Explicitly teaching about the correspondence of words' ending sounds either by defining a rhyme or the use of examples.
Beginning sound awareness	Identifying/isolating initial sound of words.
Other phonological awareness	Blending, segmenting, letter-sound correspondence
Conventions of text	Time spent teaching children about concepts of print as well as conventions of book-reading.
Other Language Art Skill	Time spent engaged in any other language art skill.
Colors	Labeling, comparing, or working with colors and/or color concepts.
Numbers/ Counting	Identifying, reciting, or quantifying numbers (with or without context).
Shapes	Identifying, categorizing, or in other ways working with shapes and/or shape concepts.
Patterns	Identifying, predicting, or working with patterns and/or pattern concepts.
Science	Time spent on instruction related to science activities.

Appendix H

Comparison of Mean Time Spent During Circle Times on Instructional Activities

in Treatment and Comparison Classrooms

	Treatment Circle Times: Number of Minutes Spent on Activity	Comparison Circle Times: Number of Minutes Spent on Activity
Overall Time Spent in Circle	31.19	32.28
Book reading	8.38	6.56
Calendar	1.22	1.78
Weather	.34	.47
Attendance	1.80	.39
Daily News	.17	1.07
Singing	4.96	6.37
Group-writing	.37	0
Alphabet activity	.69	1.87
Vocabulary	.06	.38
Rhyming	.08	.19
Beginning sound awareness	.61	.07
Other phonological awareness	.64	.56
Conventions of text	.40	.57
Other Language Art Skill	.44	.32
Colors	.22	1.93*
Numbers/ Counting	2.80	1.58
Shapes	.36	.43
Patterns	.10	.47
Science	.07	.07

*p<.05; **p<.01; ***p<.001

Appendix I: Results Across Different Analytic Approaches

Table A. OLS Regressions Using the Full Sample and PPVT as a Covariate

	General Receptive Vocabulary (PPVT)	General Expressive Vocabulary (GGG)	Word Knowledge Time 3	Word Knowledge Time 2	More Difficult Word Use Time 3	Conceptual Knowledge Use Time 3	Conceptual Categorization of Items Taught Time 2	Conceptual Categorization of Items **Not Taught** Time 2	Justification of Conceptual Categorization-Items Not Taught
Treatment	-.19* (.09)	.31** (.11)	.40*** (.08)	.90*** (.10)		.56*** (.12)	.68*** (.12)	.61*** (.12)	.72*** (.11)
PPVT Time 1	.48*** (.06)	.18* (.07)	.26*** (.05)	.21*** (.05)		.22** (.08)	.28*** (.06)	.26*** (.07)	.23*** (.06)
Word Knowledge Time 1	.39*** (.07)	.22** (.08)	.52*** (.06)	.39*** (.06)		.09 (.09)	-.02 (.07)	.04 (.08)	.11 (.07)
Pretest		.26*** (.07)				.30*** (.06)			
Word Knowledge Time 2							.32*** (.07)		
Word Knowledge-Not Taught Items								.24*** (.08)	.20** (.07)
Minority	-.09 (.09)	.06 (.10)	.02 (.07)	.04 (.09)		-.12 (.12)	.02 (.09)	-.02 (.10)	-.07 (.09)
Age	-.02** (.01)	.01 (.01)	.02* (.01)	.01 (.01)		.01 (.01)	.03** (.01)	.04*** (.01)	.03*** (.01)
Female	.02 (.08)	-.22* (.10)	.00 (.07)	.06 (.08)		.14 (.11)	-.07 (.08)	-.08 (.09)	.05 (.09)
Absences	-.01 (.01)	.00 (.01)	-.01 (.01)	-.01~ (.01)		-.01 (.01)	-.03*** (.01)	-.02* (.01)	-.02** (.01)
PM	.07 (.09)	.22* (.10)	.11 (.07)	.04 (.08)		.10 (.12)	-.17* (.08)	-.02 (.09)	.20* (.09)
R^2	.54***	.40***	.69***	.65***		.40***	.64***	.57***	.60***

