Developing Vocabulary and Conceptual Knowledge for Low-income Preschoolers: A Design Experiment

Susan B. Neuman

Julie Dwyer

University of Michigan

School of Education

610 E. University

Ann Arbor, MI 48109

sbneuman@umich.edu

PH: 734-615-4655

Developing Vocabulary and Conceptual Knowledge for Low-income Preschoolers: A Design Experiment

Abstract

The purpose of this design experiment was to examine two facets of a curriculum targeted to the needs of low-income preschoolers: word learning and concept learning through semantic clusters and category development. Six Head Start teachers in A.M. and P.M. programs and their children (N=89) participated in the treatment and six (N=89) in the comparison group. Phase I of the design experiment focused on treatment children's retention of words and the use of categorical information to act as a bootstrap to word learning and inference generation. Based on initial evidence from Phase I, adaptations to the instructional design were implemented in Phase II of the study. In Phase II, learning outcomes from instruction were compared between treatment and comparison groups. Our findings indicated that in Phase II (with advanced instructional design techniques), children demonstrated learning gains and retained words at a higher rate than easy words. Further, they not only learned these more difficult words but the categories and concepts they represented, enabling them to generate inferences to words not taught. Significant differences were recorded between treatment and comparison groups on word knowledge and category development. These results suggest that instructional design features may work to accelerate word learning for low-income children.

Developing Vocabulary and Conceptual Knowledge for Low-income Preschoolers: A Design Experiment

The essence of all instruction is helping children learn new concepts and the words that signify them. In particular, word knowledge—oral language vocabulary--plays a critical role in children's reading achievement (Dickinson, McCabe, & Essex, 2006; Kibby, 1995). Extensive research demonstrates that the size of a child's vocabulary in kindergarten is an effective predictor of reading comprehension in the middle elementary years (Biemiller, 2005). Further, orally tested vocabulary at the end of Grade 1 is a significant predictor of reading comprehension in high school (Cunningham & Stanovich, 1997; Stahl & Nagy, 2006).

Children's knowledge of word meanings is cumulative (Chall, Jacobs, & Baldwin, 1990). The more words children know, the easier it is to learn new words. Children with highly elaborated semantic knowledge are likely to have more ready and more fluent access to this information and it is this rich interconnected knowledge of concepts, not just individual words, that drives comprehension, and reading proficiency in later grades (Vellutino et al., 1996).

Consequently, the well-documented gap in vocabulary knowledge between economically disadvantaged children and their middle-class peers prior to entering the elementary school years (Hart & Risley, 1995) becomes of great concern if we are to improve reading achievement and decrease the knowledge disparities among poor and middle income children (Farkas & Beron, 2004). Moats (1999), for example, estimates that the difference at entry into first grade may be as large as 15,000 words, with linguistically disadvantaged children knowing about 5,000 words, compared to their advantaged peers who have 20,000 words. Hart and Risley (2003) argue that the accumulated experiences with words for children who come from poor circumstances compared with children from professional families may constitute a 30 million word catastrophe that is difficult, if not impossible to close over time.

Therefore, the earlier children can acquire a large and richly-structured vocabulary, the greater their reading comprehension is likely to be in the later grades (Hirsch, 2003). Nevertheless, available evidence (Beck & McKeown, 2007) indicates that there is little emphasis on the acquisition of vocabulary in school curriculum. For example, in a recent content analysis of 10 published early literacy programs adopted by Early Reading First recipients (Authors, 2009), we found little evidence of an instructional regime, a deliberate effort in curriculum materials to teach vocabulary to preschoolers. Unfortunately current instructional materials appear to offer little guidance to teachers who want to do a better job of teaching vocabulary to young children.

Despite the importance of vocabulary in predicting later achievement (Senechal & LeFevre, 2002; Storch & Whitehurst, 2002), few intervention studies have made significant gains in closing the gap in word knowledge between middle- and low-income students (Juel, Biancarosa, Coker, & Deffes, 2003). To date, storybook reading (Mole, Bus, de Jong, 2009) has been regarded as the most potent source for teaching vocabulary in early childhood. Studies, however, suggest that the effects of reading aloud to children may not be powerful enough to enhance low-income children's word knowledge (Elley, 1989; Penno, Wilkinson, & Moore, 2002). In a now-classic study, Elley (1989) demonstrated that 7-year olds showed an average vocabulary gain of 15% from an oral storybook reading when the words in the text were frequently mentioned, depicted in illustrations, and redundant in the surrounding context. However, less than half of these gains were demonstrated in a second story with different characteristics. Further, in a recent meta-analysis of 31 experiments (Mol, Bus & DeJon, 2009), researchers found that the strongest effect sizes appeared in highly controlled settings executed by examiners, not classroom teachers. Teachers seemed to have difficulty with fostering growing in young children's language and literacy skills. Together, this evidence suggests more intensive interventions might be needed to narrow the gap for less advantaged children.

In this article, we report on a supplemental multimedia vocabulary curriculum, known as the World of Words (Authors, 2007), designed to engage low-income preschoolers in these kinds of learning activities. It is based on a framework that capitalizes on word learning through category formation. Considered a major component of word learning (Bloom, 2000), category membership is one of the first pieces of information a child learns about a word (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Whitmore, Shore, & Smith, 2004). Learning to identify a furry, four-legged animal as a dog, for example, involves not just mapping the label 'dog' to one's household pet, but actually establishing a *concept* of 'what is' and 'what is not' dog. Existing evidence suggests that children use categories to gain information about unfamiliar terms (Gelman & O'Reilly, 1988; Kalish & Gelman, 1992) which therefore, may potentially help bootstrap word learning.

Traditionally, studies of curriculum development like ours and educational research have been considered two distinct enterprises. Goals related to curriculum development has been to produce instructional materials; scientific research, on the hand, the creation of knowledge. Such distinctions, however, have not served the educational community well, and could be a reason that curriculum development has not reliably improved (Clements, 2007). For example, although knowledge is usually created during curriculum development, this knowledge has seldom been explicated, published, or used to better understand the mechanisms that underlie learning.

