Journal of Educational Pyschology

An Effective Method for Building Meaning Vocabulary in Primary Grades

Andrew Biemiller and Catherine Boote

University of Toronto

Journal of Educational Psychology, 2006, 98, 44-62.

An Effective Method for Building Meaning Vocabulary in Primary Grades

Abstract

Teaching vocabulary to primary grade children is essential. Previous studies of teaching vocabulary (word meanings) using story books in the primary grades reported gains of 20% to 25% of word meanings taught. The present studies concern possible influences on word meaning acquisition during instruction (Study 1), and increasing the percentage and number of word meanings acquired (Study 2). Both studies were conducted in a working class school with approximately 50% English Language Learners. The regular classroom teachers worked with their whole classes in these studies. In Study 1, average gains of 12% of word meanings were obtained using repeated reading. Adding word explanations added a 10% gain for a total gain of 22%. Pretesting had no effect on gains. In Study 2, results showed learning of 41 % of word meanings taught. At this rate of learning word meanings taught, it would be possible for children to learn 400 word meanings a year if 1000 word meanings were taught. The feasibility of teaching vocabulary to primary grade children was discussed.

Key words: vocabulary, instruction, primary, repeated reading, pretest, transfer

An Effective Method for Building Vocabulary in Primary Grades

As many researchers have noted, vocabulary—knowledge of word meanings--is a powerful predictor of reading comprehension. Current reading instruction is apparently premised on the view that children can build the vocabulary they need *after* learning to read (decode) fluently, as little or no vocabulary instruction occurs during the primary grades (National Reading Panel, 2000). However, as Biemiller (2005) showed, by the end of grade 2 (i.e., by the end of the primary years) while average children have acquired around 6000 root word meanings, children in the lowest quartile acquired around 4000 root words while those in the highest quartile acquired around 8000 root words. After grade 2, average children add 1000 word meanings per year. A gap of 2000 root word meanings is roughly equal to grade 2 grade levels. This gap persists throughout the elementary years. As outlined in a recent issue of *American Educator* (2003), by 4th grade many children experience a "slump" in reading comprehension caused by below-grade vocabulary levels (Becker, 1977; Chall & Jacobs, 2003; Chall, Jacobs, & Baldwin, 1990). Until schools are prepared to emphasize vocabulary acquisition, especially in the primary grades, less advantaged children will continue to be handicapped even if they master reading written words.

Recent research has also shown that vocabulary and language skills are to a considerable degree quite separate from skills leading to word identification skills. Both Storch and Whitehurst (2002) and Dickinson, McCabe, Anastasopoulos, Peisner-Feinberg, & Poe, (2003) have shown that 2 quite different groups of variables affect reading acquisition during the pre-primary and primary years: decoding skills and vocabulary.

Thus there is clearly a need to address individual differences in vocabulary in the primary grades. During this period, differences in vocabulary size grow larger (Biemiller and Slonim,

2001). While it is true that large language and vocabulary differences develop before kindergarten (Hart & Risley, 1995, 1999, 2003), current school practices allow further widening of vocabulary gaps during the primary years. The recent American "No Child Left Behind" act specifies attention to vocabulary in the primary grades, using "scientifically established" teaching methods. Unfortunately, unlike work on decoding and spelling, *there is no established method of teaching vocabulary in the primary grades* (National Reading Panel, 2000).

Becker (1977) suggested that the school emphasis on reading skills (word identification) in the early grades usually involves little challenging vocabulary. This is appropriate for reading instruction, but provides little opportunity build vocabulary. Low vocabulary results in problems for many middle elementary children's reading comprehension. Those with restricted oral vocabularies comprehend at lower levels. Other studies have shown that: (a) developed vocabulary size in kindergarten is an effective predictor of reading comprehension in the middle elementary years (Scarborough, 1998, 2001); (b) orally-tested vocabulary at the end of grade 1 is a significant predictor of reading comprehension 10 years later (Cunningham & Stanovich, 1997); and (c) children with restricted vocabulary by third grade have declining comprehension scores in the later elementary years (Chall, Jacobs, & Baldwin, 1990). In each of these studies, observed differences in vocabulary were related to later comprehension. None of these studies had any evidence that schooling was responsible for vocabulary size.

Concurrent correlations between vocabulary and reading comprehension are high. When a reading vocabulary/language test is correlated with a reading comprehension test, correlations are usually over $r = .80$ (Bloom, 1976). Teaching vocabulary can also affect reading comprehension. Studies summarized in National Reading Panel (2000) showed that teaching

text-specific vocabulary improved comprehension of texts (Carney, Anderson, Blackburn, & Blessings, 1984; Medo & Ryder, 1993; Wixson, 1986). In 2 studies (the second being a replication), general vocabulary instruction resulted in significantly increased reading comprehension (Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson, & Perfetti, 1983),

Unfortunately, several studies suggest that at present, primary school attendance is not a major source of vocabulary acquisition. Age, not school experience, apparently affects vocabulary development. Cantalini (1987), Christian, Morrison, Frazier, & Massetti (2000), and Morrison, Dow, and Ehrensberger (1995) reported that, unlike early academic skills, vocabulary is affected by age but *not* by school experience in the primary years. Thus the average vocabulary of relatively "old" kindergarten children and "young" grade 1 children is similar. (They differ in school experience by a year, but in age by 1 or 2 months.) Similarly, the average vocabulary of "old" grade 1 children and "young" grade 2 children is also about the same.

Although most primary language programs make reference to word meaning instruction, in fact few word meanings are actually identified for instruction--far fewer than would be needed to significantly impact vocabulary development. In one of the author's direct experience with both public and Catholic schools (as a teacher educator), word meaning instruction was uncommon, and when done, far too few word meanings were discussed to have an impact on vocabulary development. During the current research, teachers have objected to devoting as much as half an hour a day to vocabulary instruction.

In short, vocabulary levels diverge greatly during the primary years, and virtually nothing effective is done about this in schools. It is true that some children arrive in kindergarten with less vocabulary than other children. Schools cannot change what happens before children start

school. However, when children continue to acquire less vocabulary while in primary school, it becomes less likely that they can later "catch up". Our chances of successfully addressing vocabulary differences in school are greatest in the preschool and early primary years.

During the primary years, average children add an average of at least 840 root word meanings per year. Children in the lowest quartile add an average of 570 meanings per year during the same period (Biemiller, 2005). To have a useful impact on vocabulary growth, an intervention would need to add several hundred root word meanings per year. This is considerably more meanings than are presently addressed in classroom programs.

Context and Vocabulary Instruction

In a review of vocabulary instruction with elementary school children, Stahl and Fairbanks (1987) reported that instruction of word meanings in context is more effective than no-context instruction of word meanings. All primary grade programs reported there and elsewhere appear to be context-based. Why is context useful for teaching vocabulary?

Consider Carey's (1978) hypothesis for learning a single new word meaning. She argues that young (children ages 2 to 6) "map" (quickly associate) new words with meanings illustrated in specific concrete task contexts. She has also demonstrated that syntactic cues are used in mapping word meaning. Other studies indicate that by age 4 or earlier, new words can also be mapped to verbal narrative context referents as well as concrete task context referents. For example, studies by Senechal and her colleagues illustrate the use of narrative-based referents with preschool children (e.g., Senechal, 1997; Hargrave & Senechal, 1995). After initial "mapping", word meanings are extended as the word is encountered in other contexts (Carey, 1978).

When an unknown word is encountered in an interesting narrative, the basis for creating a verbal "referent" or meaning exists. Brief explanations (often, 1 or 2 sentences) can be sufficient to establish what new words refer to. In a narrative that is being generally followed and understood, a word referring to a new object (e.g., *castle*), agent (e.g. *trainer*), action (e.g. *paddle*), modifier (e.g. *rapid*), or setting (e.g. *circus*) can be understood in the context. In some cases, the child can construct these meanings without any explanation, simply while hearing a story. For other cases, providing a short verbal explanation is often sufficient to establish a meaning to which the word can be "fast-mapped".

Part of the advantage of direct explanation may be simply drawing attention to the novel word. It is very difficult for a young listener to attend to a word in a stream of speech and either (a) inquire about the word while interrupting the speaker's speech, especially in group situations, or (b) actively map the word independently while listening to the story.

Direct explanations of word meanings while reading stories provide explicit opportunities to relate words to meanings. Reading repeatedly a story orally to young children helps the learners both to become clearer about the actors and tasks in the narrative, and to have several exposures to taught or self-learned word meanings. As described below, repeated reading of stories with word meaning explanations result in acquisition of more word meanings than simply rereading without word meaning explanations.

The consequence of both initial spontaneous word meaning acquisition and initial taught word meaning acquisition amount to Carey's "fast mapping". Brief explanations of word meanings in context provide the beginning of word meanings acquisition, not the whole of it. (Some researchers have provided more extensive word meaning instruction, e.g. Brabham & Lynch-Brown, 2002.) As Carey stressed, if the words are to become better understood--and

used--further experience with the newly mapped words will be needed. When the same word meaning is re-encountered in another context, a richer referent is established.

Methods of Promoting Primary Vocabulary

There are very few studies of recent vocabulary instruction with elementary children (National Reading Panel, 2000). Only 4 were reported found in the primary grades, and only 2 were cited. Both involved reading stories, one with direct explanations. No non-story methods were described at the primary or preschool level. We have found 13 studies of vocabulary instruction using texts conducted with elementary children in the past 15 years. Studies of interventions to facilitate vocabulary acquisition differ in (a) the ages of children (from preschool to grade six), (b) instructional group size (from 1:1 to whole class), (c) child vocabulary size, (d) numbers of readings of stories (from 1 to 3), (e) number of word meanings tested, and (f) the number of meanings taught in the study and per reading. A summary table of these studies appears as Appendix A. These studies mainly used multiple-choice vocabulary pretests and posttests--usually administered orally. The main conclusions are:

a. In 4 studies, a single reading led to an average of 15% gain in word meanings known of those tested, (Nicholson & Whyte, 1992; Senechal, 1997; Senechal & Cornell, 1993; Stahl, Richek, & Vandevier, 1991). (In all conditions, the Senechal studies included some word explanation using text pictures.) "Percent gain", means the difference between percentage of tested words known before intervention and the percentage known after intervention.

b. In 6 studies, repeated readings *without* word meaning explanations led to an average 9% gain in word meanings known (Brabham & Lynch-Brown, 2002; Brett, Rothlein,

& Hurley, 1996; Elley, 1989; Hargrave & Senechal, 2000; Penno, Wilkinson, & Moore, 2002; and Robbins & Ehri, 1994).

c. In 6 studies, repeated readings *with* word meaning explanations led to an average 26% gain in word meanings known (Brett, Rothlein & Hurley, 1996; Elley, 1989; Hargrave & Senechal, 2000; Penno, Wilkinson, & Moore, 2002; Robbins & Ehri, 1994; Senechal, 1997; Senechal, Thomas, & Monker, 1995). When meanings were explained, the percentage of words taught that were acquired was fairly consistent near 26%, whether few or many word meanings were taught per reading or in total.

d. In 2 studies, repeated readings with interactive word *discussions* led to an average of 17% gain in word meanings known (Hargrave & Senechal, 2000; Brabham & Lynch-Brown, 2002). In Hargrave and Senechal's study, Whitehurst's dialogical reading was used. In Brabham and Lynch's study, 2 discussion treatments were used. The only difference between them was whether discussions preceded and followed story reading versus the addition of discussion while reading. (The latter treatment resulted in a 26% gain in meanings known—about the same as the simple word explanation studies.)