***p ≤ .001; **p ≤ .01; *p ≤ .05; ~p ≤ .10

Table B. OLS Regressions Using the Sample Matched on PPVT Pretest

	General Receptive Vocabulary (PPVT)	General Expressive Vocabulary (GGG)	Word Knowledge Time 3	Word Knowledge Time 2	More Difficult Word Use Time 3	Conceptual Knowledge Use Time 3	Conceptual Categorization of Items Taught Time 2	Conceptual Categorization of Items **Not Taught** Time 2	Justification of Conceptual Categorization-Items Not Taught
Treatment	-.13 (.11)	.32* (.14)	.33** (.11)	.82*** (.12)	.68*** (.13)	.50** (.16)	.56*** (.15)	.48*** (.14)	.67*** (.14)
Word Knowledge Time 1	.27*** (.08)	.37*** (.09)	.67*** (.06)	.54*** (.06)	.35*** (.08)	.24* (.09)	.13 (.08)	.18~ (.09)	.13 (.09)
Pretest	.55*** (.07)	.26** (.09)			.33*** (.08)	.35*** (.08)			
Word Knowledge Time 2							.35*** (.09)		
Word Knowledge-Not Taught Items								.32*** (.09)	.30*** (.09)
Minority	-.16 (.12)	.09 (.14)	-.09 (.11)	-.03 (.12)	-.03 (.13)	-.28~ (.15)	-.12 (.12)	-.12 (.13)	-.27* (.01)
Age	-.02 (.01)	-.01 (.01)	.02~ (.01)	.01 (.01)	-.02~ (.01)	-.00 (.02)	.02* (.01)	.03* (.01)	.02 (.01)
Female	.17 (.11)	-.24~ (.13)	.12 (.10)	.11 (.11)	.03 (.12)	.26~ (.15)	.09 (.12)	.03 (.13)	.09 (.13)
Absences	.00 (.01)	.00 (.01)	.00 (.01)	-.01 (.01)	-.01 (.01)	-.01 (.01)	-.02~ (.01)	-.02 (.01)	-.03** (.01)
PM	.05 (.11)	.19 (.14)	.10 (.10)	.02 (.11)	.12 (.13)	.04 (.15)	-.32** (.12)	-.11 (.13)	.17 (.13)
R^2	.55***	.33***	.60***	.61***	.47***	.41***	.56***	.49***	.52***

***p ≤ .001; **p ≤ .01; *p ≤ .05; ~p ≤ .10

Table C. HLMs Using the Full Sample with PPVT as a Covariate

	General Receptive Vocabulary (PPVT)	General Expressive Vocabulary (GGG)	Word Knowledge Time 3	Word Knowledge Time 2	More Difficult Word Use Time 3	Conceptual Knowledge Use Time 3	Conceptual Categorization of Items Taught Time 2	Conceptual Categorization of Items **Not Taught** Time 2	Justification of Conceptual Categorization-Items Not Taught
Treatment	-.17 (.14)	.28~ (.15)	.40*** (.08)	.90*** (.11)		.56*** (.13)	.68*** (.13)	.61*** (.14)	.72*** (.12)
PPVT Time 1	.49*** (.06)	.18* (.07)	.26*** (.05)	.20*** (.05)		.21** (.08)	.28*** (.06)	.26*** (.07)	.23*** (.06)
Word Knowledge Time 1	.37*** (.07)	.23** (.08)	.52*** (.06)	.40*** (.06)		.09 (.09)	-.02 (.07)	.04 (.08)	.12~ (.07)
Pretest		.26*** (.07)				.30*** (.06)			
Word Knowledge Time 2							.32*** (.07)		
Word Knowledge-Not Taught Items								.25** (.08)	.20** (.07)
Minority	-.04 (.09)	.04 (.11)	.02 (.07)	.02 (.09)		-.11 (.12)	.03 (.09)	-.01 (.10)	-.07 (.09)
Age	-.02* (.01)	.00 (.01)	.02* (.01)	.01 (.01)		.01 (.01)	.03*** (.01)	.04*** (.01)	.03*** (.01)
Female	.01 (.08)	-.20* (.10)	.00 (.07)	.06 (.08)		.14 (.11)	-.07 (.08)	-.09 (.09)	.05 (.09)
Absences	.01 (.01)	.00 (.01)	-.01 (.01)	-.01~ (.01)		-.01 (.01)	-.03*** (.01)	-.02* (.01)	-.02* (.01)
PM	.04 (.09)	.22* (.10)	.11 (.07)	.04 (.08)		.10 (.12)	-.17* (.08)	-.01 (.09)	.20* (.09)