Consequently, in this study we approached curriculum from a design perspective, focusing our efforts on learning how our pedagogical design might support vocabulary and conceptual development. From this perspective (Cobb et al., 2003), learning and development are both generative and dynamic; what is known about a learning process is applied to the anticipated curriculum; what is learned from enacting the curriculum is used to revise and better understand the conditions for learning. In this respect, the goal of our design experiment was not merely to empirically fine-tune 'what works." Rather, as a design experiment (Reinking & Bradley, 2008), it was equally important to understand how, when, and why the instructional design might work, and its implications for developing theory about word learning.

Overview of the Instructional Design

The World of Words (WOW) curriculum (Authors, 2007) is an embedded multimedia program designed to foster children's vocabulary and conceptual knowledge in pre-K. The 12-minute supplemental curriculum uses multimedia (video, pictures, books) to augment children's content learning at the same time it is designed to teach critical early literacy skills of vocabulary and conceptual learning. The curriculum consists of two science-based units—living things, and healthy habits, each organized across4 topics, with each lesson taught over an 8-day sequence.

As an illustration of the kind of instruction provided, consider the vocabulary instruction from the topic, "Insects." The 8-day sequence begins each day with a "tuning-in"—a rhyme, song or word play video "clip" that is shown from a DVD[1] to bring children together. The teacher follows this activity with additional examples, engaging the children in a briskly-paced 'call and response' set of interactions.

The "tuning-in" is followed by a 'content' video, introducing children to the definition of the category. The first video is designed to act a prototype of the category, a particularly salient exemplar of the topic (i.e., a katydid). After the video, the teacher engages the children, focusing on 'wh' questions. She might ask, "Where does a katydid live? What is an insect?" The words are then reinforced using an information book (i.e. in this case on insects) specially designed to review the words just learned (e.g. Tier 2 words, antennae; segments, camouflage; familiar, wings; outside), and to provide redundant information in a different medium.

On subsequent days, the teacher increasingly supports children's vocabulary learning using additional videos that focus on new words in- and outside the category, helping to build children's knowledge of the properties (e.g. insects have six legs and three body segments) that are related to the category. In addition, videos and teacher's questions deepen children's knowledge of the concept by providing information about the topic (e.g. insects live in a habitat that has the food, shelter, and weather they like). Following the video, the teacher uses the information book and picture cards to engage

[1] All clips have been specially selected from the archives of Sesame Street and Elmo's World. Clip length varies from 40 seconds to 1 ½ minutes.

children in sorting tasks, including words that are not clearly in or out of the category (e.g. is a bat an insect?), challenging children by giving them problems to solve, such as "Time for a challenge." Lastly, the children review their learning through journal writing activities that involve developmental (phonic) writing.

The 8-day instructional sequence is designed to help teachers scaffold children's learning. In the beginning, for example, the teacher lesson plan focuses on explicit instruction, helping children to 'get set'—providing background information—and "give meaning," to deepen their understanding of the topic. As the instructional sequence progresses, the teacher begins to "build bridges" to what children have already learned and what they will learn (establishing inter-textual linkages across media). Here, the teacher begins to release more control to the children. Finally, the teacher is encouraged to "step back" giving children more opportunities for open-ended discussion. At the end of the instructional sequence, children are given a "take-home" book—a printable version of the information book used in the lesson. Throughout the sequence, familiar words are used for helping children talk about a topic, and for incorporating the approximately 5-7 Tier 2 new words for each topic into more known contexts. All eight topics follow a similar instructional design format.

Overview of the Methodological Design and Research Questions

Often seen as a test-bed of innovation (Cobb et al., 2003), design research has several common features: It is interventionist, iterative, and takes place in naturalistic contents. Conceptualized by Ann Brown (1992) and Allan Collins (1992), it was originally introduced with the expectation that researchers would systemically adjust various aspects of a design so that each adjustment served as a type of experimentation, allowing researchers to test and generate theory in naturalistic contests (Barab & Squire, 2004).

In this research, we systematically focused on two facets of our pedagogical design for teaching vocabulary: word selection and word organization. Specifically, our goal was adjust our instructional design to promote greater word learning in categorical structures that would enable children to inference and make generalizations. We describe each pedagogical issue in greater depth below.

<u>Word Selection.</u>

Surprisingly little research has focused on word selection in vocabulary training (Beck & McKeown, 2007). More often than not, curriculum builders have selected words subjectively based on what might be considered unfamiliar to children, or opportunistically based on the existing instructional materials in hand. Recently, however, there have been two approaches proposed for word selection. Beck and her colleagues (2002), for example, has argued that words for vocabulary instruction should be selected from the portion of word stock that comprises high utility sophisticated words (Tier 2) that are characteristic of written language (e.g. commotion for noisy). Biemiller (2006), on the other hand, has argued for greater breadth of word knowledge, focusing instruction on words children will learn more readily--words that constitute between 40-70% of a target student groups' knowledge--because the greatest gains can be made on these words. Although both approaches have been independently examined (Beck, McKeown & Kucan, 2002; Biemiller & Boote, 2006), the relationship of the degree of word selection, and its effects on retention has not been explored.

To examine this issue, we selected words in each topic based on content areas (e.g. science and health) designated in early learning standards. Of these words, 5-7 were considered Tier II words, or academically sophisticated words, and 10 partially familiar words. These words were analyzed for difficulty level using the lexical norming sample of the MacArthur-Bates Communicative Developmental Inventories (MCDI) (Dale & Fenson, 1996) to identify words considered known (‘acquired) by normally-developing 3-year-olds. This database is a set of parent-report inventories of child language and communication designed to yield information on the course of language development within a population. The MCDI has strong concurrent and predictive associations with other measures of vocabulary, language and cognitive development.

We also used a set of corpora from the CHILDES database (MacWhinney, 2000), a database consisting of transcriptions of adult-child spoken interactions in different home and lab settings around the world. We selected a combination of English-American corpora focusing on young children under 5-years of age from a variety of socio-economic backgrounds ranging from high-risk families to professional families. From this, we created a norming database of word difficulty.