There were no clear effects of the ages of the children, or the size of instructional groups (ranging from 1:1 in Senechal's studies to whole class in several other studies as noted in Appendix A).

Vocabulary gains associated with differences in children's vocabulary (PPVT) were examined in 3 studies (Nicholson & Whyte, 1992; Robbins & Ehri, 1994; Senechal, Thomas & Monker, 1995). In 2 of these studies, children with relatively low vocabularies acquired few word meanings from repeated readings *without* explanations, unlike those with larger

vocabularies. In the third study (Senechal et al), pictures associated with each word meanings were pointed out. In this study, when verbal explanations were added, there was little difference between lower and higher vocabulary children's gains. Explanations were not used in the other 2 studies of the effect of differences in children's vocabulary.

Overall, individual children learned the meanings of an average 26% of the word meanings explained, across a range of 3 to 20 different meanings taught *per day*. (Thus if 10 words were taught, an average of 2.6 words were learned.) Results were similar for children from age 3 to age 10 (grade 4). Note that when children are learning a quarter of the words explained, *different children* can be learning different words. Some will have known more of the words to begin with.

Those studies which compared immediate posttests with posttests 6 weeks to 3 months after reading reported word knowledge to be either about the same at both times or higher at delayed testing (Brett et al, 1996; Senechal & Cornell, 1993; Senechal et al, 1995). In Appendix A, when delayed testing was used, delayed tests are reported. In short, when word meanings are learned in context and explained, they appear to be retained well.

There appears to be no information on vocabulary taught in the primary grades using non story-based methods.

To summarize this review, there are repeated findings that encouraging vocabulary acquisition in the primary grades using repeated reading combined with word meaning explanations works. It is possible that other methods would work as well, but at present there is no evidence for this.

Purpose of Research

We report here 2 new primary grade studies of vocabulary instruction embedded in children's stories. In both studies, we worked with the regular classroom teachers who provided our vocabulary instruction to kindergarten, grade 1, and grade 2 children in whole-class sessions.

Study 1. In Study 1, 3 factors which might affect word acquisition examined: the effect of pretesting, reading books 2 or 4 times, and word explanations, Most prior studies used pretesting to establish a baseline for word meaning acquisition. We hypothesized that pretesting could be an effective way to increase word learning from hearing stories, both with and without direct word explanations.

The role of number of readings used in story-based vocabulary instruction concerned how effective different numbers of readings were effective. Hence it was of practical significance to compare learning word meanings with 2 versus 4 readings.

Reading stories with word explanations has been shown to be more effective than simply reading stories, even when read repeatedly. This contrast was included to see if the other 2 factors, pretesting and number of reading, interacted significantly with the presence or absence of word meaning explanations.

Study 2. Study 1 provided information about the role of pretesting and numbers of times stories were read for vocabulary building. Gains were of the same magnitude as previous studies of word meaning instruction in conjunction with story reading. Study 2, concerned modifications to instruction procedures to increase the amount of learning of word meanings during classroom sessions by modifying instruction procedures. In addition, transfer of word meaning knowledge from original pretest items taken from the instruction story texts to the same words in new sentence context were studied. Finally, retention of learned word meanings over time was examined.

Study 1: The Effects of Pretesting, Number of Times Books were Read, and Direct Explanation

Pretesting. Most prior studies used pretesting to establish a baseline for word meaning acquisition. Other researchers have been concerned with this effect, but none have studied word meaning acquisition with and without pretesting. Some have avoided pretesting, using separate control groups (Robbins & Ehri, 1994; Stahl, Richek, & Vandevier, 1991). The others summarized in Appendix A all used pretests.

We hypothesized that pretesting could be an effective way to increase word learning from hearing stories, both with and without direct word explanation. Furthermore, our method of assessing word meaning knowledge involved presenting such words in context sentences (see below) drawn directly from the books read. This might sensitize children to words appearing in a story and thus be a useful tool for fostering vocabulary growth.

Number of readings. The role of number of readings used in story-based vocabulary instruction has 2 practical issues. First, the percentage of word meanings learned could become better with more readings because children have been exposed to each word meaning of interest more often. Alternatively, more reading could reduce learning word meanings because children become disinterested with too many re-readings. Second, in the authors' experience children dislike interruptions for word meaning explanations on the first reading but not subsequent readings. When different word meanings are explained each day, and if the first reading does not include many word meaning explanations, there will be more opportunities to explain word meanings with 4 readings (3 with meaning explanations) than with 2 readings (1 with meaning explanations). Hence it is of practical significance to compare learning word meanings with 2 versus 4 readings.

Reading with explanations versus reading without meaning explanations. It is established that reading with explanations is more effective. We were interested in the interaction of the other factors with this one. In particular, pretesting and more readings might prove to be more effective in the no explanation condition.

Method

The School System

This study was conducted in a publicly-supported Catholic school system. In Ontario, separate Catholic school systems are supported in all school districts. About one third of Ontario children attend Catholic school boards. The schools follow the same mandated curriculum as the public schools, although they also include some religious instruction. Teaching methods do not differ systematically between public and Catholic school boards.

Sample

The children in this study attended a Toronto Catholic District School in an area serving mainly children from Portuguese families in a working class district. There were few, if any, families in public housing or on welfare, and few parents with college degrees. About 50% of grade 3 children in this school first learn a language other than English as reported by the Ontario Quality and Accountability Office. Similar percentages can be assumed for younger children. Information on language status of individual children is not available. A total of 43 kindergarten (24 girls), 37 grade 1 (13 girls), and 32 grade 2 children (14 girls) were included. At each grade, children were in 2 classrooms. In each grade, the 2 classroom teachers read the books and provided word explanations.

Books Used

Books were selected in consultation with the collaborating teachers and the school librarian. We attempted to select books which would be (a) interesting and (b) include a number of word meanings that the children might not know. All books were narrative fiction. In kindergarten, books selected were: *There is Much More* (Alexander, 1987), *Mud* (Ray, 1996), and *Something from Nothing* (Gilman, 1992). In grade 1, books selected were: *Going Down the Road* (Schertle, 1995), *Seashore* (Zolotow, 1992), and *The Tree that Grew to the Moon* (Fernandes, 1994). In grade 2, books selected were: *Drac and the Gremlin* (Baillie, 1998), *Alexander and the Wind-up Mouse* (Lionni, 1969), and *Brenda and Edward* (Kovolski, 1984).

Note that different books were appropriate to the different grades. This means that conclusions about absolute levels of word knowledge from these class books cannot be drawn. Conclusions about *gains* in word meaning knowledge and comparing these across grades can be drawn.

Word Meanings Selected for Instruction

In each grade, 3 books were used. Twelve word meanings from each book read twice and 24 word meanings from the book read 4 times were selected. Thus a total of 48 word meanings were used in each grade. Initial word selection was based on our judgement of word meanings likely not to be known by all children in the class. Meanings were then checked in Dale and O'Rourke's (1981) *Living Word Vocabulary (LWV)*. Word meanings were omitted if reported known by more than 80% of children at grade 4 in *LWV*, as prior research had shown that such meanings are actually known by most grade 2 children (Biemiller & Slonim, 2001). Note that children below grade 4 were not tested for the *LWV* since the tests presupposed fluent reading skill. The reported grade level for each selected word meaning (grade at which 67% – 80% of children are reported to know a particular word meaning in *LWV*) was used to match the

levels of word meanings used with 2 versus 4 readings. Different word meanings were explained during each reading. Grade levels of word meanings in the taught and not taught conditions were also matched using the *LWV*. This provided a rough indexing of word difficulty and ensured that levels of difficulty were similar in each test condition.

Testing Knowledge of Word Meanings

General vocabulary test. A general vocabulary test was needed to create matched cohorts of children. An abbreviated version of test B used in Biemiller and Slonim (2001) was used, with 40 rather than 57 items. This test correlated .80 with the PPVT in a separate study (Biemiller, 2005). Approximately every 3rd item in the original test was deleted. The test contained 7 items at *Living Word Vocabulary* (grade) levels 2, 4, 6, and 8, and 6 items at levels 10 and 12. As in the normative study (Biemiller & Slonim, 2001), there were some items at *LWV* grade levels 4 and 6 that were known by some grade 1 or 2 children. Following procedures in Biemiller and Slonim (2001), children were tested orally at individual sessions. Each tester had been trained to read the sentences as they appeared in the test and ask, "What does … mean in this sentence?". The tester wrote down exactly what the child said. The general vocabulary test and the pretests and posttests each took about 20-25 minutes. Children were told they could say, "don't know".

Coding of responses involved categorizing children's word explanations as "known" (score 1), "possibly known" (score .5) or "not known" (score 0). ("Possibly known" meant that the students response *could* be interpreted as knowing the meaning but wasn't conclusive.) Examples of scoring items are given in Appendix B. Generally, any plausible superordinate word (e.g. for "dog", child reports that "A dog is an animal"); or clear example (e.g., "My dog

plays with me") sufficed to demonstrate knowledge of the word's meaning. Gestures or pointing were also accepted when these were plausible. In other words, word meanings were considered known if the child knew enough about a meaning to make sense in the story. Good dictionary definitions were not required. Reliability was $r = .90$ between 2 raters on the same words across 3 children. This is similar to coding results reported in (Biemiller & Slonim, 2001). Virtually all children were able to give some plausible responses. Note that this measurement of vocabulary is more demanding than multiple choice tests (using either verbal or pictured alternatives). Just as the case is with multiple choice methods (e.g., with pictures as in the Peabody Picture Vocabulary Test, or with verbal alternatives as used for the *LWV*), correct scores do not discriminate degrees of word meaning knowledge. Lower levels of scores were to be expected using children's verbal explanations than children's multiple-choice scores (Curtis, 1987). Biemiller (2005) reported similar results. Examples of test items and scoring responses are given in Appendix B.