***p ≤ .001; **p ≤ .01; *p ≤ .05; ~p ≤ .10

Table D. HLMs Using the Sample Matched on PPVT Pretest

	General Receptive Vocabulary (PPVT)	General Expressive Vocabulary (GGG)	Word Knowledge Time 3	Word Knowledge Time 2	More Difficult Word Use Time 3	Conceptual Knowledge Use Time 3	Conceptual Categorization of Items Taught Time 2	Conceptual Categorization of Items **Not Taught** Time 2	Justification of Conceptual Categorization-Items Not Taught
Treatment	-.11 (.15)	.28 (.18)	.34** (.10)	.84*** (.15)	.71*** (.14)	.55* (.19)	.62** (.17)	.55* (.19)	.74** (.19)
Word Knowledge Time 1	.21** (.07)	.35*** (.08)	.72*** (.05)	.56*** (.06)	.29*** (.07)	.22* (.09)	.12 (.08)	.18* (.09)	.12 (.09)
Pretest	.57*** (.07)	.25** (.08)			.33*** (.07)	.37*** (.11)			
Word Knowledge Time 2							.36*** (.09)		
Word Knowledge-Not Taught Items								.32*** (.09)	.31*** (.09)
Minority									
Age							.03** (.01)	.03** (.01)	.03* (.01)
Female									
Absences									-.02* (.01)
PM							-.31** (.12)		

***p ≤ .001; **p ≤ .01; *p ≤ .05; ~p ≤ .10

Table E. Influence of Treatment on Each Outcome in Four Different Analytic Approaches

	General Receptive Vocabulary (PPVT)	General Expressive Vocabulary (GGG)	Word Knowledge Time 3	Word Knowledge Time 2	More Difficult Word Use Time 3	Conceptual Knowledge Use Time 3	Conceptual Categorization of Items Taught Time 2	Conceptual Categorization of Items **Not Taught** Time 2	Justification of Conceptual Categorization-Items Not Taught
Full Sample									
OLS Regression	-.19*	.31**	.40***	.90***		.56***	.68***	.61***	.72***
	(.09)	(.11)	(.08)	(.10)		(.12)	(.12)	(.12)	(.11)
HLM	-.17	.28~	.40***	.90***		.56***	.68***	.61***	.72***
	(.14)	(.15)	(.08)	(.11)		(.13)	(.13)	(.14)	(.12)
Matched Sample									
OLS Regression	-.13	.32*	.33**	.82***	.68***	.50**	.56***	.48***	.67***
	(.11)	(.14)	(.11)	(.12)	(.13)	(.16)	(.15)	(.14)	(.14)
HLM	-.11	.28	.34**	.84***	.71***	.55*	.62**	.55*	.74**
	(.15)	(.18)	(.10)	(.15)	(.14)	(.19)	(.17)	(.19)	(.19)

***p ≤ .001; **p ≤ .01; *p ≤ .05; ~p ≤ .10

References

Anderson, R.C., & Freebody, P. (1981). Vocabulary knowledge. In J.T. Guthrie (Ed.), *Comprehension and Teaching* (pp. 77-117). Newark, DE: International Reading Assocation.

Anderson, R.C., & Pearson, P.D. (1984). A schema-theoretic view of basic processes in reading comprehension. *Handbook of literacy research (Vol. I)* (pp.). Mahwah, NJ: Erlbaum.

Bauer, P.J., & Mandler, J.M. (1989). Taxonomies and triads: Conceptual organization in one- to two-year-olds. *Cognitive Psychology, 21,* 156-184.

Baumann, J. F. (2009). Vocabulary and reading comprehension: The nexus of meaning. In S.E. Israel & G.G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 323-346). New York, NY: Routledge.

Beals, D.E. (1997). Sources of support for learning words in conversation: Evidence from mealtimes. *Journal of Child Language, 24,* 673-694.

Beck, I.L., McCaslin, E.S., & McKeown, M.G. (1980). *The rationale and design of a program to teach vocabulary to fourth-grade students.* Pittsburgh: University of Pittsburgh, Learning Research and Development Center.

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,* 251-271.

Beck, I.L., McKeown, M.G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York, NY: Guildford Press.

Becker, W. C. (1977). Teaching reading and language to the disadvantaged-What we have learned from field research. *Harvard Educational Review, 47,* 518-543.