We selected approximately equal proportions of familiar and unfamiliar words (based on the above corpora), with 56% of the primary words considered “unfamiliar” to preschoolers in Unit 1, and 54% in Unit 2. As part of our iterative design, our goal was look at the types of words learned immediately following the topics in the first Unit, and then to adjust our instructional design if necessary in the second Unit based on the pattern of learning we found.

Word Organization

The organization of knowledge is a central feature of cognitive ability in early and later learning and conditions of instruction have been shown to significantly influence the kinds of knowledge structures that are acquired (Gelman & Kalish, 2004). Since category development in an important factor in word learning (Bloom, 2000), our instructional design proposed that learning words like antennae, segments and wings, all taught as properties of a particular topic, could help children develop both a mental representation and a strategy for efficiently storing new information.

To examine this instructional design feature, we semantically grouped words in each topic that helped to explain the properties of a topic; we then also included words that helped teachers talk about the topic. For example, semantically grouped words in exercise included muscles, bones, heart, lungs, and intestines. Following the topics in Unit 1, children were asked to sort words into categories. Based on what we found, our goal was to adjust the instructional design to promote inference and generalization.

Therefore, the purpose of the design experiment was to research, test, and iteratively derive principles of word learning and word organization that could help to engineer a learning process, and theoretically advance our understanding of vocabulary development for low-income preschoolers. Specifically, we asked the following questions: 1) How might word difficulty influence knowledge and retention of words? 2) Could learning words in categories support vocabulary and inference generation? 3) How might treatment children differ from a comparison group who did not receive the intervention?

THE DESIGN EXPERIMENT: PHASE I

Method

Research sites and participants

This study was conducted in two elementary school Head Start programs. The programs were selected on the basis of proximity, initial enthusiasm for and commitment to the project, and the population served specifically low-income preschoolers. Three teachers participated in each site in the treatment; all had A.M. and P.M. classrooms.

Design experiments do not usually require comparison or control classrooms as in conventional experiments. However, to better determine the intervention's potential to enhance vocabulary, we selected six additional classrooms from the same sites to serve as a comparison group. Demographic characteristics of the sample are provided in Table 1. All teachers were female; one was African-American, the others, Caucasian. All had a bachelor's degree and considerable teaching experience (ranging from 4 to 20 years). Class size ranged from 15-18 children in each session (Treatment=89; Comparison=89), for a total sample size of 178 children. Prior to the start of the study, we individually administered the Peabody Picture Vocabulary Test (PPVT) to examine children's receptive vocabulary in treatment and comparison classrooms. T-tests revealed no significant differences between groups, $F(1, 176) = .78$, p=n.s.

The Early Language and Literacy Observation (ELLCO) (Smith & Dickinson, 2002) was conducted to examine the quality of the literacy supports in the environment in classrooms. Two trained research assistants observed classrooms for approximately 1 ½ hours prior to the intervention. Inter-rater reliability was .90. As shown in Table 1, there were no significant differences in classrooms quality.

Insert Table 1 about here

Approach to Analysis

The first phase of our design experiment was conducted in collaboration with the six teachers in the treatment group, four graduate research assistants and a project director. Throughout the experiment, we viewed our relationship as a shared learning enterprise. As part of the process of refinements, we held monthly debriefing sessions in which we shared and interpreted past events, and discussed and planned for prospective events.

Our research team visited classrooms twice each week, generating a comprehensive record of activities in progress throughout the experiment. Using the lesson plans as our guide, we documented the evolving issues to be discussed within our research team as supporting or questioning our conjectures, and then shared these with teachers. Teachers did the same, keeping detailed diaries of their activities throughout the project. We also provided teachers with video cameras to document examples of children's activities and their evolving conjectures.

In the first phase of the study we were interested in how WOW worked exclusively for the treatment group. We devised three sets of instruments to determine how the instructional design of WOW might influence word selection and word organization.

Expressive vocabulary. To examine the extent to which children learned instructed words, we developed an expressive vocabulary test. Five words were randomly selected from each of the instructional topics. Children were shown picture cards for each of the 35 words from a single randomized set and were asked to name each picture. Scores represented the number of cards correctly identified out of 35. Cronbach's alpha indicated acceptable reliability (alpha= .80).

To examine children's immediate recall, 10 words, five 'easy' (as identified by Biemiller's criteria) and five harder words (as identified by Beck's criteria) were randomly selected and assessed after topics. These "end of topic" expressive measures were similar in format to the pre-and post assessment, and were designed to measure immediate retention for children in the treatment group.

Word Organization. The second task was known as "Picky Peter," was designed to tap a) growth in conceptual/categorical knowledge; and b) the use of categorical knowledge to bootstrap children's learning of unfamiliar words. Adapted from a puppet task used by Waxman and Gelman (1986), children were shown a puppet, **Peter**, and were told, "**Peter** is picky and only likes insects" (a specific category label). Children were then shown 20 items (some of the items were category exemplars; others did not belong to the category) and were asked to help **Peter** find the items he would like and to justify their choices. Ten of the items were from the curriculum, and ten items were novel, and not taught in the curriculum. Items were depicted on picture cards with some contextual background (e.g. a raccoon in the woods) because children were expected to infer category membership based on cues from the item. Scores were calculated independently for taught items and not-taught items, allowing us to examine what was learned from the curriculum, and what could be inferred from the curriculum.

Children were asked "What is this?" If the child was unable to name the referent, the assessor provided the label for the child. Then, the assessor asked, "Is it a wild animal? And "Why do you think a…is/is not a wild animal? Children's justifications were recorded and transcribed. Cronbach's alpha indicated acceptable reliability (α = .89).

Tell me. The "Tell Me" task was an open-ended measure designed to examine children's use of words and properties of categories, prior to and after the unit. Four picture cards, depicting a category (such as pets) for each topic were shown to child. The assessor said, "I'm going to show you a picture. Tell me what you know about it." Children were given one minute to talk about the picture. These conversations were audio-taped and transcribed verbatim. Children received a point for each correct word label, and one point for each property in the categories (α=.85).