Test of instruction words. The same method was used with target words from the books being read. Tests were constructed using context sentences from books to be read and which included one of the target words. Scoring procedures were the same as those for the general vocabulary test.

Design

In this study, the effects of repeated reading with direct instruction of word meanings versus repeated reading without word explanations, reading stories 2 or 4 times, and pretesting or not were examined. The basic design was a pretest/posttest assessment of the effect of word meaning instruction during book reading versus repeated reading without instruction on the acquisition of word meanings . This was done in 3 primary grades: kindergarten, grade 1, and

grade 2. In addition, in each grade, 2 books were read twice, while another book was read 4 times. Assessing the effect of pretesting on word meaning acquisition was done by dividing children in each classroom into 2 matched cohorts. At pretest, half of the total set of 48 word meanings (in each grade) was given to one cohort in each classroom, the other half of the meanings was given to the other cohort. At posttest in each grade, all children received all 48 word meanings.

Pre/post test gains. To test pre/post gains, children were tested in 2 matched cohorts, based on the general vocabulary test given one month before the pretest. Children in each cohort were pretested and posttested on 24 word meanings. Meanings of 12 of these word meanings were instructed, while another 12 word meanings, matched for grade levels in the *LWV,* were not instructed. A second cohort in each class, matched for general vocabulary, was pretested and posttested on an additional 24 word meanings. Different books and words were used in different grades. The basic focus of this study was the significance of pre- versus posttest results and the interaction of pre- versus posttest scores when instructed and not instructed.

Effect of number of times a book was read. In each grade, 12 word meanings were taught from books read twice in 1 week . There were 6 meanings from each of 2 books (combined across 2 cohorts). Twelve matched non-instructed word meanings from these books were also pretested and posttested. Twelve more word meanings were taught while reading a third book 4 times during the second week of instruction. Again, 12 additional matched word meanings from the book read 4 times were also tested. The effect of the number of readings on pre/posttest gains could be examined.

Effect of pretesting. In order to assess the effect of pretesting on the acquisition of word meanings, within each class children were divided into 2 equal cohorts matched on general

vocabulary (using a 40-word version of the Root Word Inventory (Biemiller & Slonim, 2001). The general vocabulary test was given one month before the pretests. As above, in each grade, the word meanings used in the study were divided into 2 groups: 24 instructed meanings and 24 uninstructed meanings. Each set of 24 words was further divided into 12 "group A" meanings and 12 "group B" meanings. One cohort was pretested on the "A" meanings but posttested on both A and B meanings. The other cohort was pretested on the "B" meanings but also posttested on both A and B meanings. Words in each subgroup were matched on *LWV* grade levels. The same tests were used (noting that for any meaning tested, half of the children had not been pretested on the word). In each grade gains due to instruction vs. not were replicated across 2 different cohorts of children in 2 different sets of word meanings. Effects of pretesting was examined by considering all posttested words and examining the interaction between cohorts and word lists. If pretesting facilitated word acquisition (whether taught or not), there should be a significant interaction for gains between cohorts and word lists.

Classroom Vocabulary Instruction

Prior to testing word meaning acquisition in regular classrooms, the procedure of explaining word meanings while reading a book was pretested with children in the laboratory school at the Institute of Child Study. In several small groups with kindergarten and grade 1 children, on the first reading, one or more children expressed complaints about interruptions for explaining word meanings. However, on second and subsequent readings, they did not object to interruptions for word meaning explanations. Dickinson & Smith (1994) also reported that reading without interruption was valuable. Their study did not involve repeated reading.

Thus in the regular classroom study, on the first reading a story was read as is normally done in a classroom, with questions about comprehension at the end but without vocabulary interruptions. The teacher usually brought the class to a carpet area surrounding a reading chair.

One or 2 word meanings were explained before the first reading. For example (from kindergarten, *Clifford at the Circus*): "Before we begin, I want to tell you what the word **circus** means--a travelling animal show with clowns and usually takes place in a large tent. (The picture of a circus on the book's cover is shown to the class.)

The book was then read one or 3 more times with different word meanings explained during each reading. On the second day, teachers began the session with instructions to their class, e.g., "We are going to read the story again and this time I am going to stop and tell what some of the words mean." On these days, 4-6 different word meanings were explained on each reading. As each sentence with a word to be explained was reached, the teacher would re-read the sentence and ask if anyone knew what the word meant. If a plausible explanation was given, the teacher agreed and went on. Otherwise, the teacher provided an explanation. Explanations of word meanings were brief. For example from a grade 2 story, *"It seemed like a good **solution***. What does **solution** mean?" If no plausible answer was supplied, the teacher would say, "A **solution** is the answer to a problem."

The teacher continued reading the story until the next "target word" for that reading came up. That sentence would be repeated, and meaning of the target word sought or supplied. If there were a third reading, the teacher would begin the session by saying, "We are going to read the story again and I am going to stop and tell you what some different words mean. Listen for the words we talked about yesterday." The procedure continued as in the second reading. The

same procedure continued for 4th readings. Different word meanings were taught with the 2nd, 3rd, and 4th readings.

After reading each story, we included 2 open-ended comprehension questions. This was done to maintain some focus on the content of the story even when word meanings were the focus of the instruction. No effort was made to assess comprehension in this study.

Results

General Vocabulary

At the pre-test (April, 2002), knowledge of a sample of grade level 2 to grade level 12 word meanings from the *Living Word Vocabulary* was somewhat lower than a normative population in kindergarten and grade 2, as seen in Table 1. Grade 1 was at about the same percentage of word meaning knowledge. The lowered results in kindergarten and grade 2 would be expected as this is a working class district with about half of the children not speaking English at home. In Biemiller & Slonim (2001), a large apparent increase in vocabulary was noted between grades 1 and 2 (possibly reflecting increased skill in explaining words as well as actual acquisition of new word meanings). In Study 1, no such increase was observed between grades 1 and 2. This population did not add many word meanings drawn from *LWV* grade levels 4 and above.

Table 1 about here

Pretest vs. Posttest Scores

Statistical analysis was conducted with the Systat program (Wilkinson, Hill, Miceli, Howe, & Vang (1992). The first concerned the significance of differences or gains between pretests and posttests. The ANOVA was conducted on pretest and posttest scores from 12 instructed word meanings and 12 non-instructed word meanings from each cohort at each grade.

There were 3 between-group factors (grade, sex, and cohort) and one within-group factor (pre- vs. posttest).

In Table 2, in all grades children had higher scores on posttests. Overall, 25% of all words were known at pretest and 42% were known at the posttest. The effect size, Cohen's *d,* was 1.21 (Cohen, 1988). The main effect for pre/posttest result was highly significant, $F(1, 100) = 182.726$, $p < .001$.. Pre/posttest differences had a nearly significant interaction with grade, $F(2, 100) = 2.986$, $p < ,06$. This reflected smaller pre/post difference in kindergarten (posttest-pretest = 14%) compared to grade 1 (18%) or grade 2 (20%). Otherwise, there were no significant interactions between pre/post scores and sex or cohort, or 3 way or 4 way interactions.

There was a main grade effect, $F(2, 100) = 3.287$, $p < .05$. However, grade differences were determined by the books used. It would be possible to choose books and word meanings creating any grade order effect (e.g. lowest or highest at kindergarten). There was also a significant main sex effect, $F(1, 100) = 4.260$, $p < .05$. Girls had higher scores overall (36%) than boys (32%). There were no sex interactions with pretest/posttest, grade, or cohort.

Table 2 about here

Gains for Instructed vs. Non-instructed Word Meanings

Having established that pre/post gains were significant, further analysis was made of factors affecting pre/post gains. An ANOVA was conducted on gain scores (posttest - pretest) from instructed word meanings and non-instructed word meanings. There was 1 between-group factor (grade) and 2 within-group factors (instruction or not, and number of readings--2 vs. 4). Sex and cohort were omitted as factors since all interactions with pre/post scores were not significant.

In Table 2, across grades, pre-post gains were 22% (*SD* 19%) for instructed words and 12% (*SD* 15%) for non-instructed words. ($d = 0.53$) Thus instruction makes a difference, $F(1,109) = 19.715$, $p<.001$. An additional 10% gain occurred when word meanings were instructed in addition to repeated reading. There was no significant interaction between grade and gains on instructed versus non-instructed meanings.

Reading Books Twice Versus Four Times.

Table 2 also shows data for reading books 2 or 4 times. Overall, there was no significant difference in gains when reading text 2 versus reading 4 times. However, there was a significant interaction in gains between the number of readings and grade, $F(2, 109) = 3.489$, $p < .05$. When text was read 4 times, kindergartners gains scored 6% higher than when text was read only 2 times. For grade 1 students, this difference was 7%. But for grade 2 students, gains after 4 readings were actually 5% lower than scores after only 2 readings. (See Table 2).

There was a near significant 3-way interaction between instruction, grade, and reading 2 or 4 times, $F(2, 109) = 2.817$, $p < .06$. In kindergarten, there was little difference between 2 versus 4 readings. Gains were larger when meanings were instructed (23%) versus non-instructed (8%). In grade 1, there was no difference between gains for instructed word meanings when stories were read 2 or 4 times, but non-instructed gains were larger for word meanings when a story was read 4 times (21%) than when stories were read twice (6%). In grade 2, there was little difference in word meaning gains between instructed words read 2 or 4 times. Anomalously, word meaning gains for non-instructed words read 4 times were lower (14%) than for words read twice.

Effects of Pretesting on Word Acquisition

Table 3 shows posttest word scores—both when meanings were pretested, and when they were not pretested. An ANOVA was conducted on posttest scores for the 2 wordlists. The 2 between-group factors were grade and cohort. The 2 within-group factors were wordlist and instruction versus not (children in each cohort had both wordlists at the posttest). For example, in grade 1, cohort 1 was pretested on wordlist A while cohort 2 was pretested on wordlist B. While wordlist A meanings proved to be easier than wordlist B meanings, there was no difference between pretested and non-pretested words (e.g., there was no significant interaction between cohort and word list). While there was a significant difference between wordlists in grade 1 (but not in the other grades), there were no significant interaction between cohort and wordlist, $F(1, 107) = .003$, *ns*. There was also no significant interaction between grade, cohort, and wordlist, $F(2, 107) = .090$, *ns*. There was also no significant interaction between wordlist, cohort, and instruction/not.