Biemiller, A. (2003). Oral comprehension sets the ceiling on reading comprehension. *American Educator, Spring*.

Biemiller, A. (2004). Teaching vocabulary in the primary grades: Vocabulary instruction needed. In J.F. Baumann, & E.J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp. 28-40). New York: Guilford.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research (Vol. 2)* (pp. 41-51). New York: Guildford Press.

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98*, 44-62.

Biemiller, A., & Slonim, N. (2001). Estimating root word vocabulary growth in normative and advantaged populations: Evidence for a common sequence of vocabulary acquisition. *Journal of Educational Psychology, 93*, 498-520.

Blachowicz, C. & Fisher, P. (2000). Vocabulary instruction. In M. Kamil, P. Mosenthal, D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3)(pp. 503-523). Mahwah, NJ: Lawrence Erlbaum Associates.

Blewitt, P., Rump, K.M., Sealy, S.E., & Cook, S.A. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology*, *101*, 294-304.

Bloom, P. (2000). How *children learn the meanings of words*. Cambridge, MA: MIT Press.

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language, 33*, 759-790.

Brabham, E.G., & Lynch-Brown, C. (2002). Effects of teachers' reading-aloud styles on vocabulary acquisition and comprehension of students in the early elementary grades. *Journal of Educational Psychology, 94*, 465-473.

Bus, A.G., van IJzendoorn, M.H., & Pelligrini, A.D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, *65,* 1-21.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G.A. Miller (Eds.*), Linguistic theory and psychological reality*. Cambridge, MA: MIT Press.

Carey, S. (2009). *The origin of concepts*. Oxford, UK: Oxford University Press.

Chall, J.S. (1996). *Learning to read: The great debate* (3rd ed.). Fort Worth: Harcourt Brace.

Chall, J.S., & Jacobs, V.A. (2003). The classic study on poor children's fourth grade slump. *American Educator, Spring*, 14-15, 44.

Chall, J.S., Jacobs, V.A., & Baldwin, L.E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Brookline Books.

Clark, E. V. (1983). Meanings and concepts. In J.H. Flavell, & E.M. Markman (Eds.). *Cognitive development* (Vol. III). New York, NY: Wiley & Sons.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Cunningham, A.E., & Stanovich, K.E. (1997). Early reading acquisition and its relation to reading experience and ability ten years later. *Developmental Psychology, 33*, 934-45.

Denton, K., West, J., & Walston, J. (2003). *Young children's achievement and classroom experiences: Findings from the condition of education.* Washington, D.C.: National Center for Education Statistics.

Dickinson, D.K., & Smith, M.W. (1994). Long-term effects of preschool teachers' book readings on low-income children's vocabulary and story comprehension. *Reading Research Quarterly, 29*, 104-122.

Dodge, D.T., Colker, L., & Heroman, C. (2002). The creative curriculum for preschool. Cengage Learning.

Dunn, L.M., & Dunn, L.M. (1997). *Peabody Picture Vocabulary Test* (3rd ed)(PPVT III). Circle Pines, MN: American Guidance Service.

Early Childhood Research Institute on Measuring Growth and Development (1998). Research and development of individual growth and development indicators for children between birth and age eight (Tech. Rep. No. 4), Minneapolis, MN: Center for Early Education and Development, University of Minnesota.

Elley, W.B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, *24*, 174-187.

Ewers, C.A., & Brownson, S.M. (1999). Kindergarteners' vocabulary acquisition as a function of active vs. passive storybook reading, prior vocabulary, and working memory. *Reading Psychology*, *20*, 11-20.

Fenson, L., Dale, P., Reznick, J., Bates, E., Thal, D., & Pethick, S. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59* (5, Serial No. 242).

Gelman, S.A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65-95.

Gelman, S.A., & Coley, J.D. (1990). The importance of knowing a Dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26,* 796-804.

Gelman, S.A., & Kalish, C.W. (2006). Conceptual development. In D. Kuhn., & R.S. Siegler (Volume Eds.). *Handbook of Child Psychology* (Vol. 2) (6th ed.)(pp. 687-733). Hoboken, NJ: John Wiley & Sons, Inc.

Gelman, S.A., & Markman, E.M. (1986). Categories and induction in young children. *Cognition, 23,* 183-209.

Gelman, S.A., & Markman, E.M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58,* 1532-1541.