Procedures

We introduced WOW to teachers in the winter term through a day-long workshop that explained the approach and the instructional design behind its development. Materials were provided including: DVD player, DVD with video clips, information books, picture cards, instructional guides for each of the topics. Teachers agreed to use the supplementary curriculum during the whole group circle time for the 10-12 minute instructional period each day.

A primary goal for our design experiment was to improve the initial design by assessing and revising our conjectures informed by our analysis of both children's gains in word knowledge, and concepts. We used an iterative process that included three sources of data: information from teachers' feedback; observations of the enactment of the curriculum, and children's assessment scores.

Prior to the beginning of the study, pretests were individually administered to children. We then began our iterative cycle: before each topic, pretests individually assessed children's productive labeling and understanding of the properties of the categories; following each topic, they received an end of topic word assessment. Following the unit, they received a post-Tell me assessment, and a Picky Peter task. We reviewed observations among the research team, and met with teachers bi-weekly.

During the enactment of the topic, we conducted weekly observations specifically focused on the alignment between the written lesson plan and its enactment. Our efforts were not designed to examine fidelity to the lesson plan, in particular; rather, we wanted to learn how teachers might use, or adapt the lesson to meet children's needs. Using the lesson plan as our guide, we took notes on the implementation of the lesson; what seemed to engage children's attention; and when interest appeared to fall off. We reviewed these notes weekly, comparing our observations across settings.

Phase 1 of this design experiment concluded with a series of focus group discussions to review and debrief with teachers regarding their experiences with WOW. We then began to analyze the quantitative evidence, focusing specifically on our theoretical conjectures about word learning and concept development.

Results: Phase I

<u>Word selection and word retention</u>

Our first analysis was designed to examine which words were most likely learned and retained. Using our norming criteria, pretest words were grouped into two categories reflecting Biemiller's criteria (e.g. partially familiar or "easy"), and Beck's (e.g. sophisticated or "hard"). Shown in Table 2, the pretest average percent correct for easy words exceeded the upper range of Biemiller's criteria (e.g. 70%), with children demonstrating knowledge of 78% of these words. In contrast, only 29% of the hard words were correctly identified, indicating that the word selection was, in fact, more difficult. Immediately following instruction, greater gains were made in learning hard words. More than 28% of the children remembered the meaning of these words compared to an additional 9% for easy words.

At posttest, however, the trajectory changed: children retained more of the easy words than the hard. There was a drop-off of 4% of the easy words compared to 11% of the hard. These results portray both the growth potential and the retention problem for hard words. Put simply, children appeared to learn them and lose them more frequently than easy words.

Insert Table 2

Word Organization

The next analysis was designed to examine the degree to which children used categorical knowledge for word learning and inference generation. Using the "Tell Me" task, we coded the average number of words used per category to determine how or if words were incorporated in their open-ended responses to a contextually-based picture.

As shown in Table 3, prior to instruction, the average number of words produced per category was minimal; on average, children used less than three words related to the category prior to instruction.

However, following instruction, post-tests scores showed little improvement. In fact, scores were essentially flat. A similar pattern was evident in the average number of properties children produced prior to and following instruction. On average, children could name less than one property of a category prior to instruction; following instruction, although some growth was recorded, scores showed little improvement. These results seemed to suggest that the instructional design did not promote the organizational prosthetic that could help children accelerate word learning and inference generation.

Insert Table 3 about here

At the same time, results from the Picky Peter task seemed to disconfirm this finding, or at least call it into question (see Table 3). Unlike the "Tell Me" task, the Picky Peter measure did not require expressive language; rather, children responded by physically sorting or pointing to a particular category. On the basis of this task, the results seemed to indicate that children were able to use categories to make inferences. They correctly sorted more than 7 of the 10 words taught in the appropriate category indicating that they had learned what had been taught. But they also sorted more than 7 of the 10 words that had not been specifically taught. These results appeared to suggest that children used their understanding of the properties of categories to infer category membership for new words. In other words, categories seemed to serve as a strategy for inference generation.

However, sorting activities are prone to guessing. Given these conflicting results, it was evident that our assessment techniques needed further refinement to examine children's use of categories as a bootstrap for word knowledge. Further, teachers' feedback and observations indicated that we had under-specified the properties of categories and the relation to the words children were learning. We decided to address both of these issues of assessment and instructional design in Phase II of our design experiment.

Therefore, in preparation for Phase II we developed a series of adaptations to the instructional design to both exploit the opportunities provided in the WOW instruction, and to meet its challenges. The purpose of the Phase II research was to examine the outcomes of these adaptations, extending our understanding of the theoretical premise underlying the instructional design.

THE DESIGN EXPERIMENT: PHASE II

Adaptations to the Instructional Design

Word selection: Given evidence of the differential retention rates of easy and hard words, several adaptations were made to the instructional design. From our observations and teacher reports, children appeared to enjoy and use many of the hard words taught throughout the lessons. Teachers' feedback suggested that they liked the complexity and the sounds of many of the words like 'camouflage' and 'habitat." Consequently, we did not believe it was due their lack of enthusiasm for learning hard words. Rather, we conjectured that hard words, defined by their nature as outside children's existing lexicon, might need additional practice and review. Previous research by Beck and her colleagues (2002) and Nagy, Anderson, and Herman (1987), for example, had shown that the frequency of exposure to words was tied to word growth. Children needed to hear, say, and practice these words more frequently and in different contexts.

Therefore, we made two design changes to lessons in the second phase of our design experiment. The first was to include a review of hard words. Together with teachers we developed a strategy that would involve them in quick call- and response questions designed to elicit each of the difficult words. This strategy could allow for a brief, but frequent review of words, keeping the daily lesson within the 10-12 minute time constraint. For example, in the insect topic, the teacher might ask: What body part do insects use to feel? Where do insect live most of the time? How do moths hide from their enemies? Teachers would repeat the responses and cue children to do so as well.

