2007

Fostering the Development of Vocabulary Knowledge and Reading Comprehension Though Contextually-Based Multiple Meaning Vocabulary Instruction

J. Ron Nelson
University of Nebraska - Lincoln, rnelson8@unl.edu

Scott A. Stage
University of Washington, scott_stage@ncsu.edu

Fostering the Development of Vocabulary Knowledge and Reading Comprehension Though Contextually-Based Multiple Meaning Vocabulary Instruction

J. Ron Nelson

University of Nebraska, Lincoln

Scott A. Stage

University of Washington

Abstract

The primary purpose of this study was to assess the effects of contextually-based multiple meaning (i.e., words with multiple meanings) vocabulary instruction on the vocabulary knowledge and reading comprehension of students. Third and 5th grade students received either contextually-based multiple meaning vocabulary instruction embedded in the standard language arts instruction offered to all students over a three-month period or the standard language arts instruction alone (i.e., non-specific treatment). Students who received the contextually-based multiple meaning instruction generally showed statistically and educationally significant gains in their vocabulary knowledge and reading comprehension relative to students who did not. These gains were most evident in reading comprehension. Additionally, students with low initial vocabulary knowledge and reading comprehension achievement tended to show greater gains than those with average to high achievement. These effects were more pronounced in the case of 3rd grade students. The results and limitations are discussed.

Vocabulary knowledge plays a critical role in people's lives and future possibilities (Beck, McKeown, & Kucan, 2002). In fact, "it is clear that a large and rich vocabulary is the hallmark of an educated individual. A large vocabulary repertoire facilitates becoming an educated person to the extent that vocabulary is strongly related to reading comprehension in particular and school achievement in general" (Beck et al., p. 1).

Armbruster, Lehr, and Osborne (2003) reported two ways in which vocabulary is learned: indirect and direct vocabulary instruction. Indirect vocabulary building pertains to learning words

Correspondence to J. Ron Nelson, Ph.D., Center for At-Risk Children's Services, 202 Barkley Center, Lincoln, NE 68583-0732; e-mail: rnelson8@unl.edu.

primarily through exposure—through conversations with others, being read to, or reading on one's own (Beck et al., 2002; Cunningham & Stanovich, 1998; Nagy, Herman, & Anderson, 1985). Thus, the more children participate in rich oral and reading vocabulary experiences, the greater their vocabulary knowledge and reading comprehension (Greene & Lynch-Brown, 2002; Robbins & Ehri, 1994; Stahl, Richek, & Vandevier, 1991). Unfortunately, we know that many children may have limited indirect experiences in vocabulary development for a variety of reasons (see Hart & Risley, 1995 for further details).

Students also learn vocabulary directly through explicit instruction. Researchers have studied the effects of a wide range of explicit approaches to vocabulary instruction across the Pre-K-11th grades. These instructional approaches include: (a) key word (Levin, Levin, Glassman, & Nordwall, 1992; Levin, McCormick, Miller, & Berry, 1982), (b) repeated multiple readings (Leung, 1992; Senechal, 1997), (c) rich contexts (Kameenui, Carnine, & Freschi, 1982; McKeown, Beck, Omanson, & Pople, 1985), (d) computer-based (Heise, Papelweis, & Tanner, 1991; Heller, Sturner, Funk, & Feezor, 1993), (e) pre-instruction (Brett, Rothlein, & Hurley, 1996; Carney, Anderson, Blackburn, & Blessing, 1984), and (f) restructuring the task (Malone & McLaughlin, 1997; Scott & Nagy, 1997). Overall, the results of this research suggest that explicit vocabulary instruction methods improve vocabulary knowledge and reading comprehension, and the effects are greatest for students with low initial vocabulary knowledge levels (see Fukkink & deClopper, 1998; Klesius & Searls, 1990; NICHD, 2000; Stahl & Fairbanks, 1986 for reviews of the vocabulary instruction research literature).

Although researchers have applied a wide range of explicit instruction approaches to teach a single meaning for a word, it appears that to date they have not yet attempted to teach students the multiple meanings for a word. Explicitly teaching students that most words they encounter have multiple meanings that may fall into different semantic categories (e.g., verb, noun, adjective), depending upon the context in which they are used (Anderson & Nagy, 1991; Biemiller, 1999; Chall & Dale, 1995; Dale & O'Rourke, 1981), should have a positive effect on reading comprehension because it would encourage students to attend more closely to contextual clues that influence word meanings. Further, teaching students that the meanings and semantic categories of many words are influenced by context would provide them a word learning strategy that could be used beyond the words being taught. Word-learning strategies are helpful because we cannot teach students the definition of every word they will encounter.

Armbruster et al. (2003) recommend teaching word meanings

in context (also supported by the NICHD, 2000). In addition to other instructional activities (e.g., word histories, semantic mapping), contextually-based instruction begins with the presentation of word meanings in clear and simple language. Paired with this presentation, every meaning of every word is clearly illustrated in sentences to help the learner grasp the different meanings conveyed by a word. The student then produces contextually relevant sentences that depict word meanings. In much the same way that it is important to provide multiple exposures to a word and associated meanings, the context in which a word is learned is critical (McKeown et al., 1985). The learner is more likely to learn the meanings for words if they are presented in contextually-relevant sentences that convey the different meanings for words.