Table 3 about here

Classroom differences

There were significant differences between teachers' effectiveness in teaching word meanings. 3 of the teachers (one in each grade) had substantially larger gains with instructed word meanings (23% – 33%) than the other 3 teachers (6% - 16%). The same 3 teachers had larger gains for the non-instructed words as well (14% - 19%) than the other 3 (6% - 9%). Clearly, teachers varied in their effectiveness. Each teacher was visited twice while teaching vocabulary, but no clear evidence was seen that might explain these differences. At each visit in each classroom, the lesson was taught as planned. Each teacher stated that she was implementing the program as specified and when visited, was doing so.

Discussion: Study 1

In this study, average gains of 12% of word meanings were obtained using repeated reading. Adding word explanations added a 10% gain for a total gain of 22%. Reading books 2 or 4 times had different effects in different grades, with kindergarten children profiting most from 4 readings, whereas by grade 2 there was no apparent benefit of 4 versus 2 readings. No effect on gains due to pretesting was found.

Repeated reading. Reading stories several times is normally welcomed by preschoolers and kindergartners. The kindergarten children gained 23% of instructed word meanings when stories were read 4 times, but only 16% when stories were read twice. In grades 1 and 2, 21% to 24% of instructed words were acquired, when stories were read either 2 or 4 times. Note that different words were taught on each reading. Grade 1 children acquired more non-instructed word meanings when stories were read 4 times (20% gain) rather than 2 times (6% gain). More readings increased the probability that non-instructed word meanings would be acquired. Thus there was a clear benefit for kindergarten and grade 1 children to hear stories several times even when meanings were not taught explicitly, although more words were learned when instructed. It is less clear that many readings were needed by grade 2 children. In grade 2, similar percentages of words were acquired when stories were read 2 or 4 times.

Pretesting. Pretesting had no measurable impact on the acquisition of word meanings in any grade. The lack of a pretest effect indicates that pretesting vocabulary did not influence posttest gains. This is good for research—testing need not be seen as an influence on interventions. However, the same result is unfortunate for educators. If pretesting had facilitated word meaning acqusition, it could have been incorporated into instruction for that purpose.

Teacher differences. Clearly, in this study some of the teachers were more effective than others in teaching vocabulary to children. Interestingly, in those classes with larger gains for intstructed word meanings, there were also larger gains for non-instructed word meanings. This may have been the result of convincing children that word meanings are important. At any rate, bringing all teachers to the level of effectiveness reached by the most effective teachers would clearly result in more word meaning learning.

Study 2: More Intensive Word Instruction and Transfer of Word Meanings to New Contexts

Increasing the percentage of word meanings learned. Average children have acquired 6000 root word meanings at the end of grade 2 (Biemiller, 2005). Children in the lowest quartile have acquired an average of 4000 root word meanings. If instruction in primary classrooms is to have a significant impact on vocabulary acquisition, low vocabulary children need to acquire a meaningful proportion of the 2000 root word meaning difference. For example, if low vocabulary children could acquire 10 new root word meanings per week while at school, approximately 400 could be added in a year. If 25% of word meanings taught are learned (as seen in studies reviewed here and Study 1), 40 word meanings would have to be taught weekly. If a higher percentage of word meanings could be learned, proportionately fewer word meanings would have to be taught.

A 5 day sequence was developed for each story used. Suggestions for modifications to improve the percentage of word meanings learned were solicited from the participating teachers. The teachers suggested adding reviews each day, and that children might not have been attending to other children's explanations as well as they attended to teacher explanations. In addition, we decided to increase the number of word meanings taught each day from 4 – 6 to 7 – 9, and to add

a final review day using new context sentences. Thus the revisions to increase word meaning acquisition involved (a) increasing the numbers of word meanings taught by teaching more meanings on each reading and using four readings, (b) using vocabulary reviews of word meanings taught during each reading of a story, (c) using an additional review with new context sentences in a final review, and (d) using only teacher explanations of word meanings.

Retention of learned word meanings. As was done in some of the studies reviewed, immediate and delayed posttesting were contrasted. This was to determine if there was any loss of word meanings learned. For Study 2, children were tested 2 weeks after the instruction period, and again for a "delayed posttest" 6 weeks after instruction. At the time of planning the study, we had not noticed that previous studies showed gains rather than losses for delayed posttests. Some losses of word meaning knowledge were expected. The question was how large the losses would be.

Transfer of word meanings to new contexts. In Study 2, transfer of learned word meanings to new verbal contexts was also examined. Word meanings learned in a particular sentence/story context must be useable outside of the original story. Two kinds of delayed posttests were contrasted. One posttest used the pretest items--context sentences taken directly from the books used for instruction. The other posttest used new and different context sentences which were not related to the books used for instruction. This allowed some examination of the extent the children could generalize their understanding of word meanings.

Method

Sample

Study 2 was conducted the following year (2003) in the same school as Study 1 and with the same teachers (except for 1 kindergarten teacher who left and was not replaced due to

declining enrolment). The teachers included 1 kindergarten teacher with 2 classes, and 2 teachers each in grades 1 and 2. Twenty-eight kindergarten (16 girls), 37 grade 1 (16 girls), and 42 grade 2 children (21 girls) participated in all 3 sessions of the study. (Seven additional children missed either the immediate or delayed posttest. Their data were not included in the analysis.) As in Study 1, approximately half of the children in this school were "English Language Learners" or ESL children as reported by the board. These were children who reported using another language more than half of the time at home.

Books Used for Reading to Classes

The same approach for choosing books was used as in Study 1. Some of the books used in previous study were used again in this study. Two books were used at each grade level. Books were chosen by the teachers librarian, and researchers together. Our choices were more influenced by the school librarian than in Study 1, as she had much knowledge of books rich in vocabulary. In kindergarten, *Clifford at the Circus* (Bridwell, 1985) and *Jillian Jiggs* (Gillman, 1977) were used. In grade 1, *Going Down the Road* (Schertle, 1995), and *Julian* (Khalsa, 1989) were used. In grade 2, *The Chicken Cat* (Lionni, 1969), and *Alexander and the Wind-up Mouse* (McLellan, 2000) were used.

Selecting Word Meanings for Explanation

Word meanings were selected which many children wouldn't know. As in Study 1, these choices were confirmed by checking to see if the meanings were above 80% at (grade) level 4 in the *Living Word Vocabulary* (Dale & O'Rourke, 1981). Most word meanings in the stories that were above Dale and O'Rourke's word meanings which test above 80% in fourth grade testing were used for instruction. Thus many more word meanings were instructed than in Study 1. In Study 2, word meaning scores at pretest were also used to eliminate meanings which were

known by 85% or more of the children. After deleting word meanings with high scores at pretest, across 2 weeks, there were 42 word meanings to teach in kindergarten, 55 word meanings in grade 1, and 46 word meanings in grade 2. These included all possible words for instruction that we identified and that were below 85% at pretest. (In the discussion, we will return to the issue of word meaning selection.)

Testing Word Knowledge

Testing procedures were the same as those in Study 1 and Biemiller & Slonim (2001). The same 40 word general vocabulary test was administered. In each grade, tests of words being taught in that grade were used as pretests and immediate posttests. These items were constructed using text sentences or abbreviations of text sentences. Each of these posttest items was also used by half of the children at the delayed posttest. (The pretests were longer as they included some word meanings that were excluded from instruction.)

A second set of posttest items using the same word meanings was designed providing new test sentences, not derived from the books. These were also used at the delayed posttest by half of the children. These test sentences were different from the review sentences used on the 5th day of instruction with each book. Appendix C includes sample words, *LWV* definitions, text-based test sentences, alternative non-text-based test sentences, and additional non-text illustrative sentences used in the final review of word meanings.

Design

The 3 questions of interest were the effect of improved vocabulary instruction, retention of word meanings, and transfer of understanding of word meanings when presented in new context sentences not taken from in the instruction texts. The basic design was a simple pretest-posttest-delayed posttest study of the effectiveness of changes in instruction to increase gains in

word meaning acquisition. Comparisons could be made with gains in the previous year and with results from other studies. The first posttest was made 2 weeks after 2 weeks of vocabulary instruction, and used the same test as the pretest in each grade.

At this point, a general vocabulary test was also administered. This was the same test used in Study 1. This test was used to assign children to cohorts.

In order to examine retention of word meanings over six weeks, a delayed posttest was conducted. At this time, children in each grade were divided into 2 cohorts, matched on the general vocabulary test. Each cohort received half of the original test sentences again, while half of the word meanings were tested with new sentences, not taken from the story books used for instruction. Thus all word meanings were tested with both the original sentences and alternative sentences at the delayed posttest. If understanding of word meanings was lower when presented in new non-text sentences, problems with transfer would be shown in a significant interaction between cohorts and lists of pretest items versus new-context items . Sample pretest and alternative posttest sentences are given in Appendix C.

Study 2 also included a small no-intervention group in an additional class. Eleven grade 2 children in this class were tested, using the same pretest and posttest as the intervention grade 2 children, 6 weeks after the pretest. This was the same interval as the experimental pretest – posttest period. Shortages of research staff prevented testing most of the grade 1 children in the same classroom. This no-intervention group provided a rough estimate of vocabulary gains in the absence of reading stories with the test vocabulary and instruction on word meanings.

Classroom Vocabulary Instruction

As in Study 1, on the first day, each story was read as is normally done in a classroom, with questions about comprehension at the end but no vocabulary interruptions. One or 2 word

meanings might be explained prior to the first reading. Such word meanings would be critical to the story. For example, *circus* was pre-taught in *Clifford at the Circus* (used in kindergarten). For the following 3 days, 7-10 word meanings were explained at each reading. Different word meanings were explained on each of these 3 days. As in Study 1, the teacher explained that there would be interruptions for teaching word meanings. As each sentence with a word to be explained was reached, the teacher would re-read the sentence. In Study 2, the teacher then provided an explanation of the word meaning. Again, explanations of word meanings were brief. For example from a grade 2 story, ***"****It seemed like a good* ***solution***. What does **solution** mean? A **solution** is the answer to a problem." In Study 1, the class had been asked what the target word meant, and if a plausible answer was offered, the teacher agreed and continued reading. The teachers and we suspected that children did not attend to another child's definition as well as they attended to teacher definitions. So for Study 2, children were not asked to provide meaning explanations while the teacher was reading the story. Children were invited to provide meanings on the final review day (below).

The teachers suggested adding a review each day of word meanings taught that day. (Grade 1 teachers also had the children learn to read the words explained each day.) This review consisted of reading aloud the context sentences and explanations again.

On the fifth day, the story was not re-read. All words that had been explained that week were reviewed again. On this day, new context sentences not based on the book were used. (See examples in Appendix C.) Children were asked to raise their hands if they knew what the word meant. Correct child responses were confirmed by the teacher ("yes, _____ means…"). Incorrect children's responses were corrected. The teachers could choose to review all the words

at once, or at different times on the 5th day. Thus the complete cycle for each book took 5 days, for about half an hour a day.