The second design change was to include more review and practice across topics. Teachers believed that a process of continual review, essentially building meaning of these words throughout the topics would be useful. Since topics within units were tied to super-ordinate categories, the repetition and practice of words would allow teachers to extend children's understanding, build bridges between topics, and provide opportunities for greater frequency of word use as well as greater depth of understanding. We decided jointly that the "time for a challenge" activity, asking children why words either belong in or out of category" could prove to be an optimal time to practice hard words in different contexts. Together, these adaptations would provide teachers with approximately a 20% increase in opportunity to review and practice hard words more frequently.

Word Organization. Given the conflicting results about the potential power of taxonomic categories to scaffold word learning and inference generation, we made several adaptations to the instructional design. Further, to better understand children's use of categories in word learning we made an additional adaptation to our assessment approach.

The first instructional design change was to highlight the properties that were integral to each topic. For example, with insects, we included properties such as: insects always have six legs; they have three different parts or segments; they have antennae that they use to smell and feel things, and they most often live outdoors. These properties had been integrated throughout the teacher-child interactions in lessons; however, now we placed them on the front page of the instructional guide in a special box to call attention to their importance.

The second design change was to place a special column on the left-hand side throughout the lesson plan to give teachers the rationale for each phase of the activity and its importance for conceptual development. For example: "This part of the lesson links the category-related words and provides them with additional information and additional vocabulary words that they can use to describe and explain a category." By making these design changes our goal was to help teachers focus on the 'big ideas,' or concepts to essentially prioritize certain aspects of the lesson. In other words, we wanted to ensure that teachers recognized that word learning and word organization worked hand-in-hand,

and that the words were semantically clustered. These changes were discussed and demonstrated in a workshop with teachers prior to the beginning of the topics in Unit 2.

To examine how these changes might affect children's ability to make inferences, we also made an adjustment to the Picky Peter task. First, we went back to our first set of lessons and qualitatively coded children's justifications for making sorting decisions. For example, children received a point for each justification that included properties common to the category; a total content score was calculated for each child, and averaged across classrooms (inter-rater reliability was .95).

Second, using words not taught, we sought to ask children to justify their responses to words in categories. According to the MCDI index, each of these words would be outside of children's average working vocabulary. Our goal was to understand how categorical instruction might support children's ability to talk about a topic compared to those who had not.

Together, these adaptations in word selection and word organization were designed to better our understanding of whether our instructional design could enhance, and potentially accelerate vocabulary development.

Participants

In Phase II of the study, we examined growth in word knowledge and categorical learning for both treatment and comparison classrooms.

Procedures

We revised lessons for the treatment group according to the adaptations described. Materials were distributed to teachers, and similar procedures were followed. Before the intervention, treatment and comparison children were assessed on WOW expressive language assessment, the Tell Me and the Picky Peter task. However, only treatment children were assessed on word knowledge following each topic to examine immediate gains in word knowledge as a result of the intervention. In addition, we continued to collect observational data during this second phase.

Results: Phase II

Word Selection. Our first analysis was to examine differences between treatment and comparison groups on word knowledge. Given that words were curriculum specific, it was not surprising to find that there were significant differences between groups. Analysis of covariance, with pretest as covariate indicated that treatment children scored significantly higher on the expressive language assessment than those children in the comparison group, $F(2, 175)=16.68$, $p < .001$; further these gains were educationally meaningful (Cohen's *d*= .64) as demonstrated by the strong effect size.

Next, we analyzed word growth specifically for the treatment group after the curriculum adaptations were put in place. Table 4 describes the differences in children's word knowledge at pretest. As in the first Phase, differences in word knowledge for easy words compared to hard words are stark. Children knew almost double the number of easy words compared to hard words. Similar to the first Phase, following instruction, children made greater growth for hard words than easy. Average percent growth from pretest to end of unit for hard words was 22% compared to easy words at 8%.

Insert Table 4

Retention of word knowledge, instead of declining as in Phase I, however, increased for both easy and hard words. Average growth in knowledge actually continued to increase: For easy words, increases were modest, but for hard words, more substantial.

Lastly, we looked at a comparison of gains made in Phase I compared to Phase II. Here we found that right after instruction, children seemed to have learned a similar proportion of hard words in both Phase I and II. However, at posttest, children actually continued to gain word knowledge; this was true for easy words as well, albeit less pronounced than with hard words.

Figure 1 about here

These results indicated that initially in both units, hard words were predictably known by a smaller proportion of children at pretest than easy words. However, in both Phases, greater growth was recorded for hard words. With revision and added frequency of exposure to hard words, a greater proportion of children retained their knowledge of these words in Phase II. Further, it seemed that the cross-topic reviews might have demonstrated greater utility for retention than immediate review. These results suggest that instructional design features of review and practice in different contexts enhanced word knowledge. It also showed that children were capable of learning and retaining hard words.

Word Organization. Next we examined differences between treatment and comparison children's ability to use words and properties to describe pictures (e.g. "Tell

Me), and their ability to identify words in categories. Shown in Table 5, the differences between groups were significant and educationally meaningful, $F(2, 175)=26.46$, $p < .001$. Children in the treatment group used words and properties of categories in their descriptions; further, they were significantly more likely to sort words in appropriate categories, $F(2, 175)=45.13$, $p < .001$.

Insert Table 5 about here

We then conducted a more stringent analysis of children's ability to use categories. We compared treatment and comparison children's justifications, focusing particularly on words that were not taught in the curriculum. For example, shown a picture card of an insect--in this case a spider, a word that had not been taught--a child was asked "how do you know that it is not an insect?" To receive points, a child would have to provide a justification that included a property of the category, such as "because it doesn't have six legs."

Our analyses indicated that treatment children were better able to correctly justify their inferences than children in the comparison group. In short, they were able to talk about why they made their choices in ways that demonstrated an understanding of how these words semantically clustered. At the same time, they also used these properties to describe the exclusion of certain words and terms for a particular category, demonstrating cognitive flexibility across tasks and topics.