Gelman, S.A., & Medin, D. (1993). What's so essential about essentialism? A different perspective on the interaction of perception, language, and conceptual knowledge. *Cognitive Development, 8,* 157-168.

Gelman, S.A., Coley, J.D., Rosengren, K.S., Hartman, E., & Pappas, A. (1998). Beyond labeling: The role of maternal input in the acquisition of richly structured categories. *Monographs of the Society for Research in Child Development, 63* (Serial No. 253, No. 1).

Gelman, S.A., & Wellman, H.M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition, 38,* 213-244.

Gopnick, A. & Meltzoff, A. (1986). Relations between semantic and cognitive development in the one-word stage: The specificity hypothesis. *Child Development, 57,* 1040-1053.

Gopnick, A. & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58,* 1523-1531.

Gopnick, A., & Meltzoff, A. (1992). Categorization and naming: Basic-level sorting in eighteen-month-olds and its relation to language. *Child Development, 63,* 1091-1103.

Hargrave, A.C., & Senechal, M. (2000). Book reading intervention with language-delayed preschool children: The benefits of regular reading and dialogic reading. *Journal of Child Language*, *15,* 765-790.

Hart, B. & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young Amercian children*. Baltimore, MD: Brookes Publishing.

Hayes, D.P., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese'. *Journal of Child Language, 15,* 395-410.

Heit, E. (1997). Knowledge and concept learning. In K. Lamberts, & D. Shanks (Eds.), *Knowledge, concepts, and categories* (pp. 7-41). London: Psychology Press.

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative settings. *Child Development, 62,* 782-796.

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experiences and language development. *Applied Psycholinguistics, 19,* 603-629.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27,* 236-248.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. London, England: Routledge.

Jencks, C., & Phillips, M. (Eds.)(1998). *The black-white test score gap*. Washington, D.C.: Brookings Institution Press.

Keil, F.C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Keil, F.C. (1991). Theories, concepts, and the acquisition of word meaning. In S.A. Gelman, & J. P. Byrnes (Eds.), *Perspectives on language and thought* (pp.197-224). Cambridge, UK: Cambridge University Press.

Langer, J.A. (1984). Examining background knowledge and text comprehension. *Reading Research Quarterly, 19,* 468-481.

Lonigan, C.J., & Whitehurst, G. (1998). Relative efficacy of a parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13,* 263-290.

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

Markman, E.M. (1994). Constraints on word learning in early language. In L. Gleitman & B. Landau (Eds.). *The acquisition of the lexicon* (pp. 199-227). Cambridge: MIT Press.

Markman, E.M., & Hutchinson, J.E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16,* 1-27.

McKeown, M.G., Beck, I.L., & Omanson, R.C., & Pople, M.T. (1985). Some effects of the nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20,* 522-535.

McNamara, D.S., Kintsch, E., Songer, N.B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1-43.

Merton, R. (1968). The Matthew effect in science. *Science,* 56-63.

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. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Erlbaum.

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research, 53,* 253-279.

Mol, S.E., Bus, A.G., & deJong, 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,* 979-1007.

Murphy, G.L. (2002). *The big book of concepts.* Cambridge, MA: MIT Press.

Murphy, G.L., & Lassaline, M.E. (1997). Hierarchical structure in concepts and the basic level of categorization. In K. Lamberts, & D. Shanks (Eds.). *Knowledge, concepts, and categories* (pp. 93-132). Cambridge, MA: MIT Press.

Nagy, W.E., & Scott, J.A. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3)(pp. 269-284). Mahwah, NJ: Lawrence Erlbaum Associates.

National Early Literacy Panel (2009). *Developing early literacy: A scientific synthesis of early literacy development and implications for intervention.* Washington, DC: National Institute for Literacy.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific literature on reading and its implications for reading instruction.* Washington, DC: US Department of Education.

Neuman, S.B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in pre-k. *Reading Teacher, 62*, 384-392.

Neuman, S.B., Dwyer, J., Koh, S., & Wright, T. (2007). The World of Words (WOW) Vocabulary Curriculum.

Neuman, S.B., & Roskos, K. (2005). The state of state pre-kindergarten standards. *Early Childhood Research Quarterly, 20,* 125-145.

Noldus Information Technology (2001). The observer XT.

O'Donnell, C.L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K12 curriculum intervention research. *Review of Educational Research, 78,* 33-84.