Results

General Vocabulary

As in Study 1, kindergarten children knew fewer word meanings than the normative sample, grade 1 children had results similar to the normative sample in Biemiller and Slonim (2001), but again grade 2 children showed no meaningful gain compared to grade 1 (Table 4). In this case, the grade 2 children were mainly those who were in grade 1 in Study 1. (Therefore, these results did not represent results from one anomalous low-vocabulary grade 2 group.) It appears that this population was acquiring the most basic word meanings well. This included meanings in *LWV* levels 2 and a few meanings in levels 4 and 6 of word meanings in Table 4, but not acquiring many word meanings at *LWV* levels 4 and 6 or above these levels, unlike grade 2 children in the normative population.

Table 4 about here

Reliability of story word meanings. Correlations between *pretest* children's scores for the 2 books were correlated between $r = .72$ and $r = .87$ across the different grades. Thus there was substantial stability in children's initial vocabulary size on the 2 different books.

Gains in Word Knowledge

Table 5 shows Study 2 results for pretest, posttest, and delayed posttest by grade sex and cohort. The overall gain from pretest to delayed posttest was 41% (*SD* 14.3). Averages of 8.2 word meanings were gained *per week* in kindergarten, 12.4 word meanings in grade 1, and 8.5 word meanings in grade 2. (To some extent, these differences in word meaning gains reflect differences in the number of word meanings taught: kindergarten, 42; grade 1, 55; and grade 2,

46.) Differences were tested with a mixed-model ANOVA with between-group factors grade, cohort, and sex, and within-group factor pre/post/delay tests (Wilkinson, L, Hill, M, Miceli, S., Howe, P., & Vang, E., 1992). Pretest, posttest and delayed posttest percentages of word meanings known differed significantly ($F(2, 192) = 528.597$, $p < .001$. (This was calculated without the small grade 2 control sample.) The effect size between pretest and delayed posttest was $d = 2.97$. There was a significant interaction between grade and pre/post/delay tests, $F(4, 192) = 4.617$, $p < .002$. This interaction reflected larger gains by children in grade 1. There were no other significant main effects or interactions. Thus sex and cohort had no impact on gains. Differences between pretest-posttest and posttest-delay tests are presented below.

An ANOVA of pretest vs. (immediate) posttest differences was significant, $F(1,105) = 530.875$, $p < .001$, reflecting a gain of 35% ($d = 2.50$). (Cohort and sex were not included as factors as they were non-significant in the first ANOVA). There was also an interaction between grade and pre/post test scores, $F(2, 105) = 6.664$, $p < .002$. As above, children in grade 1 made larger gains (42%) than children in kindergarten (32%) or grade 2 (30%). An ANOVA of immediate post- and delay-test means showed further significant gains, $F(1, 105) = 31.694$, $p < .001$ ($d = 0.26$). There was no significant interaction between grades and post/delay scores. The overall post-delay gain was 6%. Note that this means children continued to gain vocabulary for four weeks without further instruction.

In Study 1, there was a pre/post gain of 13% for repeated reading and 22% for repeated reading plus word meaning explanations or 10% above simple repeated reading. In the other studies reviewed which included repeated reading with word meaning explanations, the average gain was 26%. In Study 2, the gain of 41% represented an instruction gain of 28% over the 13% no instruction treatment in Study 1.

Table 5 about here

The small no-intervention control group gained an average of 4% or 1.8 word meanings over the six weeks between pretest and immediate posttest or about a third of a word meaning per week. This group was not retested 4 weeks later at the delayed posttest. Based on their gains over the 6 week period examined, the control group children might have gained another 1.2 word meanings during the subsequent period or a total of 3 word meanings while the average gain for instructed children was 19.4 word meanings. Most of the difference must have occurred during the 2 weeks of vocabulary instruction.

Transfer of word meanings to new contexts

Table 6 shows delayed posttest results for cohorts 1 and 2 by wordlists A and B. In each grade, for cohort 1, wordlist B was posttested with new sentences. For cohort 2, wordlist A was posttested with new sentences. An effect of new versus storybook-based context sentences would yield a significant interaction between cohort and wordlist. A mixed model ANOVA of delayed posttest scores was conducted with grade and cohort as between-group factors and wordlist as the within-group factor. The overall interaction between wordlist and cohort was significant at $F(2, 102) = 4.153$, $p<.05$. There was also a significant interaction of the wordlist by cohort by grade ($F(2, 100) = 3.663$, $p<.05$). This interaction reflected the fact that in grade 2, cohort 1 was lower on the new sentence than on the old sentences. In grades 1 and 2, the interaction between cohort and wordlist was not significant, $F(1, 62) = .638$, *ns*. Overall, there was little difference between accuracy on old and new context sentences.

Table 6 about here

Words Learned

All words taught are shown by pretest and delayed posttest percentage correct in Appendix D. At each grade, some words were either little learned by most children or much learned by most children. (See *gain* column in Appendix D.) Words were ordered by percentage of "possible gain". In other words, possible gain was defined as 100-pretest percentage, and determined the percentage of actual gain out of possible gain. A few words were completely learned. Many others were well learned (reaching 80 percent or more of possible learning). Unfortunately, we did not find any rationale for why certain words were better learned than others. In recent work, Isabel Beck and Margaret McKeown reported the same problem (personal communication).

Percentages of words known in different classrooms were highly correlated at pretest. Correlations between word means in different classrooms were: kindergarten, $r = .90$; grade 1, $r = .90$; and grade 2, $r = .81$. This replicated Biemiller and Slonim's (2001) findings showing the existence of a strong word order effect.

Classroom Differences

Gains in word knowledge between pretest and delayed posttest words scores were fairly highly correlated between classrooms. Correlations between word meaning gains were $r = .64$ in kindergarten (between classes, same teacher), $r = .63$ in grade 1 (between word means in 2 classrooms with different teachers), and $r = .37$ in grade 2 (also between word means in 2 classrooms with different teachers). Classroom pretest-delay posttest instruction gains in kindergarten were similar in both classrooms (40% and 39%--same teacher). Classroom gains were also similar in grade 1 (44% and 46%--different teachers). In grade 2, one teacher also had a gain 40%, but the other was lower at 34%. The teachers had participated more in the planning of the vocabulary intervention in Study 2. As a result, they probably had a stronger commitment

to the program than in Study 1. As in Study 1, the classrooms were visited during instruction, but we have no further hypotheses re the classroom with the smaller gain.

Discussion: Study 2

Study 2 has demonstrated that a substantial number of word meanings can be learned using repeated oral reading of stories combined with explanations of a substantial number of words and reviews of words explained. Eight to 12 word meanings were learned per week. In Study 2, results show that larger percentages of words taught were learned than in previous studies (see Table 1 and Study 1). This was the result of added reviews during instruction, and added numbers of word meanings taught per week in comparison to previous studies. This study also showed that word meanings were not lost during 4 weeks after the first posttest. In fact, there were further gains. Finally, this study showed that children could understand word meanings when tested using context sentences different from the story used for instruction.

Increased percentage of gains from modified instructional procedure. Adding 2 reviews of each word meaning taught and using teacher-supplied word meanings resulted in an increase from 22% gain in meanings known in Study 1 to 41% in Study 2. We suggest that adding 400 word meanings per year is a reasonable goal. With acquisition of 22% of word meanings taught, 1800 word meanings would have to be taught for 400 meanings to be learned. With acquisition of 41% of word meanings taught, about 1000 word meanings would have to be taught. This would mean teaching 25 meanings per week—a realistic number.

Delayed posttest gains. There was some additional word meaning acquisition during the four week interval between the posttest and delayed posttest in Study 2. The teachers did not report any additional instruction during this period. However, when study words were

encountered, it seems likely that they were noticed by the children and/or the teachers. In other words, "word consciousness" was probably high for these words (Scott & Nagy, 2004). In addition, we have suggested that instruction of our type has the effect of creating initial "mappings" (Carey, 1978) between words and meanings. These may be "partially" mapped as a result of instruction, so that if the words are encountered again during the “delay” interval, the mapping may be better consolidated so that the child can now report a meaning.

Differences in how well specific meanings are learned. There were large differences in how much learning of specific word meanings occurred (Appendix D). Note that the highest correlation of gains between teachers in a grade was around $r = .65$. This means that more than half of the variance in gains was not specific to particular word meanings. We do not have any explanation of these differences at present. In our view, until the possibility that the specific instructions may be at fault is tested, it would probably not be worthwhile to develop more advanced theories about why certain word meanings were not well learned.

Differences between grades. Children in grade 1 made larger gains than children in the other grades (45% compared to 39% in kindergarten and 37% in grade 2. There was one difference in instruction in grade 1. Teachers had children read new vocabulary at each daily review, using a simple chart of the words. We do not know if this was responsible for the larger gains in grade 1, but clearly, learning to read the words did not hinder the children’s acquisition of new meanings. For children with some reading skill, adding reading of vocabulary words is probably desirable.

Classroom differences. Four of the 5 teachers had word meaning gains of 39% or more. The 5th, at 34%, had nonetheless shown substantial improvement in instruction compared to Study 1. Without having detailed observations of fidelity of instruction, it is clear that teachers

could use the method as introduced, and could achieve substantial gains in word meanings taught.

English Language Learners. In this study, about half of the children were English Language Learners (i.e., spoke another language at home—mainly Portuguese). It was not possible to identify which individual children were English Learners. However, pretest scores were to some extent an indicator of English Learner status. The overall correlation between pretest scores and pretest-delayed posttest gains was negative at $r = -.20$ (N = 108). Thus *lower* pretests were slightly associated with higher gains--about 4% of pretest variation. The correlations for specific grades vared. Kindergarten and grade 2 classes showed negative correlations between pretest word knowledge and instruction gains, while grade 1 showed a small positive effect of pretest word knowledge ($r = .28$ or 8% of gain variance). Overall, it appears that levels of initial word knowledge had little effect on the amount of word knowledge gained.

Discussion: Educational Implications

In this section, general implications for educational practice will be considered. The most basic is whether enough word meanings can be learned to be worth teaching. In addition, this discussion will include the number of repeated readings required, teaching of many word meanings versus intensive teaching of few word meanings, direct instruction of word meanings versus more interactive teaching, selection of word meanings for instruction, and the effectiveness of long term vocabulary instruction.