The use of contextually-based instruction is especially true for words that have multiple meanings (Beck et al., 2002). For example, knowing that *just* means "based on reason and fairness" will provide little help when *just* is used in the following ways: "Jason just broke the high school track record in the 100-yard dash" (the meaning now relates to "little time or distance") or "My parents said that whatever movie we wanted to rent was just fine" (the meaning now relates to "simply"). Additionally, students understanding of the meanings for a word are strengthened when they are asked to generate contextually-correct written narratives using its meanings (Armbruster et al., 2003).

Along these lines, Stahl (1999) discussed the importance of using examples and non-examples when teaching vocabulary, particularly during guided practice with words like those with multiple meanings. In terms of instruction, the teacher may pose scenarios where vocabulary words and their meanings are used as expected (examples) or not used as expected (nonexamples). Interestingly, Stahl and Fairbanks (1986) noted that approaches providing only definitional information did not significantly affect students' reading comprehension. In contrast, methods that provided both definitional and contextual information did significantly improve comprehension. It would appear that contextually-based multiple meaning vocabulary instruction should have a positive effect on the reading comprehension of students. Stahl (1999) also called for semantic mapping activities where students practice mapping vocabulary words to other related words. These activities expand students' knowledge base and teach students that words and their meanings are highly interrelated. "It is generally beneficial to teach words so that students learn more than just single words." (Stahl, 1999, p. 51). Thus, related words practice is critical in vocabulary development.

The primary purpose of this study was to assess the effects of contextually-based multiple meaning vocabulary instruction on the vocabulary knowledge and reading comprehension of students. Classrooms received either contextually-based multiple meaning vocabulary instruction embedded in the standard language arts instruction offered to all students or the standard language arts instruction alone (i.e., non-specific treatment) and were pre- and post-tested using a standardized assessment of vocabulary knowledge and reading comprehension. The target words and associated related words taught were systematically chosen to ensure that they were used frequently and widely across content areas as well as the 3rd through 6th grade levels (selection procedures are described below).

Method

Participants

A total of 283 third ($n = 134$) and fifth ($n = 149$) grade students enrolled in a small Midwestern public school system were participants. The students were drawn from 16 third ($n = 8$) and fifth ($n = 8$) grade classrooms. Third and fifth grade classrooms were randomly assigned to an experimental or non-specific treatment condition. To estimate treatment effects, all students who were present at pre- and post-testing and who had appropriate test protocols were sampled. Appropriate test protocols were those in which students made a reasonable attempt to complete the test. For example, a protocol in which the student did not respond or responded inappropriately (the same response for all items) were not included. This resulted in a total of 134 (or 86% of 156 students) third grade and 149 (or 84% of 168 students) fifth grade students. Approximately 32% of the students qualified for free or reduced lunch.

Students were classified into two groups based on their initial overall vocabulary and comprehension achievement. Students' initial overall vocabulary and comprehension achievement was based on their pre-test Gates-MacGinitie Reading Tests (4th Edition) (GMRT) Total scale normal curve equivalent (NCE) score (MacGinitie, MacGinitie, Maria, & Dreyer, 2000): low (NCE $\leq$ 30) and average to high (NCE >30). The Total scale score is a composite of the GMRT Vocabulary and Comprehension scale scores (described below). A separate high group was not established because few students ($n = 7$) received scores more than one standard deviation above the mean. Student race, language status, and special education status by experimental condition and students' initial vocabulary and comprehension achievement status are presented in Table 1. With one exception, chi-square analyses on these nominal data showed no effect for condition. The difference in

Table 1
Student Demographics By Achievement Status and Condition

Grade/ Status/ Variable			Condition			
			Experimental		Non-Specific	
			N	(%)	N	(%)
Third			71		63	
A to H	Sex (male)		25	(35)	22	(35)
	Race	European American	34	(48)	33	(52)
		Hispanic	9	(13)	6	(10)
		Other	3	(4)	1	(2)
	English Second Language		9	(13)	6	(10)
	Special Education		—	—	—	—
Low	Sex (male)		16	(23)	14	(22)
	Race:	European American	13	(18)	13	(21)
		Hispanic	10	(14)	9	(14)
		Other	2	(3)	1	(2)
	English Second Language		10	(14)	9	(14)
	Special Education		8	(11)	6	(10)
Fifth			72		77	
A to H	Sex (male)		38	(53)	39	(51)
	Race:	European American	51	(71)	60	(78)
		Hispanic	14	(19)	15	(19)
		Other	7	(10)	2	(3)
	English Second Language		14	(19)	15	(19)
	Special Eduction		2	(3)	1	(1)
Low	Sex (male)		10	(14)	6	(8)
	Race:	European American	5	(7)	8	(10)
		Hispanic	9	(13)	1	(1)
		Other	2	(3)	—	—
	English Second Language		9	(13)	1	(1)
	Special Education		5	(7)	5	(8)

Note. A to H = average to high. Percentages based on the number in each respective group at each grade level.

the percentages of fifth grade European Americans and Hispanic students who were in the low initial overall vocabulary and reading comprehension in the experimental and non-specific treatment conditions was statistically significant (X^2 = 7.73, 2, N = 25, p < .05). Students in the experimental condition were less likely to be of European American descent than those in the non-specific treatment condition.