Insert Figure 2 about here

Table 6 provides several examples of the differences between groups in justifying their choices. In response to the question, “Is a heel a part of the body?” a child who had received instruction reported, “Yes, because it helps you walk” while a comparison child not receiving instruction said, “cause.” Similarly, treatment children were able to apply their categorical information, suggesting that they were using the semantic information about categories to make inferences and generalizations. Categories, therefore, appeared to give children a way to organize words which became helpful for learning new words. In contrast, children who did not have such information often searched for a rationale that was most immediate to them.

Insert Table 6 about here

Taken together, these results suggest that the design enhancements appeared to enable children to identify common properties associated with categories, and to use this information to make inferences beyond what was taught in the curriculum. By increasing the explicitness of category membership, lessons appeared to better support children’s reasoning and knowledge of the concepts words represented. Further, it seemed like taxonomic knowledge acted like a bootstrap for making inferences.

Discussion

Children’s vocabularies play an enormously important role in their lives and future possibilities (Beck & McKeown, 2007). A large and rich vocabulary is strongly associated with reading proficiency (National Reading Panel, 2000). Recent cognitive models of reading (Cunningham & Stanovich, 1997; Storch & Whitehurst, 2002) have demonstrated that facility in vocabulary makes a critical contribution to comprehension.

Nevertheless, starting as early as two years old (Halle et al., 2009) there are profound differences in vocabulary knowledge among learners from different socioeconomic groups. Particularly disheartening is the finding that once established, differences in vocabulary knowledge remain throughout schooling (Cunningham & Stanovich, 1997). Consequently, there is an emerging consensus that early intervention is critically needed if we are to substantially improve children's achievement and begin to close the gap on reading performance.

Given that more intensive instruction is needed to increase vocabulary, decisions must be made about the kinds of activities that can potentially accelerate its acquisition. Selecting which words to teach may seem like a primary issue; however it has received strikingly little attention in curriculum—especially at the preschool level (Authors, 2009). In fact, Coyne and his colleagues (2007) point out that, although knowledge about how to teach vocabulary is accumulating, what to teach remains elusive.

We used the iterative process of a design experiment to extend our theoretical understanding of word selection and word organization. Both Biemiller (2006) and Beck and colleagues (2002) have proposed a heuristic for word selection, though neither has established a normative definition. Using the extant data bases of MacArthur-Bates and the CHILDES, we established a normative estimate of each heuristic for words in our curriculum, and subjected these to systematic analyses.

Initial evidence suggested that while greater growth occurred for hard words, these words were less likely to be retained. In this respect, it seemed to provide support for Biemiller's position that words should be selected from the portion of word stock that was partially familiar. However, following our instructional design changes in Phase II which involved additional review and practice, growth and retention for hard words substantially improved. Children learned more difficult words and retained them at a higher rate than easy words. These differences were significant when we compared growth in word knowledge between treatment and comparison groups.

Given that instructional time is precious, these results suggest that it may be most facilitative to teach hard words-- if sufficient practice and review are provided. These are words that are not only characteristic of written language, they are critical to content learning. Our words, for example, were selected on the basis of content standards regarded by the Fordham Foundation as exemplary (Finn et al., 2006). As Beck and her colleagues have argued (2007), it is precisely these words comprising sophisticated words of high utility in content areas that are that are least likely to be learned outside of school.

The iterative process of the design experiment also allowed us to make conjectures about the role of categorical learning and word knowledge. Studies of early language acquisition (Gopnik & Meltzoff, 1987; Huttenlocher et al., 1991) have shown a simultaneous growth in the ability to categorize and vocabulary acquisition. It has been suggested that these two phenomena, the ability to learn new words and knowledge of categories may be related in a synergistic fashion. Borovsky and Elman (2004), for example, recently tested this assumption through computational simulations, finding in each that improvements in category structure were tightly correlated with subsequent improvements in word learning ability. We proposed in this design experiment that by teaching words in semantic clusters along with an articulated set of categorical properties, we could potentially improve children's word learning and conceptual development.

The first test of our theory found equivocal results; the expressive task indicated little improvement on word labeling or identifying properties associated with categories. At the same, children's ability to sort words in categories, both taught and not taught seemed to support our thesis. By making the properties of categories more explicit and central to the instructional design, and by seeking children's justifications for sorting we attempted to learn more about their thinking process in the assessment selection.

Based on these changes, results of the category tasks indicated that treatment children were able to slot familiar words into appropriate categories, and provide a sound rationale for why they were doing so. Further, their knowledge of categories enabled them to better slot words that were not taught into appropriate conceptual groupings. Examining their justifications, these differences appeared especially striking when compared with other children who had not had such training. These findings suggest that teaching words in categories may represent an important instructional design scaffold for efficiently and economically storing vocabulary. It might also provide greater capacity to attain new information. Schema theorists (Anderson & Pearson, 1984; Rumelhart, 1980), for example, have argued that such frameworks act as a kind organizational prosthetic, serving to diminish information processing load. Given the stark differences in word knowledge between middle-income and low-income children (Hart & Risley, 2003), these results could suggest a significant pathway for accelerating vocabulary development.

The decision to use a design experiment was shaped by the daunting nature of challenges in curriculum development. Too often, research has merely compared the effectiveness of one instructional program against another (Reinking & Bradley, 2008). In such experiments, researchers work to control the influence of design factors rather than to understand them. In contrast, in this research we worked from a strong theoretical foundation to guide our iterative process toward the goal of improving vocabulary development for low-income children. Our efforts were designed to expand and deepen knowledge of instructional practice and how instructional design features can be changed to reach the targeted outcomes. Therefore, our goal was to focus on two dimensions of information: the outcomes of new instructional practice and what instructional features are required to engage in that practice. Knowledge along both dimensions is critical if we are to create evidence-based instructional materials.

References

Anderson, R. C., & Pearson, P. D. (1984). A schema-theoretic view of basic processes in reading comprehension. In P. D. Pearson (Ed.), *Handbook of reading research* (pp. 255-291). NY: Longmann.

Authors. (2009).

Authors (2007).

Authors. (2007b).

Barab, S. & Squire, K. (2004). Design-based research: Putting a stake in the ground. *Journal of the Learning Sciences*, *13*, 1-14.

Beck, I., & McKeown, M. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *Elementary School Journal*, *107*, 251-271.