Can sufficient vocabulary be taught in the primary grades? Acquiring word meanings is a function of the percentage of word meanings taught that are learned, and the number of word

meanings taught. With kindergarten and grade 1 children, the number of readings influenced the learning of word meanings *not* taught directly. In Study 1, about 4 word meanings per week were acquired, including both 12 instructed word meanings and 12 non-instructed meanings. In Study 2, the number of word meanings taught was increased to an average of 21 – 27.5 word meanings per week. This included all word meanings we identified in each story used as being above the *LWV* grade level 2 (see Biemiller & Slonim, 2001) and which were below 85% on the pretest. Eight to 12 word meanings were acquired per week, the highest being in grade 1, where the most meanings were taught.

Biemiller and Slonim (2001) and Biemiller (2005) reported that words appear to be learned in largely the same order by children from various groups. Thus if appropriate word meanings could be taught at a successful rate and continued to do so for the 3 primary years, a child could acquire 1000-1500 additional word meanings. This would be enough to significantly improve the vocabulary of children with initially low vocabularies. However, using whole class methods would also increase the vocabulary of children with initially greater knowledge of words. Note that both English Language Learners and children raised in English will benefit from such instruction.

How many times should texts be reread? In grades 1 and 2, the percentage of meanings learned was about the same when books were read twice than when books were read 4 times (Study 1). (In kindergarten, reading books several times was more effective, and is generally welcomed by children.) However, more word meanings can be taught in a week if the same book is read 4 times rather than 2 times. This is because different word meanings are taught on each of 3 readings rather than 1. If vocabulary can be taught *without* providing a first reading with a minimal word meaning explanation, then the decision to use 1 or 2 versus 4 readings is

not important, and other considerations could be used. For example, teaching some vocabulary *after* a story is read, again revisiting the relevant text (and page and pictures) might be as effective as teaching during instruction. (We do not believe that teaching many word meanings *prior* to reading would be as effective.) The present method was effective, but the issue of repeated reading procedures requires more research.

Interactive instruction. The method of vocabulary instruction used in these studies involved simple direct instruction, and in Study 2, no children's responses were sought except in the final review. These lessons were taught on a whole-class basis. With smaller groups, we believe more discussion and student-supplied meaning explanations could be used. In addition, stories for vocabulary instruction could be targeted for specific groups. Whether the time requirements of such approaches could be justified remains to be seen.

Teaching many versus few meanings. A closely related issue is the choice of teaching many word meanings briefly versus teaching a few more intensively. We have chosen to teach many meanings without extensive discussion or use of each meaning taught. One practical issue is that different children prove to know different meanings. Hence many meanings must be taught to ensure opportunities for each child to learn. Others have preferred to teach a few meanings intensively (e.g. Beck, McKeown, & Kucan, 2002). In a study using their approach (Beck & McKeown, 2003, May), they presented 6 word meanings per week for 7 weeks with many activities with each word meaning. Approximately 30% gains of the word meanings taught were acquired, or about 2 per week. At this point, introducing many meanings appears to be more effective.

Teaching vocabulary to preliterate versus literate children. Children in the primary grades are generally "preliterate"—they do not understand language in print as well as they

understand oral language. Most children in grade 2 and below are "preliterate" as defined here. Such children are more likely to encounter new vocabulary in stories and other texts read to them than in books they read. Building vocabulary often involves acquiring word meanings from others. The methods tested in Studies 1 and 2 were intended for preliterate children.

For children who are literate, other approaches become relevant. A literate child can stop reading and ponder or ask about an unfamiliar word (unlike a listener, especially a listener in a group situation). Such words can be marked in a book or put in a list for subsequent attention. In general, literate children can take much more responsibility for building their vocabulary.

Word meanings to teach. It is one thing to determine an effective means of teaching vocabulary and quite another to determine *what* vocabulary will be needed during the primary grades. Although Biemiller and Slonim (2001) showed that there is a substantial order to word acquisition, they did not have an accurate listing of the 4000 or so root word meanings which are likely to be learned during the primary years. (An individual child will learn on average about 2500 root word meanings during this period.)

We are currently attempting to identify word meanings for use in the primary grades. Our strategy involves identifying most root word meanings that are known by children with advanced vocabularies at the end of grade 2, but not by some or all of the rest of grade 2 children. Word meanings are learned in a rough order (Biemiller & Slonim, 2001). Thus concentrating on meanings that children are likely to learn next, and attempting to speed up this process by teaching such meanings may best help low vocabulary children build larger vocabularies.

Research needed. The evidence in Study 2 strongly supports the possibility of teaching enough root word vocabulary in the primary grades to be useful. Research in the near future should include; (1) identifying word meanings appropriate for instruction in the primary grades, (2) determining if higher levels of word meaning acquisition be sustained over a substantial period of time, (3) determining if children who acquire a substantial number of relevant word meanings in the primary grades prove to have significantly improved reading comprehension by grade 3 or 4. We will primarily be working on identifying words learned by average and advanced primary-grade children, as without this, more extended teaching is hard to justify.

References

Alexander, S. (1987). *There is More Much More.* Fairbanks, AK: Gulliver Books.

Baillie, A., & Tanner, J. (1991). *Drac and the gremlin.* Victoria, Australia: Puffin at Penguin Books.

Beck, I. L. & McKeown, M. G. (2003, May). Promoting vocabulary development in the early grades. A presentation at the annual conference of the International Reading Association, Orlando, FL.

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

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74(4),* 506-521.

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., & 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.

Biemiller, A. (2005). Size and sequence in vocabulary development: Implications for choosing words for primary grade vocabulary instruction. In A. Hiebert. & M. Kamil, (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp 223-242). Mahwah, NJ: Erlbaum.

Bloom, B. (1976). *Human characteristics and school learning*. New York, NY: Mcgraw-Hill.

Brabham, E. G., and 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.

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal*, 96, 415-422.

Bridwell, N. (1985). *Clifford at the circus.* New York, NY: Scholastic Inc.

Cantalini, M. (1987). *The effects of age and gender on school readiness and school success*. Unpublished doctoral dissertation. Ontario Institute for Studies in Education. Toronto, Canada.

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan, and G. A. Milller (Eds.), *Linguistic theory and pyschological reality* (pp 264-293). Cambridge, MA: MIT Press.

Carney, J. J., Anderson, D., Blackburn, C., & Blessing, D. (1984). Preteaching vocabulary and the comprehension of social studies materials by elementary children. *Social Education*, *(48(3),* 195-196.

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

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

Christian, K., Morrison, F. J., Frazier, J. A., & Massetti, G. (2000). Specificity in the nature and timing of cognitive growth in kindergarten and first grade. *Journal of Cognition and Development, 1(4),* 429-448.

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

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

Curtis, M. E. (1987). Vocabulary testing and vocabulary instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquistion,* pp 37-52. Hillsdale, NJ: Lawrence Erlbaum Associates.

Dickinson, D. K., McCabe, A., Anastasopoulos, L., Peisner-Feinberg, E. S. & Poe, M. D. (2003). The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. *Journal of Educational Psychology, 95,* 465-481.

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.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*, 174-186.

Feitelson, D., Goldstein, Z., Iraqi, J., and Share, D. I. (1991). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly, 28*, 70-79.

Fernandes, E. (1994). *The tree that grew to the moon.* Toronto, Canada: Scholastic Canada.

Gillman, P. (1977). *Jillian Jiggs.*. New York, NY: Scholastic Inc.

Gilman, P. (1992). *Something from nothing.* Toronto, Canada: Scholastic Canada.

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.* 765-90.

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

Hart, B. & Risley, T. R. (1999). The social world of children learning to talk. Baltimore, MD: Paul Brookes Pub.

Hart, B. & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3. *American Educator, 27*(1), 4-9.

Khalsa, D. K. (1989). *Julian*. Montreal, Canada: Tundra Books.

Lionni, L. (1969). *Alexander and the windup mouse.* New York, NY: Scholastic Inc.

Lovolski, M. A., (1984). *Brenda and Edward.* Toronto, Ont., Canada: Kids Can Press.

McClellan, S. S. (2000). *The chicken cat.* Markham, ON, Canada: Fitzhenry and Whitside.

McKeown, M. G., Beck, I., Omanson, R. C., & Perfetti, C. A. (1983). The effects of long-term vocabulary instruction on reading comprehension: A replication. *Journal of Reading Behavior, 15(1),* 3-18.

Medo, M. A., Ryder, R. (1993). The effects of vocabulary instruction on readers' ability to make causal connections. *Reading Research Quarterly, 33(2),* 119-134.

Morrison, F. J., Smith, L., & Dow-Ehrensberger, M. (1995). Education and cognitive development: A natural experiment. *Developmental Psychology, 31,* 789-799.

National Reading Panel (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* Washington, D.C.: National Institute of Child Health and Human Development.

Nicholson, T., & Whyte , B. (1992). Matthew effects in learning new words while listening to stories. In C. K. Kinzer & D. J. Leu (Eds.*), Literacy research, theory, and practice:*

Views from many perspectives: Forty-first Yearbook of the National Reading Conference (pp 499-503). Chicago, IL, The National Reading Conference.

Penno, J. F., Wilkinson, A. G., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanaiton and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology, 94,* 23-33.

Ray, M. W. (1996). *Mud.* Voyager Books Inc. 1996

Robbins, C., & Ehri, L. C. (1994). reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86(1),* 139-153.

Scarborough, H. S. (1998). Early identification of children at risk for reading disabilities: Phonological awareness and some other promising predictors. In B. K. Shapiro, P. J. Accardo, & A. J. Capute (Eds.), *Specific reading disability: A view of the spectrum* (pp. 75-119). Timonium, MD: York Press.

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities): Evidence, theory, and practice. In S. B. Neuman and D. Dickinson (Eds.) *Handbook of early literacy research* (pp. 97-110). New York, NY: Guilford Press.

Schertle, A. (1995). *Going down the road.* Orlando, FL: Harcourt.

Scott, J. A., & Nagy, W. E. (2004). Developing word consciousness. In J. F. Baumann & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice* (pp 201-217). New York, NY: Guilford Press.

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

Senechal, M., & Cornell, E. H. (1993). Vocabulary acquisition thorugh shared reading experiences. *Reading Research Quarterly, 28(4),* 360-374.

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.

Stahl, S. A., Richek, M. A., & Vandevier, R. J. (1991). Learning meaning vocabulary through listening: A sizth grade replication. In J. Zutell & S. McCormick (Eds.), *Learner factors/teacher factors: Issues in literacy research and instruction: Fortieth Yearbook of the National Reading Conference.* (pp. 185-192).

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

Tomesen, M., & Aarnoutse, C. (1998). programme for deriving word meanings. *Educational Studies, 24(1),* 107-128.

Wilkinson, L, Hill, M, Miceli, S., Howe, P., & Vang, E. (1992). *Systat for the Macintosh*, Version 5.2. Evanston, IL: Systat, Inc.