Selection of Words

Two interrelated processes were used to identify two levels of multiple meaning target words: Level I (i.e., words with two mutually exclusive meanings) and Level II (i.e., words with three or four mutually exclusive meanings). The commonly accepted "stages of vocabulary knowledge" (Dale & O'Rourke, 1986) were used in the first process to identify multiple meaning target words. The four stages include:

1. I never saw the word before.
2. I've heard of the word, but I don't know what it means.
3. I think I know it—it has something to do with.
4. I know the word—it means "…" in this context.

The aforementioned stages provided the initial criterion for selecting multiple meaning target words from *The Living Word Vocabulary* (Dale & O'Rourke, 1981). *The Living Word Vocabulary* is a national level vocabulary inventory of the familiarity of 44,000 words represented by their meanings. This vocabulary inventory provides objective familiarity scores for students in grades 4, 6, 8, 10, 12, 13, and 16 on each of the 44,000 word meanings. The target words included in Levels I and II were those that students in the 4th through 6th grades are likely to struggle with (Stages 1-3). Thus, the typical target word included two to four meanings, some of which students may have some (Stage 3) or little (Stage 2) familiarity with one or more meanings, and some of which students were unlikely to know at all (Stage 1).

The Educator's Word Frequency Guide (Zeno, Ivens, Millard, & Duvvuri, 1995) was used in the second process to further refine the list to ensure that the target words and associated related words (i.e., 3 per meaning) were used frequently and widely across content areas as well as the 3rd through 6th grade levels. This guide provides a frequency and breadth index for words used in written text based on 17 million words. The final words used in the study were included in the 1000 most frequently and widely used words in 3rd through 6th grades. The two interrelated processes resulted in a total of 80 target words at each level, and approximately 480 (Level I) and 800 (Level II) related words. The Level I and II words and associated related words along

with the corresponding frequency and breadth indices are available from the first author.

Design, Core Instruction, and Conditions

Design. A pre/post experimental and non-specific treatment group design was used to assess the effects of the multiple meaning vocabulary instruction on the vocabulary knowledge and reading comprehension skills of students. Sixteen third (n = 8) and fifth (n = 8) grade classrooms were randomly assigned to the experimental conditions. The four 3rd and 5th grade classrooms assigned to the experimental condition received contextually-based multiple meaning vocabulary instruction on the Level I and II words, respectively. All instruction was delivered by classroom teachers across a contemporaneous four month time span. Teachers self-selected the words they believed to be most relevant to their students. The four 3rd and 5th grade classrooms assigned to the non-specific treatment condition received the standard language arts instruction provided to all students. No attempt was made to alter the instructional practices of teachers. Additionally, the multiple meaning vocabulary instruction was embedded in the time normally allotted for standard language arts instruction in the experimental classrooms.

Core instruction. All 16 teachers used and followed the district's core language arts curriculum. Teachers used the Scott Foresman Basal Reading program (Scott Foresman, 2001) to guide their instructional activities each week. No direct observations were conducted to describe or contrast the vocabulary and reading comprehension instructional activities used by teachers.

Experimental condition. Students in the experimental condition received contextually-based multiple meaning vocabulary instruction on 36 target words (SD = 2.4) as well as three related words per meaning. This instruction was embedded within the language arts instruction they provided to students. Third grade students received the Level I words (i.e., two mutually exclusive meanings); whereas fifth graders received the Level II words (i.e., three or four mutually exclusive meanings). Teachers guided students through the instructional activities (described below) using overhead masters of each activity. Students completed each activity in their student workbook. Teachers participated in a two hour training session. A three-step training process was used to train educators to implement the contextually-based multiple meaning vocabulary instruction. First, the trainer (first author) provided educators an overview of the rationale and formats for the instructional activities. Second, the trainer modeled and practiced the instructional activities with educators. Finally, following training,

a question and answer session was conducted approximately one month after teachers began to implement the contextually-based multiple meaning vocabulary instruction to address implementation issues.

Each target word and associated set of related words was taught over 2 days for approximately 20-30 minutes each day. The meanings for the target words were presented nine times in six varied contextually-based instructional contexts. These contextually-based learning opportunities began in the pre-lesson activity (Day 1) and extended across all the instructional activities included in the lesson (Day 2). The set of contextually-based instructional contexts are available from the first author.

On Day 1, the meanings of each of the target words were introduced through related words to activate students' prior knowledge in a pre-lesson activity (entitled, **"**Meanings of Related Words**"**). For example, *accident* has two meanings: (a) unexpected happening and (b) event that causes damage. Thus, students' prior knowledge for the meanings of the target word was activated by getting a chance to learn words that were related to the two meanings of *accident*. For example, one meaning of accident (i.e., unexpected happening) has the following related words: *fluke, mishap,* and *by chance*. Students examined and discussed sentences that used the related words in context with the teacher ("Regina's plane was delayed. This mishap caused her to miss her sister's party") and then wrote sentences of their own using the related word ("Write a sentence using *mishap*").

On Day 2, the first activity was labeled "Word Meaning in Context.**"** This activity began with the word history of the target word. For example, using the word *accident*, "It all began with a Latin phrase meaning 'fall.' Later in the English language, this meaning changed to its two current meanings to include 'unexpected happening' and 'event that causes damage.'" Students then practiced their knowledge of these meanings within the context of sentences using each of the meanings.