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

Biemiller, A. (2005). Addressing developmental patterns in vocabulary: Implications for choosing words for primary grade vocabulary instruction. In E. H. Hiebert & M. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 223-242). Hillsdale, NJ: Erlbaum.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In Authors, 2006.

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

Blair, R. C., Higgins, J., Topping, M., & Mortimer, A. (1983). An investigation of the robustness of the *t* test to unit of analysis violations. *Educational and Psychological Measurement, 43*, 69-80.

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.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2,* 141-178.

Chall, J., Jacobs, V., & Baldwin, L. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.

Clements, D. (2007). Curriculum research: Toward a framework for "research-based curricula." *Journal for Research in Mathematics Education, 38*, 35-70.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*, 9-13.

Cohen, D., Raudenbush, S., & Ball, D. (2003). Resources, instruction and research. Educational Evaluation and Policy Analysis, 25, 119-142.

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15-22). New York: Springer-Verlag.

Coyne, M., McCoach, B., & Kapp, S. (2007). Vocabulary intervention for kindergarten students: comparing extended instruction and incidental exposure. *Learning Disabilities Quarterly, 30*, 74-88.

Coyne, M., Simmons, D., Kame'enui, E., & Stoolmiller, M. (2004). Teaching vocabulary during shared storybook readings: An examination of differential effects. *Exceptionality, 12*(3), 145-162.

Cunningham, A. E., & Stanovich, K. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934-945.

Dale, P., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28*, 125-127.

Dickinson, D., McCabe, A., & Essex, M. (2006). A window of opportunity we must open to all: The case for preschool with high quality support for language and literacy. In Authors, 2006.

Volume II (pp. 11-28). New York: Guilford Press.

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

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

Elley, W. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*, 375-411.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: differences by class and race. *Social Science Research, 33*, 464-497.

Finn, C., Petrilli, M. & Julian, L. (2006). *The state of state standards*. Washington, D.C.: Fordham Foundation.

Gopnik, A. & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development, 58*, 1523-31.

Gelman, S., & Coley, J. (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., Coley, J., Rosengren, K., 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*.

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

Gelman, S., & O'Reilly, A. (1988). Children's inductive inferences within superordinate categories: The role of language and category structure. *Child Development, 59*, 876-887.

Halle, T., Forry, N., Hair, E., Perper, K. Wander, L., Wessel, J. et al. (2009). Disparities in early learning and development: Lessons from the Early Childhood Longitudinal Study-Birth Cohort. Washington, D.C.: Child Trends.

Hargrave, A., & Senechal, M. (2000). A book reading intervention with preschool children who have limited vocabularies: The benefits of regular reading and dialogic reading. *Early Childhood Research Quarterly, 15*, 75-90.

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

Hart, B., & Risley, T. (2003). The early catastrophe. *American Educator, 27*, 4,6-9.

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

Hohmann, M., & Weikart, D. (1995). *Educating young children*. Ypsilanti, MI: High/Scope Press.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology, 27,* 236-48.

Juel, C., Biancarosa, G., Coker, D., & Deffes, R. (2003). Walking with Rosie: A cautional tale of early reading instruction. Educational Leadership, 52, 12-18.

Kalish, C., & Gelman, S. (1992). On wooden pillows: Multiple classification and children's category-based inductions. *Child Development, 63*, 1536-1557.

Kaplan, A., & Murphy, G. L. (2000). Category learning with minimal prior knowledge. *Journal of Experimental Psychology-Learning Memory and Cognition, 26*, 829-846.

Kibby, M. (1995). The organization and teaching of things and the words that signify them. *Journal of Adolescent and Adult Literacy, 39*(3), 208-223.

Moats, L. (1999). *Teaching reading IS rocket science*. Washington, D.C.: American Federation of Teachers.

Morrow, L. M. (1988). Young children's responses to one-to-one readings in school settings. *Reading Research Quarterly, 23*, 89-107.

Murphy, G. L., & Lassaline, M. E. (1997). Hierarchial structure in concepts and the basic level of categorization. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 93-131). East Sussex, UK: Psychology Press.

Nagy, W., Anderson, R. & Herman (1987). Learning word meanings from context during normal reading. *American Educational Research Journal, 24*, 237-270.

Pan, B., Rowe, M., Spier, E., & Tamis-Lemonda. (2004). Measuring productive vocabulary of toddlers in low-income families: concurrent and predictive validity of three sources of data. *Journal of Child Language, 31,* 587-608.

Penno, J., Wilkinson, A., & Moore, D. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 86*, 139-153.

Rehder, B., & Hastie, R. (2004). Category coherence and category-based property induction. *Cognition, 91*, 113-153.

Reinking, D., & Bradley, B. (2008). *Formative and design experiments*. New York: Teachers College Press.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R.J. Spiro, B. C. Bruce & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 34-58). Hillsdale, NJ: Erlbaum.

Senechal, M., LeFevre, J., Thomas, E., & Daley, K. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly, 33*, 96-116.

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

Smiley, S. & Brown, A. (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, E. E. (1995). Concepts and categorization. In E. Smith & D. Osherson (Eds.), *An invitation to cognitive science* (2nd Edition: Thinking ed., pp. 3-33). Cambridge, MA: MIT Press.

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

Storch, S. & Whitehurst, G. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934-947.

Vellutino, F., Scanlon, D., Sipay, E., Small, S., Pratt, A., Chen, R., et al. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes for specific reading disability. *Journal of Educational Psychology, 88,* 601-638.

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

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

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

Whitehurst, G. J., Epstein, J., Angell, A., Payne, A., Crone, D., & Fischel, J. (1994). Outcomes of an emergent literacy intervention in Head Start. *Journal of*

Educational Psychology, *86*, 542-555.

Whitehurst, G. J., Falco, F., Lonigan, J., Fischel, J., DeBaryshe, B., Valdez-Menchaca, M., et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology*, *24*, 552-559.

Whitmore, J., Shore, W., & Smith, P. (2004). Partial knowledge of word meanings: Thematic and taxonomic representations. *Journal of Psycholinguistic research*, *33*, 137-164.