Wixson, K. K. (1986). Vocabulary instruction and children's comprehension of basal stories. *Reading Research Quarterly, 21(3),* 317-329.

Zolotow, C. (1992).The Seashore Book. Toronto, Ont., Canada: Harper Collins.

Author Note

Andrew Biemiller is a professor emeritus at the Inst. of Child Study, Dept. of Human Development and Applied Psychology, University of Toronto. He also is an educational consultant. He can be reached at abiemiller@oise.utoronto.ca. Catherine Boote is a research associate at the Institute of Child Study. This research was conducted under a grant from the Ontario Ministry of Education, and with the support of the Toronto Catholic District School Board and the teachers and principal of a school in that Board. Without their support and participation, this project could not have been carried out during the past 2 years. In particular, we want to thank the teachers who carried out the actual interventions, and who contributed to improvements in vocabulary instruction reported here. We also thank the many graduate students at the Inst. of Child Study who along with us carried out the pretests and posttests. Finally, we are indebted to Monique Senechal and the anonymous reviewers for their many helpful suggestions.

Appendix A
A Comparison of Words Learned in Different Elementary Vocabulary in Story Instruction Studies

Grade/Study	Condition, instruction group size (1:n)	Nu. Chil.	Nu. Books	Nu. of Repeated Readings	Nu. of Words Tested	Nu. Of Words Taught Per day	Pretest Percent Correct [a]	Posttest Percent Correct	Gain Percent Correct	Total Nu. of Words Learned	Average Nu. of Words/ day
Age, 4-5 yr. Senechal &Cornell (1993)	Single reading, 1:1 reading using										
	pictures re meanings	39	1	1	10	10	16%	38%	22%	2.2	2.2.
	added word repetition	39	1	1	10	10	14%	37%	23%	2.3	2.3
	added recasting	39	1	1	10	10	18%	40%	22%	2.2	2.2
	added questioning	39	1	1	10	10	20%	43%	23%	2.3	2.2
Preschool, 4 yr. Senechal, Thomas, & Monker (1995)	Repeated reading using pictures re meanings:1:1										
	low PPVT group	8	1	2	13	0	28%	28%	0%	0	0.0
	high PPVT group	8	1	2	13	0	26	37	11	1.4	0.7
	Addition of verbal explan.:1:1										
	low PPVT group	8	1	2	13	13	31	45	14	1.8	0.9
	high PPVT group	8	1	2	13	13	29	47	18	2.4	1.2
Preschool, 4 yr., delayed voc. dev. Hargrave & Senechal (2000)	Repeated reading: 1:9										
	"Regular" reading	18[h]	10	2	18	0	17%	23%	6%	1.1	0.06
	"Dialogic" reading	18[h]	10	2	18	?	12	24	12%	2.2	0.11

Grade/Study	Condition	Nu. Chil.	Nu. Books	Nu. of Repeated Readings	Nu. of Words Tested	Nu. Of Words Taught Per day	Pretest Percent Correct [a]	Posttest Percent Correct	Gain Percent Correct	Total Nu. of Words Learned	Average Nu. of Words/ day
Presch. 3 & 4 yr. Senechal (1997)	Single reading, reading using pictures re meanings :1:1	20	1	1	10	10	28	32	4	0.4	0.4
	Repeated reading using pictures re meanings:1:1	20	1	2	10	10	33	46	13	1.4	0.7
	Addition of verbal explan.:1:1	20	1	2	10	10	27	56	29	3.0	1.5
Kindergarten Robbins & Ehri (1994)	Story not heard:										
	Lo PPVT	11	na	na	11		note g	26%			
	Mid PPVT	11	na	na	11			27			
	High PPVT	11	na	na	11			34			
	Story heard: 1:11										
	Lo PPVT	11	1 (of 2)	2	11	0	Note g	27%g	1%g	0.1	0.05
	Mid PPVT	11		2	11	0		37	10	1.1	0.55
	Hi PPVT	11	1 (of 2)	2	11	0		55	21	2.3	1.15
			1 (of 2)								
Grade 1 Penno et al (2002)	Repeated reading: 1:11	47	1[b]	3	10	0	32&	37%	5%	0.6	0.2
	Reading + explan: 1:11	47	1[b]	3	10	10	28	53	26	2.7	0.9
Grade 1 Elley (1989)	Repeated reading 1;30	157	1	3	20	0	46%	61%	15%	3.0	1.0
Grade 2 Elley (1989)	Control-no story	51	na	na	20[c]		Pretest	Posttest	2%	(0.2)	0.1
	Repeated reading 1:30	125	1[b]	3	20	0	data not	data not	10	2.0	0.7
	Reading + explan. 1:30	125	1[b]	3	20	20	avail.	avail.	29	5.8	1.9

Grade/Study	Condition	Nu. Chil.	Nu. Books	Nu. of Repeated Readings	Nu. of Words Tested	Nu. Of Words Taught Per day	Pretest Percent Correct [a]	Posttest Percent Correct	Gain Percent Correct	Total Nu. of Words Learned	Average Nu. of Words/ day
Grade 1	Repeated reading:ca 1:20	40	2	3	40	5	29%	34%	5%	2.0	0.3
Brabham &	Reading + discuss										
Lynch-Brown	before/after	38	2	3	40	5	33	45	12	4.8	0.8
(2002)	Reading + discuss										
	before/during/after	39	2	3	40	5	32	57	25	10.0	1.7
Grade 3	Repeated reading:ca 1:20	46	2	3	40	5	40%	52%	12%	4.8	0.8
Brabham &	Reading + discuss										
Lynch-Brown	before/after	41	2	3	40	5	43	64	21	8.4	1.4
(2002)	Reading + discuss										
	before/during/after	42	2	3	40	5	49	76	27	10.8	1.8
Gr. 3 to 5	Single reading, ca 1;19										
Nicholson &	below-avg. reading	18	1	1	10	10	17%	26%	9%	0.9	0.9
Whyte (1992)	avg. reading	19	1	1	10	10	28	45	17	1.7	1.7
	above-avg. reading	20	1	1	10	10	31	54	23	2.3	2.3
Grade 4	Control-No story	61	na	na	20		30%	32%	2%	(0.4)	0.1
Brett, Rothlein	Repeated Reading	56	2	(5)[e]	20	20	29	28	-1	-0.2	0.0
& Hurley (1996)	Reading + Explan.	58	2	(5)	20	20	24	57	33	6.6	1.3
Grade 6	Single reading[g]:										
Stahl, Richek, &	not heard	43	1	0	13		note g.	46.9%			
Vandevier	heard	43	1	1	13	0		53.5	6.6%	0.9	0.9
(1991)											

Note a. All receptive vocabulary tests in involved 4-alternative multiple choice tests. Some used pictures while others used words or both. In all cases, 25% is the random score (i.e., no knowledge of word). There is, of course, some random variation around 25%.

Note b. Two different books were tested in this study. However, each child heard 1 book in each reading condition.

Note c. Data for words from 1 of the books was reported for the control group.

Note d. Pre- and posttest data for 1 of the 2 books was not given, only the gains were reported.

Note e. These stories were considerably longer than those typically read in primary classrooms. Teachers took five days to complete these stories

Note f. In each class, same children heard 12 words from each story without explanation and another 12 words with explanation.

Note g. Half of the children heard book A but not B. The other half heard B. but not A. All were tested on words from A and words from B. No pretest was given. Gains were based on comparison of heard and not heard books.

Note h. Tested words were not explicitly taught, but were sampled from the books read. In essence, this proved to be more of reading without explanation. The “dialogic reading” did increase word acquisition, but at a low rate for the time used

Appendix B

Samples of Scoring Children's Word Meanings

(Note, these examples come from pretests for Study 2 but were the same for Study 1 and the general vocabulary test.)

1. (test item) She told him she **longed** to fly.

Full score (1.0):	"really wanted to"
Half score (0.5):	(no instances)
Wrong response score (0.01):	"takes 2 hours", "flew well", "took a long time"
No answer score (0.0)	"don't know", "pass"

2. Guinivere **barely** gave Merlin any attention.

Full score (1.0)	"not giving lots of", "Merlin doesn't get much"
Half score (0.5)	(no instances)
Wrong response score (0.01)	"giving", "something you like", "something you like"
No answer score (0.0)	"don't know", "pass"

3. Cups, **saucers**, and spoons went flying.

Full score (1.0)	"plate where you put a cup on", "plate under a teacup"
Half score (0.5)	"plate"
Wrong response score ((0.01)	"cups and saucers" (repeats what was supplied)
No answer score (0.0)	"don't know", "pass"

Appendix C

Examples of Pretest Test Sentences, Alternative Review Sentences, and Delayed Test Sentences.

Kindergarten

Word	Definition (*LWV*)	Sentences
obey	Do what they are told	**Pre test** (story sentence) The lions and tigers wouldn't **obey**.
		Review (new sentence) We all try to **obey** the rules.
		Post test (same as pre test)
		Delayed test A and B (new sentence) You really should **obey** your mother.

slip	Put on quickly and easily.	**Pre test** (story sentence) So Clifford **slipped** into an elephant suit.
		Review (new sentence) She **slipped** into her clothes.
		Post test (same as pre test)
		Delayed test A and B (new sentence) **Slip** into your shorts.

Grade 1

Word	Definition (*LWV*)	Sentences
rapidly	Going fast	**Pre test** (story sentence) The vegetable garden was **rapidly** being eaten.
		Review (new sentence) The water was **rapidly** going down the drain.
		Post test (same as pre test)
		Delayed test A and B (new sentence) The water was moving **rapidly**.

solution	Answer to problem	**Pre test** (story sentence) Since I had always wanted a dog it seemed like a good **solution**.
		Review (new sentence) I wonder what the **solution** will be.
		Post test (same as pre test)
		Delayed test A and B (new sentence) That is the **solution**.