Next, a "Word Meaning Map" activity was conducted. In this exercise, students matched the related words that appeared in the pre-lesson activity (covered in Day 1) with the appropriate meanings of the target word. These meanings appeared in a graphic organizer format with spaces provided for the students to write the related words. Next, students completed a definition activity for the multiple meanings of the target word in a section entitled "Complete Each Definition." Following this short exercise, there was an "Understanding Check" where students examined short reading passages to see if the target word was used as they *expected* or *not expected*. For example, "Jasmine worked hard to earn enough money to buy a new car. Jasmine's new

car was perfect in every way. It looked like it had been in an *accident*." This sentence would be labeled as *not expected* by students. Finally, students wrote short stories or scenarios using each of the meanings of the target word in a section entitled "Create Stories."

Non-specific treatment condition. Students in the non-specific treatment condition received the core language arts instruction offered in the classroom. No attempt (staff development activities directed at vocabulary and reading comprehension development) was made to change any of the language arts instructional activities provided to students by teachers.

Dependent Measure

The GMRT (4th Edition) (MacGinitie et al., 2000) was used to measure students' vocabulary and reading comprehension skills. The GMRT is a timed multiple-choice test administered in groups. Levels 3 and 5 were used to assess third and fifth grade students' vocabulary and reading comprehension skills, respectively. Furthermore, alternate forms were used at pre- (Form S) and post-testing (Form T). The alternative form reliabilities of the different levels of Forms S and T were .90 or higher (MacGinitie et al., 2000). The GMRT Vocabulary scale is a test of vocabulary knowledge. The student chooses the word or phrase that means most nearly the same as the test word. The administration time for the GMRT Vocabulary scale is 30 minutes. The GMRT Comprehension scale consists of fiction and nonfiction prose passages. The passages are drawn from various content areas and written in a variety of styles. The administration time for the GMRT Comprehension scale is 50 minutes. The test-retest reliabilities of all levels and forms of the GMRT Vocabulary and Comprehension scales ranged from .58 to .91 with only two coefficients below .70 (MacGinitie et al., 2000). All analyses were based on the GRMT NCE Vocabulary and Comprehension scale scores. NCEs are normalized transformations of percentile rank scores in which the range is divided into 99 equal parts with a mean of 50 and a standard deviation of 21.06.

Treatment Fidelity

Teacher self-evaluations and permanent product assessment of lessons (i.e., completed student worksheets) were used to assess treatment fidelity. Both measures assessed the total number of program components implemented correctly. The program elements included eight items that focused on the specific instructional activities: (1) following the two day lesson sequence, (2) using the pre-lesson activity, (3) reviewing and discussing the meanings of the target words with students during the pre-lesson activity, (4) writing sentences for each

of the related words, (5) reviewing and discussing the word history with students, (6) completing the word meaning map, (7) completing the understanding check, and (8) writing a short story for each of the meanings for a target word. Teachers completed a self-evaluation during the 10th week. Teachers rated the extent to which they completed the eight components for each lesson on a 4-point Likert-type scale (i.e., 1 = never, 2 = sometimes, 3 = usually, 4 = always).

The permanent product assessment included 20 completed student lessons randomly selected from a group of low achieving (n = 10) and average to high achieving (n = 10) students. Equal numbers of third and fifth grade student lessons were assessed. A three-step procedure was used to select the completed student lessons. In the first step, teachers collected all student essays during the 6th week of the intervention. The teachers sorted the completed lessons into two groups depending on their pre-intervention skill levels (i.e., low achieving and average to high achieving) and removed all identifying information in the second step. The first author then selected randomly the completed student lessons from the two groups of papers.

Two independent raters assessed the number of components completed by students to establish inter-rater agreement. The raters were graduate students in education and were unaware of the purpose of the study. Inter-rater agreement was calculated by taking the percentage of agreements divided by the total number of agreements and disagreements multiplied by 100. Inter-rater checks conducted on the 20 completed student lessons was 100%.

Teachers' Perception of Efficacy

Teachers completed a questionnaire that focused on their perceptions of the efficacy of the contextually-based multiple meaning vocabulary instruction. Teachers responded on a 5-point Likert-type scale (i.e., 1 = strongly disagree, 2 = disagree, 3 = undecided, 4 = agree, 5 = strongly agree) to the following four efficacy items: (1) The exercises challenged students; (2) Students learned key vocabulary knowledge and reading comprehension skills; (3) Students can apply the lesson content in other areas; and (4) Students responded enthusiastically to the lessons.

Results

Treatment Fidelity

With one exception (i.e., respondent indicated "usually"), all teachers indicated that they "always" (a) followed the two day lesson sequence, (b) used the pre-lesson activity, (c) reviewed the meanings of the target words with students during the pre-lesson activity,

(d) had students write sentences for each of the related words, (e) had students complete the word meaning map, and (f) had students complete the understanding check ($X = 3.9$ and $SD = .31$ in all cases). Teachers' mean responses on the two remaining components—(1) reviewed and discussed the word history with students and (2) had students write a short story for each of the meanings for a target word—were 3.8 ($SD = .42$) and 3.7 ($SD = .48$), respectively. Permanent product assessments of lessons completed by students revealed that the percentage of program components implemented correctly was 100% in all cases.