Pan, B.A., Rowe, M.L, Singer, J.D., & Snow, C.E. (2005). Maternal correlates of growth in toddler vocabulary production in low income families. *Child Development, 76*, 763-782.

Penno, J.F., Wilkinson, 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 Pscyhology, 94,* 23-33.

Pressley, M. (2000). What should comprehension instruction be the instruction of? In M. Kamil, P. Mosenthal, D. Pearson, & R. Barr (Eds.), *Handbook of reading research* (Vol. 3)(pp. 545-561). Mahwah, NJ: Lawrence Erlbaum Associates.

Pressley, M., Disney, L., & Anderson, K. (2007). Landmark vocabulary instructional research and the vocabulary instructional research that makes sense now. In R.K. Wagner, A.E. Muse, & K.R. Tannenbaum (Eds.). *Vocabulary acquisition: Implications for reading comprehension* (pp. 205-232). New York, NY: Guilford Press.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2^{nd} ed.). Newbury Park, CA: Sage.

Robbins, C., & Ehri, L.C. (1994). Reading storybooks to kindergarteners helps them learn new vocabulary words. *Journal of Educational Psychology, 86,* 54-64.

Rumelhart (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B.C.Bruce, & W.F. Brewer (Eds.). *Theoretical issues in reading comprehension* (pp.). Mahwah, NJ: Erlbaum.

Ryder, R.J., & Graves, M.F. (1994). Vocabulary instruction presented prior to reading in two basal readers. *Elementary School Journal, 95,* 139-153.

Scarborough, H. S. (1998). Early identification of children at risk of reading disabilities: Phonological awareness and some other promising predictors. In Shapiro, Accerdo, & Capute (Eds.). *Specific reading disability: A view of the spectrum* (pp. 75-199). Timoniuum, MD: York Press.

Scarborough, H., & Dobrich, W. (1994). On the efficacy of reading to preschoolers. *Developmental Review, 14,* 245-302. Scarr, S., & Weinberg, R.A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review, 43,* 674-692.

Scarr, S., & Weinberg, R.A. (1978). The influence of "family background" on intellectual attainment. *American Sociological Review, 43,* 674-692.

Schwartz, R.G., & Terrell, B.Y. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language, 10,* 57-64.

Senechal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Child Language, 24,* 123-138.

Senechal, M., Ouelette, 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).

Senechal, M., Thomas, E., & Monker, J.A. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87,* 218-229.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Boston, MA: Houghton Mifflin.

Shavelson, R.J. (1996). *Statistical reasoning for the behavioral sciences* (3rd Edition). Needham Heights, MA: Allyn & Bacon.

Smiley, S.S. & Brown, A.L. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28,* 249-257.

Smith, M.W., Dickinson, D.K., Sangeorge, A., & Anastasopoulos, M.P.P. (2002). *Early language and literacy classroom observation (ELLCO) toolkit*. Baltimore, MD: Paul H. Brookes.

Snow, C.E., Arlmann-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., & Vorster, J. (1976). Mothers' speech in three social classes. *Journal of Psycholinguistic Research, 5,* 1-20.

Snow, C.E., Burns, S., & Griffin, P. (1998). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.

Spycher, P. (2009). Learning academic language through science in two linguistically diverse kindergarten classrooms. *Elementary School Journal, 109,* 359-379.

Stahl, S.A., & Fairbanks, M.M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research, 56,* 72-110.

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

Stanovich, K. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360-407.

Swanborn, M.S.L, & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*, 261-285.

Wasik, B.A., Bond, M.A., & Hindman, A. (2006). The effects of a language and literacy intervention on Head Start children and teachers. *Journal of Educational Psychology, 98,* 63-74.

Waxman, S. R. (2004). Everything had a name, and each name gave birth to a new thought: Links between early word learning and conceptual organization. In D.G. Hall, & S.R. Waxman (Eds.). *Weaving a lexicon*. Cambridge, MA: MIT Press.

Waxman, S. R., & Gelman, R. (1986). Preschooler's use of superordinate relations in classification. *Cognitive Development, 1,* 139-156.

Weizman, Z.O., & Snow, C.E. (2001). Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37,* 265-279.

White, T.G., Graves, M.F., & Slater, W.H. (1990). Growth of reading vocabulary in diverse elementary schools: Decoding and world meaning. *Journal of Educational Psychology, 82,* 281-290.