Grade 2

Word	Definition (*LWV*)	Sentences
appetite	hungry	**Pre test** (story sentence) Merlin had no **appetite** for the birds and mice she caught.
		Review (new sentence) Around dinner time we usually get an **appetite**.
		Post test (same as pre test)
		Delayed test A and B (new sentence) James did not have much of an **appetite**.

obvious	Easily seen	**Pre test** (story sentence) By late spring, it was **obvious** to all the barn animals that Merlin wasn't doing well.
		Review (New sentence) You were **obvious** in that hiding place.
		Post test (same as pre test)
		Delayed test A and B (New sentence) Next time try not to be so **obvious**.

note: *LWV*, *Living Word Vocabulary* (Dale & O'Rourke, 1981)

Appendix D

Study 2 Word Meanings, Sorted by Percent Possible Gain

KINDERGARTEN

	Pretest	Delay Posttest	Gain	Percent Possible Gain
CIRCUS	61	100	39	100
SNIFF	77	100	23	100
BULLSEYE	06	95	89	95
SPRAIN	47	96	49	93
STRAIGHT	25	93	68	91
LEAD	08	91	83	90
DRESSUP	63	96	33	89
TIP	25	89	64	85
SLIP	08	86	78	85
TRAINER	22	86	64	82
TOGETHER	79	96	17	81
LANDLUBBER	07	79	72	77
PIRATE	71	93	22	75
NEAT	57	89	32	74
MOP	73	93	20	74
COOP	22	77	55	71
QUIT	77	93	16	70
CANARY	07	68	61	65
CACKLE	22	72	50	64
WIRE	45	79	34	62
SPOIL	32	73	41	60
OBEY	16	64	48	57
TWIRL	43	75	32	56
FAULT	68	86	18	56
TIPTOE	68	86	18	56
SIGH	04	57	53	55
ROYALTY	22	64	42	54
THROUGH	18	61	43	52
SPOIL	11	57	46	52
FAIRY	57	79	22	51
HOLLER	08	54	46	50
FINALE	00	50	50	50
GUNPOWDER	11	54	43	48
PARADE	34	66	32	48
AIM	16	47	31	37
FAINT	22	50	28	36
CANNON	15	43	28	33
ACT	20	43	23	29
BOIL	57	68	11	25
RAGE	04	23	19	20
COMMAND	25	39	14	19
GRAND	07	11	04	04

GRADE 1

	Pretest	Delay Posttest	Gain	Percent Possible Gain
DILLYDALLY	08	90	82	89
SHARP	09	88	79	87
DOZEN	07	87	80	86
SCRUFF	14	87	73	85
TIP	42	90	48	83
PILE	72	95	23	82
GROUNDHOG	48	90	42	81
WHENEVER	22	84	62	80
BUTT	08	80	72	78
WICKER	09	79	70	77
WOBBLE	58	90	32	76
OUTING	34	84	50	76
JINGLE	37	83	46	73
SAFELY	63	90	27	73
PADDLE	24	79	55	72
OBSTACLE	19	76	57	70
FIRM	08	71	63	68
RAPID	03	69	66	68
STUMBLE	29	76	47	66
SCENT	06	68	62	66
WHIMPER	15	71	56	66
SHED	24	74	50	66
CONVERSATION	42	80	38	65
FINE	19	71	52	64
MEND	13	69	56	64
TREMENDOUS	56	84	28	64
CERTINLY	29	74	45	63
WRINKLE	32	75	43	63
CHANCE	11	63	52	58
LOYAL	14	63	49	57
HUDDLE	06	57	51	54
MEADOW	08	58	50	54
ACTUALLY	26	66	40	54
SKIM	12	59	47	53
GRATEFUL	19	62	43	53
ESPECIALLY	04	55	51	53
FOND	14	58	44	51
SOLUTION	27	62	35	48
ADMIRE	09	53	44	48
SIGH	10	53	44	48
REALIZE	45	71	26	47
FRANTIC	16	55	39	46
CARTON	19	55	36	45
NUDGE	08	49	41	44
JERK	03	45	42	43
SNAG	04	45	41	43
DWINDLE	03	44	41	42
BOUND	20	53	33	41
BRISK	00	38	38	38
AISLE	35	58	23	36
STRIDE	07	40	33	36
CROUCH	31	54	23	34
PARLOR	02	29	27	28

QUIRK	00	25	25	25
PERCH	03	19	16	16

GRADE 2

	Pretest	Delay Posttest	Gain	Percent Possible Gain
DIRECTION	81	100	19	100
DUMP	41	93	52	88
PEBBLE	35	91	56	86
BASEBOARD	01	85	84	85
PRECIOUS	30	86	56	80
CUDDLE	55	90	35	78
SNEAK	78	95	17	78
SAUCER	10	76	66	73
ORDINARY	48	86	38	73
RUSTLE	11	76	65	73
BRIGHT	13	76	63	72
QUIVER	18	77	59	72
ADOPT	41	83	42	71
LONG	22	77	55	71
HATCH	13	72	59	68
HIDEOUT	48	83	35	67
BARELY	19	73	54	66
DAWN	10	69	59	66
MYSTERIOUS	50	82	32	64
FULLMOON	81	93	12	63
WOOLLY	53	81	28	60
APPETITE	50	79	29	58
BLINDING	50	76	26	52
ALAS	05	53	48	50
VAIN	00	48	48	48
RHYTHM	29	62	33	47
SCOWL	12	53	41	46
MENACE	18	55	37	45
REVEAL	13	52	39	45
ACHE	44	67	23	41
COMFORT	34	60	26	40
CAUTIOUS	16	48	32	38
CEREMONY	15	45	30	36
START	08	39	31	34
YEARN	19	46	27	33
LOUNGE	24	49	25	33
WAVER	12	41	29	33
SCRAWNY	10	38	28	31
LURE	12	38	26	29
GLANCE	26	48	22	29
WAVER	17	41	24	29
RESTLESS	40	57	17	29
ERUPT	19	41	22	27
ENVY	03	24	21	22
OBVIOUS	18	35	17	21
PANTRY	03	19	16	17

Table 1.

Study 1: *Percentage of LWV Vocabulary in the Easiest to Most Difficult Deciles of Grade 2 to 12 Word Meanings by Grade and Normative[a] vs. Current Study 1 Children.*

		Living Word Level						
	N	2	4	6	8	10	12	Pct. of 40 Words.
Kindergarten								
Normative	22	56.9%	15.4	15.7	1.6	0.8	3.5	18
Study 1	43	41.0	11.2	11.3	0.3	2.4	3.1	12
Grade 1								
Normative	18	66.7	20.8	24.1	2.6	4.2	5.5	21
Study 1	38	65.9	25.8	28.3	1.5	4.2	7.9	23
Grade 2								
Normative	25	74.7	55.1	46.3	12.1	13.2	17.5	37
Study 1	32	62.5	26.4	29.8	2.7	4.0	6.6	23

Note a. "Normative" children refers to data obtained from a normative population on the same items from the study reported by Author (2001).

Table 2

Study 1: Percentages of Words Correct by Classroom, Grade, and Instructed vs. Not Instructed.

(Standard deviations in parentheses)

		Instructed Words			Non-Instructed Words		
	N	Pretest	Posttest	Gain	Pretest	Posttest	Gain
Kindergarten	43	16% (14)	35 (25)	20 (19)	30 (16)	38 (22)	08 (16)
read 2 times		16	32	16	23	29	06
read 4 times		15	38	23	36	45	09
Grade 1	37	25 (16)	49 (24)	24 (21)	24 (16)	39 (19)	14 (15)
read 2 times		32	56	24	37	43	06
read 4 times		18	42	24	15	35	20
Grade 2	32	30 (15)	53 (23)	23 (19)	25 (18)	43 (23)	17 (14)
read 2 times		32	56	24	32	53	21
read 4 times		28	50	22	18	33	15
All	112	23 (16)	45 (25)	22 (19)	27 (16)	40 (21)	12 (15)
read 2 times		26	47	21	30	41	11
read 4 times		20	43	24	24	38	15

Table 3.

Study 1: Percentages of Posttest Words Known in Each Grade by Cohort, Instruction/Not, and Lists of Words. (List A was pretested by Cohort 1 while List B was pretested by Cohort 2. Pretested scores are italicized)

		Instructed Words		Not Instr. Words		All Words	
	N	List A	List B	List A	List B	List A	List B
Kindergarten							
Cohort 1	22	*28%*	32	*34*	30	*31*	31
Cohort 2	21	42	*42*	40	*41*	41	*42*
Grade 1							
Cohort 1	17	*59*	39	*34*	44	*47*	42
Cohort 2	20	60	*39*	34	*43*	47	*41*
Grade 2							
Cohort 1	17	*49*	54	*47*	38	*48*	46
Cohort 2	15	51	*58*	50	*38*	51	*48*

Table 4.

Study 2: Percentages of LWV Vocabulary in the Easiest to Most Difficult Deciles of Grade 2 to 12 Word Meanings by Grade and Normative[a] vs. Current Study 2 Children.

		Living Word Level						
	N	2	4	6	8	10	12	Pct. of 40 Words.
Kindergarten								
Normative	22	56.9%	15.4	15.7	1.6	0.8	3.5	18
Study 2	28	57.1	24.6	10.5	0.0	0.2	5.2	19
Grade 1								
Normative	18	65.7	25.6	22.6	2.1	6.8	2.0	21
Study 2	38	66.7	20.8	24.1	2.6	4.2	5.5	23
Grade 2								
Normative	25	74.7	55.1	46.3	12.1	13.2	17.5	37
Study 2	42	66.0	29.1	25.4	1.0	5.5	5.7	25
Study 2 (contr.)	11	62.5	19.7	20.9	0.0	0.1	3.1	20

Note a. "Normative" children refers to data obtained from a normative population on the same items from the study reported by Author (2001).

Table 5.

Study 2: Mean percentages of words correct by pre-, immendiate post-, and delayed post-tests and gains.

.		Mean Percent			Gains	
	N	Pretest	Posttest	Delay	Post	Delay
Kindergarten	28	33% (14)	64 (19)	72 (14)	32 (15)	39 (13)
boys	12	34%	62	74	28	40
girls	16	32%	66	71	35	39
Grade 1	38	20% (11)	62 (22)	65 (20)	42 (17)	45 (16)
boys	21	19%	59	63	39	44
girls	17	21%	66	68	45	47
Grade 2	42	28% (16)	58 (21)	65 (18)	30 (14)	37 (13)
boys	21	32%	62	68	30	37
girls	21	24%	53	61	30	38
All Grade	108	26% (14)	61 (23)	67 (18)	35 (16)	41 (14)
No Interv. Gr. 2 Contr. Grp.	11	19% (11)	23 (11)	na	4 (6)	na
boys	5	23%	24	na	1	na
girls	6	15%	21	na	6	na

Table 6.

Study 2: Effects of new versus previously used context sentences for percentage of correct delayed posttest word meanings. (Data from *new* test sentences are *italicized*.)

	Wordlist C	Wordlist D
Kindergarten		
Cohort 1	73%	*71 (new)*
Cohort 2	*70 (new)*	68
Grade 1		
Cohort 1	69	*63 (new)*
Cohort 2	*70 (new)*	59
Grade 2		
Cohort 1	70	*53 (new)*
Cohort 2	*68 (new)*	69
All Grades		
Cohort 1	70	*61 (new)*
Cohort 2	*69 (new)*	65