Pre-treatment Vocabulary Knowledge and Reading Comprehension Skills Levels

The pre-treatment means and standard deviations are presented in Table 2. A Condition (Experimental, Non-Specific Treatment) X Level (Low, Average to High) X Grade (Third, Fifth) Multivariate Analysis of Variance (ANOVA) applied to the pre-treatment GMRT Vocabulary and Comprehension scores revealed no statistically significant pre-treatment effects involving condition: for condition, $F(1, 290) = 0.21$, $p > .05$; for condition by initial achievement status interaction, $F(1, 290) = 1.56$, $p > .05$; for the condition by measure interaction, $F(1, 290) = 1.65$, $p > .05$; and for the 3-way interaction, $F(1, 290) = 1.73$, $p > .05$. Taken together, these results demonstrate the comparability of the treatment groups in terms of the pre-treatment vocabulary knowledge and reading comprehension skills of students.

Changes in Vocabulary Knowledge and Reading Comprehension Skills

The mean NCE pre-test, post-test, and mean change GMRT Vocabulary and Comprehension scale scores for the experimental and non-specific treatment conditions for third and fifth grades, as well as for the overall sample by experimental condition, are presented in Table 2. The mean changes in the experimental and non-specific treatment conditions on the vocabulary and reading comprehension measures were analyzed in Condition (Experimental, Non-Specific Treatment) X Level (Low, Average) X Grade (Third, Fifth) X Change (Pre-treatment, Post-treatment) ANOVAs, with the latter variable being a within-subject factor. Follow-up Newman-Kuel post hoc tests, appropriate for within-subjects analyses (Ferguson & Takane, 1989), were applied when appropriate. Additionally, effect sizes, corrected for the intercorrelation between the pre- and post-test scores, were calculated by dividing the difference between the experimental and non-specific treatment condition mean change scores by the pooled standard deviation of the improvement scores (Hedges & Olkin, 1985).

Table 2
Mean NCE Pre, Post, and Improvement Scores of Students for the Overall Sample and by Grade Level

Scale/Status/Trial	Third Grade				Fifth Grade				Overall Sample			
	Experimental		Non-Specific		Experimental		Non-Specific		Experimental		Non-Specific	
	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Vocabulary												
A to H Pre	55.69	(17.28)	52.34	(15.25)	56.43	(17.35)	55.83	(12.01)	56.07	(17.25)	54.81	(13.90)
Post	57.56	(15.84)	55.19	(15.24)	55.95	(19.79)	56.08	(13.14)	56.23	(17.92)	55.75	(14.77)
Improve	1.86	(16.83)	2.85	(9.76)	-0.48	(9.44)	0.24	(11.91)	0.17	(13.63)	0.94	(11.48)
Low Pre	14.24	(9.53)	16.90	(7.34)	13.70	(9.24)	18.31	(10.54)	13.95	(9.24)	17.51	(8.75)
Post	27.53	(13.30)	25.90	(12.73)	20.75	(11.46)	23.94	(10.52)	23.86	(12.63)	25.05	(11.72)
Improve	13.29	(14.18)	9.00	(15.80)	7.05	(11.08)	5.62	(9.53)	9.92	(12.81)	7.54	(13.39)
Reading Comprhension												
A to H Pre	51.48	(12.67)	49.07	(10.84)	49.04	(11.83)	52.96	(12.01)	55.85	(15.23)	54.81	(13.90)
Post	65.15	(14.15)	57.66	(12.19)	56.52	(11.79)	61.07	(13.14)	59.47	(15.48)	55.75	(14.77)
Improve	13.67	(8.79)	8.89	(7.74)	7.48	(6.15)	8.11	(7.85)	3.61	(11.74)	0.94	(11.48)
Low Pre	15.88	(7.89)	16.11	(7.34)	18.71	(8.09)	21.00	(11.83)	17.66	(7.94)	18.67	(11.83)
Post	40.94	(13.56)	31.33	(12.73)	35.64	(7.92)	32.50	(11.79)	35.09	(14.71)	29.22	(11.79)
Improve	25.06	(11.51)	16.22	(14.80)	16.93	(12.49)	11.50	(10.92)	17.43	(14.53)	10.56	(10.92)

Note. A to H = average to high

Table 3
Effect Sizes by Grade and Achievement Status

Scale/Status	Third Grade			Fifth Grade			Overall Sample		
		95% C.I.			95% C.I.			95% C.I.	
	Effect Size	Lower	Upper	Effect Size	Lower	Upper	Effect Size	Lower	Upper
Vocabulary	-.07	-.42	.28	-.07	-.42	.29	-.06	-.30	.18
A to H	.28	-.07	.64	.14	-.22	.49	.18	-.06	.42
Reading Comprehension									
A to H	.57	.21	.93	-.08	-.44	.27	.23	-.01	.47
Low	.67	.31	1.03	.46	.10	.82	.53	.29	.77

Note. A to H = average to high. C.I. = Confidence Interval

The obtained estimates were then corrected for bias due to sample size using a factor provided by Hedges and Olkin (1985). The 95% confidence bands for the effect sizes (*ES*) were also computed using percentiles from the standard normal distribution and the asymptotic variance of the standardized mean difference (Hedges & Olkin, 1985). The obtained effect sizes and associated 95% confidence intervals for the third and fifth grades, as well as for the overall sample by achievement status, are presented in Table 3. Effect sizes in the range of 0 to 0.3 are considered small, 0.3 to 0.8 are considered moderate, and 0.8 and above are considered large (Cohen 1988).