Ziori, E., & Dienes, Z. (2008). How does prior knowledge affect implicit and explicit concept learning? *Quarterly Journal of Experimental Psychology*, *61*, 601-62.

Table 1. Demographic Characteristics of Sample

Characteristic	Treatment (N=6)	Comparison (N=6)
Average age	41	34
Number of years teaching	10	8
Education		
B.A.	100%	100%
Ethnicity		
Caucasian	100%	83%
African American		17%
ELLCO (124 possible)	83	70
Children in classroom		
Average age (in months)	51	50
Ethnicity		
Caucasian	63%	49%
African American	25%	30%
Middle-eastern	12%	21%
PPVT	87.29	87.28

Table 2. Percent of Words Correctly Identified by Treatment Group in Phase I

Source	**Easy**	**Hard**
Pretest Average % Correct	78%	29%
End of Topic % Correct	87%	57%
Posttest Average % Correct	83%	46%
Average % Growth Pre-to-End of Topic	9%	28%
Average Drop in Retention	-4%	-11%

Table 3. Knowledge of Categories and Properties by Treatment Group in Phase 1

Source	**Pretest**	**Posttest**
Average Number of Targeted Words Used Per Category	2.59 (S.D. 1.57)	2.76 (S.D. 1.63)
Average Properties Per Category	.23 (S.D. .48)	.54 S.D. (.64)
No. of correct sorts		Taught words: 8 Not taught: 8

Table 4. Percent of Words Correctly Identified by Treatment Group: Phase II

Source	**Easy**	**Hard**
Pretest Average % Correct	80%	42%
End of Topic % Correct	88%	64%
Posttest Average % Correct	92%	70%
Average % Growth Pre-to-End of Topic	8%	22%
Average Growth in Retention	4%	8%

Table 5. Differences in Word Knowledge, and Categories by Treatment and Comparison Group

Characteristic	Treatment	Comparison	ES
Expressive language (WOW)			
Pretest	16.11 (S.D. 5.20)	14.11 (S.D.4.92)	
Posttest	20.47 (S.D. 5.68)***	16.98 (S.D. 5.27)	.64
Word labels (Tell Me)			
Pretest	17.8 (S.D.8.94)	17.0 (S.D. 10.25)	
Posttest	20.29 (S.D.10.09)	18.65 (S.D.10.05)	.16
Word properties			
Pretest	1.32 (S.D.2.02)	1.04 (S.D. 1.49)	
Posttest	3.01 (S.D.2.39)***	1.29 (S.D. 1.64)	.84
Sorting			
Taught	7.29 (S.D. 1.12)***	5.90 (S.D. 1.22)	1.19
Not taught	7.46 (S.D. 1.11)***	6.34 (S.D. 1.22)	.99

***p < .001

Table 6. Examples of Justifications for Children's Category Selections

Topic	Key Concept/Properties	
Wild Animals	They live outside/away from people; live in different habitats (grassland; jungle or water; can't be pets.	
Word	**Treatment Group**	**Comparison Group**
Racoon (in category)		
	Q: How do you know a raccoon is a wild animal?	
	C: It lives in the forest	C: Cause
	C: Cause it lives with trees	C: Cause the arrow is pointing to it.
	C: Because it don't live with People	C: Cause it is
	C: It lives in the woods	
Fish (in bowl)	Q: How do you know this fish is not a wild animal?	
(not in cat.)	C: Cause it lives in a fish bowl	C: Cause
	C: It's a pet	C: It goes there
	C: Cause it could live with people	C: I don't know

Table 7 (Continued)

Topic	Key Concept/Properties	
Insects	Are very small creatures/animals; mostly live outside; have three body parts called segments; have six legs; have special ways to protect themselves from bigger animals; many insects have wings and fly.	
Word	**Treatment Group**	**Comparison Group**
Wasp (in category)	Q: How do you know a wasp is an insect?	
	C: It lives on trees	C: Cause
	C: Because it has legs, antennae, wings, mouth, eyes	C: I don't know
	C: Because it flies	C: Because it wants to
	C: Three body segments	C: It is.
Mouse (not in category)	Q: How do you know a mouse is not an insect?	
	C: It only has one body segments	C: Cause
	C: It doesn't have 6 legs it just only has four	C: Because him want to
	C: No, cause it don't got this many legs (holds up 6 fingers)	C: Because it's not
		C: I don't know

Figure 1. Learning and Retention of Words in Phase I and Phase II by Word Type

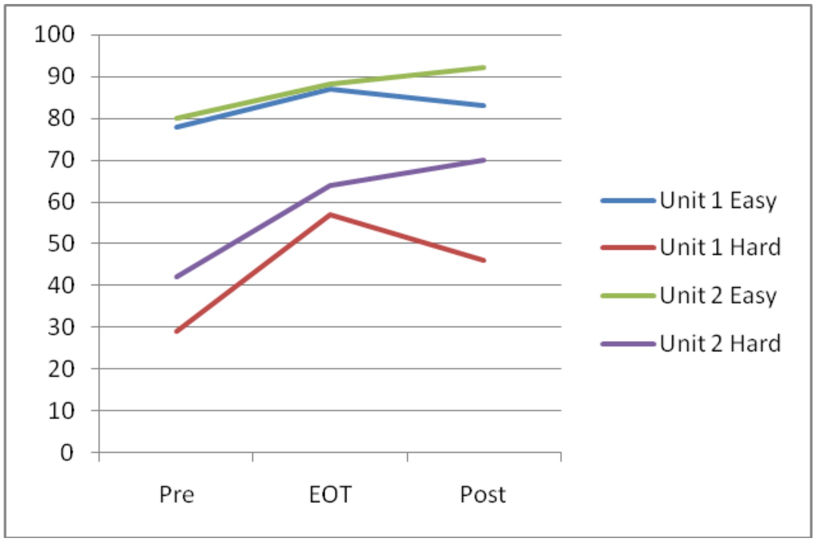

Figure 2. Differences between Treatment and Comparison Groups on Categorization