Vocabulary knowledge. Students with low initial vocabulary and comprehension achievement in the experimental condition showed small improvements in their vocabulary skills relative to students in the non-specific treatment condition (see Tables 2 and 3). Students who were average to high achieving in the experimental and non-specific conditions showed negligible changes in their vocabulary skills pre- to post-treatment. A statistically significant main effect for Change was obtained (F (1, 285) = 34.07, $p < .001$). This result revealed that students generally showed improvements in their vocabulary skills from pre- to post-treatment. Follow-up Newman-Kuels post hoc tests to the obtained statistically significant Change by Level interaction (F (1, 285) = 20.35, $p < .001$) revealed that students in the low achieving group were more likely to show improvements in their vocabulary skills than those who were in the average to high group. The relative effect sizes for students who were low and average to high achieving were .28 vs. -.07 (3rd grade), .14 vs. -.07 (5th grade), and .18 vs. -.06 (overall sample). Furthermore, follow-up Newman-Kuel post hoc tests to the obtained statistically significant Change by Grade interaction (F(1, 285) = 6.10, $p < .05$) showed that third grade students with low initial vocabulary and comprehension achievement were more likely to show improvements in their vocabulary skills than those in the fifth grade. There were no other statistically significant main or interaction effects.

Reading comprehension skills. Students in the experimental condition showed moderate to large improvements in their reading comprehension skills relative to students in the non-specific treatment condition (see Tables 2 and 3). A statistically significant main effect for Change was obtained (F (1, 285) = 34.07, $p < .001$). This result revealed that students generally showed improvements in their reading comprehension skills from pre- to post-treatment. Follow-up Newman-Kuels post hoc tests to the statistically significant Change by Condition interaction (F(1, 285) = 10.68, $p < .01$) showed that, with the exception of average to high achieving fifth graders, students in the experimental condition were more likely to show improvements in

their reading comprehension skills than students in the non-specific treatment condition. The obtained effect sizes for students with low initial vocabulary and comprehension achievement in the third and fifth grade were .67 and .57, respectively. In contrast, the resulting effect sizes for students who were average to high achieving in the third and fifth grade were .46 and -.08, respectively. Furthermore, follow-up Newman-Kuels post hoc tests to the statistically significant Change by Level interaction ($F(1, 285) = 20.76$, $p < .001$) revealed that fifth grade students with low initial vocabulary and comprehension achievement in the experimental condition were more likely to show improvements in their reading comprehension skills than those who were average to high achieving. Students who were low and average to high achieving in the third grade both showed statistically equivalent improvements in their reading comprehension skills. The obtained effect sizes for low and average to high achieving students in the overall sample were .53 and .23, respectively. There were no other statistically significant main or interaction effects.

Teachers' Perception of Efficacy

Teachers consistently rated the efficacy of the contextually-based multiple meaning vocabulary instruction as high. Teachers reported that they thought the contextually-based multiple meaning vocabulary instruction challenged their students ($X = 4.82$: $SD = .45$), helped students learn key vocabulary knowledge and reading comprehension skills ($X = 4.62$: $SD = .55$), lesson content could be applied by students in other areas ($X = 4.48$: $SD = .69$), and that students responded enthusiastically to the lessons ($X = 4.01$: $SD = .72$). The 95% confidence intervals for each of the means were calculated to establish whether teachers were significantly resolute rather than indecisive or neutral about the efficacy of the contextually-based multiple meaning vocabulary instruction. In all cases the 95% confidence interval failed to encompass the midpoint of the Likert-type scale (3=undecided), indicating no teacher indecisiveness regarding the program's efficacy.

Discussion

The primary purpose of this study was to assess the effects of contextually-based multiple meaning vocabulary instruction on the vocabulary knowledge and reading comprehension of students. Effects were studied on third and fifth grade students with low and average to high initial vocabulary and comprehension achievement. Third and fifth grade students with low initial vocabulary and comprehension achievement who received the contextually-based multiple meaning vocabulary instruction showed statistically significant gains

in their vocabulary knowledge. The magnitude of the improvements (i.e., effect sizes) for students with low initial vocabulary and comprehension achievement skills were small ($ES =.28$ and $.14$ for 3rd and 5th grade, respectively). In contrast, third and fifth grade students with average to high initial vocabulary and comprehension achievement who received the supplemental vocabulary instruction did not show statistically or educationally significant gains in their vocabulary knowledge relative to their counterparts in the non-specific treatment condition.

It is plausible that the obtained relatively modest or no change in students' overall vocabulary knowledge was a function of the relatively small number of words taught to students. The supplementary vocabulary instruction was only taught to students over a contemporaneous four month time span. Thus, on average students were only taught 36 *(SD* = 2.4) target words and approximately 220 (Level I) to 350 (Level II) related words. Students may have shown greater gains in their general vocabulary knowledge if teachers had provided the students the contextually-based multiple meaning vocabulary instruction for the entire year. It is also plausible that students who received the multiple meaning vocabulary instruction would have shown greater changes in their vocabulary knowledge relative to those in the non-specific treatment condition if a more direct measure of the words taught had been used.

Third grade students with low and average to high initial vocabulary and comprehension achievement who received the contextually-based multiple meaning vocabulary instruction showed statistically significant gains in their reading comprehensions skills. The magnitude of the improvements was moderate for students with low ($ES = .67$) and average to high ($ES = .57$) initial vocabulary and comprehension skills. In contrast, fifth grade students with low initial vocabulary and comprehension achievement showed statistically significant gains in their reading comprehension skills; whereas, those with average to high initial vocabulary and comprehension achievement did not. The magnitude of the improvements for students with low initial vocabulary and comprehension achievement was moderate ($ES = .46$); whereas, students with average to high initial vocabulary and comprehension achievement showed small negative effects ($ES = -.08$). Overall, the generally moderate improvement in students' reading comprehension skills relative to vocabulary knowledge was expected. The multiple meaning vocabulary instruction was designed to enhance students awareness of the complexity of words (i.e., multiple meanings and semantic category of meanings is dependent upon context) and to encourage them to more carefully consider contex-

tual information. The number of exposures to words with multiple meanings to achieve such awareness may not be as large as that required to build vocabulary knowledge. Of course, it is possible that students may have shown greater gains if they had been exposed to the contextually-based multiple meaning vocabulary instruction for a longer period of time.

The mixed outcomes of this study for students of differing vocabulary and comprehension levels are generally consistent with the body of research on vocabulary instruction (Klesius & Seals, 1990; NICHD, 2000; Stahl et al., 1986). This body of research has shown that various ability levels can affect the effects of vocabulary instruction. Tomesen and Aarnoutse (1998), for example, reported similar findings from a combined reciprocal and direct vocabulary instruction program provided to 4th grade students. Students who were low achieving readers showed greater gains from direct instruction in word meanings relative to those with high abilities. Although it is unclear why lower performing students tend to benefit more from vocabulary instruction, this may simply be a function of a floor effect. These students enter with such limited vocabulary and associated comprehension skills that they will benefit from any instruction that builds their vocabulary knowledge and helps them to operationalize and practice detecting word meanings in context.

There are two potential reasons for the mixed outcomes of this study for 3rd and 5th grade students. First, the words selected for this study were included in the 1000 most frequently and widely used words in 3rd through 6th grades (Zeno et al., 1995). This may have resulted in set of words that 3rd graders are less likely to know than are 5th graders. Furthermore, 5th graders may have been more likely to know the multiple meanings for a word than 3rd graders. There is evidence that students develop an ever more complete understanding of words over time (Nation, 2001). Second, the general negative or limited gains in 5th grade students may have occurred because the multiple meaning words that were taught were not content area specific. Students in the 3rd grade may have been more likely to have encountered the words taught in the text they were reading than 5th graders because the emphasis tends to focus more on content area specific reading (Meyerson, Ford, & Jones, 1991).

The fact that educators could implement the multiple vocabulary instruction reliably following a relatively short training session provides evidence of its utility. The teacher self-evaluations of the extent to which they implemented each of the instructional activities or components and followed the two day instructional sequence were high in all cases. Permanent product assessments of lessons were

consistent with the teacher self-evaluations. Teachers also reported that they found the lessons to be structured and easy to follow as well as of the right length.

Limitations and Future Research

As with all studies, this pilot study is not without limitations that should be addressed by future research. First, the study timeframe did not allow us to fully assess the effects of the contextually-based multiple meaning vocabulary instruction over the course of an entire academic year. Human subject consent and budgetary limitations restricted the study timeframe. Future research is needed to determine the effects of contextually-based multiple meaning vocabulary instruction when it is taught for an extended period of time. Second, it is certainly plausible that teacher effects may have influenced the study outcomes. Although the randomization of teachers to conditions should control for this issue, no information was collected on the core vocabulary knowledge and reading comprehension instruction practices provided to students in both experimental conditions. Future research should document the instructional practices used by teachers to clarify the "value added" effects of contextually-based multiple meaning vocabulary instruction. Third, related to this issue, teachers were allowed to select words they believed to be most relevant to their students. It is possible that the words selected by teachers varied and may not have been critical to enhancing the vocabulary knowledge of students. Future research should focus on words that students do not know (established through a pre-test). Fourth, our agreement with participating teachers did not allow us to collect observational data on treatment fidelity. Although the self-evaluations and permanent products suggest that teachers implemented the multiple meaning vocabulary instruction as prescribed, we have no way of knowing if this is the case. Future research should measure treatment fidelity more directly. Fifth, the sample of students was drawn from one school district in one geographic location and may not be representative of the general population of third and fifth graders. It is possible that the findings may not generalize to other students in other geographical regions and diverse populations. Future research should replicate these findings across varied contexts and diverse populations. Sixth, only one vocabulary knowledge and reading comprehension measure was used. In this study, vocabulary knowledge and reading comprehension skills were assessed via a standardized group administered measure (GRMT Vocabulary and Comprehension scales). It may be that students receiving contextually-based multiple meaning vocabulary instruction would have shown greater improve-

ments in their vocabulary knowledge and reading comprehension skills if measures more closely linked to the target words and instructional activities in the program had been used. Future studies could be enhanced by incorporating a range of vocabulary knowledge and reading comprehension measures. Finally, this study appears to be the first to focus on words with multiple meanings. A comprehensive program of research should be undertaken to identify the types of core and/or multiple meaning vocabulary instruction that work with a wide range of diverse students. Unfortunately, it appears that to date there is relatively little research with which to guide education decision makers regarding effective multiple meaning vocabulary instruction that can be used to meet state standards in this area. Research and discussion of vocabulary instruction typically focus on words rather than word meanings.

Implications

With the above limitations in mind, two implications are evident. First, contextually-based multiple meaning vocabulary instruction appears to produce positive outcomes. These outcomes appear to be greatest for students with low initial vocabulary and reading comprehension achievement. Second, contextually-based multiple meaning vocabulary instruction can be implemented reliably by teachers with relatively little training. This is noteworthy given the complexity of enhancing students' awareness that most words have multiple meanings that may fall into different semantic categories depending upon the context in which they are used. The key elements to achieve reliable implementation by teachers include staff development combined with a set of clear and detailed instructional activities that can be followed by both teachers and students.

Acknowledgment

References

Anderson, R.C., & Nagy, W.E. (1991). Word meanings. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (Vol. 2, pp. 690-724).

Armbruster, B.B., Lehr, F., & Osborn, J. (2003). *Put reading first: The research building blocks for teaching children to read: Kindergarten through grade 3.* Jessup, MD: National Institute for Literacy.

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

Biemiller, A. (1999). *Language and reading success*. Cambridge, MA: Brookline Books.

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

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

Chall, J.S., & Dale, E. (1995). *Readability revisited: The New Dale-Chall readability formula*. Cambridge, MA: Brookline Books.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cunningham, A.E., & Stanovich, K.E. (1998). What reading does for the mind. *American Educator, 22*, 8-15.

Dale, E., & O'Rourke, J. (1981). *The living word vocabulary*. Chicago: Worldbook/Childcraft International.

Dale, E., & O'Rourke, J. (1986). *Vocabulary building*. Columbus, OH: Zaner-Bloser.

Ferguson, G.A., & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th ed.). New York: McGraw-Hill.

Fukkink, R.G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68*, 450-469.

Greene, B.E., & 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*, 465-474.

Hart, B., & Risley, T.R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore: Brooks.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Heise, B.L., Papalewis, R., & Tanner, D.E. (1991). Building base vocabulary with computer-assisted instruction. *Teacher Education Quarterly, 18*, 55-63.

Heller, J.H., Sturner, R.A., Funk, S.G., & Feezor, M.D. (1993). The effect of input mode on vocabulary identification performance at low intensity. *Journal of Educational Computing Research, 9*, 509-518.

Kameenui, E., Carnine, D., & Freschi, R. (1982). Effects of text construction and instructional procedures for teaching word meanings on comprehension and recall. *Reading Research Quarterly, 17*, 367-388.

Klesius, J.P., & Searls, E.F. (1990). A meta-analysis of recent research in meaning vocabulary instruction. *Journal of Research and Development in Education, 23*, 226.

Leung, C.B. (1992). Effects of word-related variables on vocabulary growth repeated read-aloud events. 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. 491-498). Chicago, IL: The National Reading Conference.

Levin, J.R., Levin, M.E., Glasman, L.D., & Nordwall, M.B. (1992). Mnemonic vocabulary instruction: Additional effectiveness evidence. *Contemporary Educational Psychology, 17*, 156-174.

Levin, J., McCormick, C., Miller, G., & Berry, J. (1982). Mnemonic versus nonmnemonic vocabulary learning strategies for children. *American Educational Research Journal, 19*, 121-136.

MacGinitie, W.H., MacGinitie, R.K., Maria, K., & Dreyer, L.G. (2000). *Gates-MacGinitie reading tests (4th ed.)*. Itasca, IL: Riverside Publishing.

Malone, R.A., & McLaughlin, T.F. (1997). The effects of reciprocal peer tutoring with a group contingency on quiz performance in vocabulary with seventh- and eight-grade students. *Behavioral Interventions, 12*, 27-40.

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

Meyerson, J.J., Ford, M.S., & Jones, W.P. (1991). Science vocabulary knowledge of third and fifth grade students. *Science Education, 75*, 419-428.

Nagy, W.E., Herman, P.A., & Anderson, R.C. (1985). Learning words from context. *Reading Research Quarterly, 20*, 233-253.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

National Institute for Child Health and Human Development (NICHD). (2000). *Report of the National Reading Panel: Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of the subgroups.* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

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

Scott Foresman (2001). *World of language*. Upper Saddle River, NJ: Pearson.

Scott, J., & Nagy, W. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly, 32*, 184-200.

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

Stahl, S.A. (1999). *Vocabulary development*. Cambridge, MA: Brookline Books.

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

Stahl, S.A., Richek, M.A., & Vandevier, R.J. (1991). Learning meaning vocabulary through listening: A sixth-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) Chicago, IL: The National Reading Conference.

Tomensen, M., & Aarnoutse, C. (1998). Effects of an instructional programme for deriving word meanings. *Educational Studies, 24*, 107-128.

Zeno, S.M., Ivens, S.H., Millard, R.T., & Duvvuri, R. (1995). *The educator's word frequency guide*. New York: Touchstone Applied Science Associates.