THE EFFECTS OF REPEATED READING OF A SINGLE TEXT
AND THE READING OF MULTIPLE RELATED TEXTS ON
VOCABULARY LEARNING IN THE CONTENT AREAS

By LORELL LEVY

A dissertation submitted to
The Graduate School of Education
Rutgers, The State University of New Jersey
in partial fulfillment of the requirements
for the degree
Doctor of Education
Graduate Program in Literacy Education

Approved by

________________________________________
Erica Boling, Chair

________________________________________
Melanie Kuhn, Committee

________________________________________
Lesley M. Morrow, Committee

________________________________________
Douglas Penfield, Committee

New Brunswick, New Jersey

October, 2011
ABSTRACT OF THE DISSERTATION:
The Effects of Repeated Reading of a Single Text and the Reading of Multiple Related Texts on Vocabulary Learning in the Content Areas
By: Lorell Levy
Dissertation Chair: Erica Boling

The present study investigated the effectiveness of two different types of text reading following direct vocabulary instruction. Thirty-six third grade students were placed in one of three conditions: repeated reading of a single text, reading of multiple related texts, and control. All students were exposed to the same direct vocabulary instruction before text reading. Pre-testing, post-testing, and delayed post-testing took place using a standardized reading and vocabulary assessment (Gates MacGinitie Reading Test) and two researcher-developed assessments. The design for the analysis was a three factor mixed design with two fixed variables and one repeated variable. Qualitative data were also collected as a means of contextualizing the quantitative results. Findings indicated vocabulary and comprehension growth in all three groups as a result of the direct instruction. This growth was maintained over a three month period after the end of instruction in three out of four measures. Group differences attributable to follow-up text reading were not found in this study.

INDEX WORDS: Vocabulary, Comprehension, Repeated Reading, Wide Reading, Reading Development
DEDICATION

This dissertation is dedicated to the memory of my grandparents, Janet and Charles Letter. My grandmother always thought that I was the smartest person in the room; my grandfather actually was.
ACKNOWLEDGEMENTS

Everyone knows that it takes a village to raise a child. It actually takes one to complete a dissertation as well. I would like to thank all of the people of my village for helping me through this process.

To begin with I would like to thank my committee members, Erica Boling, Melanie Kuhn, Lesley Morrow, and Doug Penfield, who gave me the guidance and support to realize a dream. Melanie has been there from the very beginning (my first doctoral level class). Erica saw me through to the very end (thank you for being a dedicated committee chair). Lesley provided encouragement from a departmental powerhouse. Doug was my much needed “numbers guy.” You all patiently answered my many questions and were shining examples of scholarly professionalism. Thank you for accompanying me through the triumphs and the pitfalls of doing academic research.

I would like to thank the students, teachers, and administrators at Town Center Elementary School who make getting up and going to work every day a pleasure. A special thank you to my co-workers on the Child Study team who always make me laugh on those occasions when it isn’t such a pleasure.

Thank you to the Book Club girls for scheduling books and discussions around my academic schedule for the past seven years.

Pam De Riso, second scorer and best friend extraordinaire, gets an individual shout out for being my constant cheerleader for over thirty-five years.
A special thank you goes to my parents, Audrey and Lenny Levy, and my sister and brother and their families. You have all made this journey more attainable with your firm expectation that I could do anything I set my mind to.

The best teachers I have ever had in my life have been my children, Miles and Emma. How silly of me to think that I was on the planet to teach you. I have learned about resilience, grace, and joy by just being in your presence. (I’d like to think you might have learned a few things from me as well.)

Lastly, to my husband, Phil Hartman, I cannot thank you enough for the support you have given me to live my dreams. I know that all of my hours of study, research, and writing have helped your golf game, but I also know that all of that free time on the course came with a lot of late dinners and other disruptions to our domestic life. You accepted all of it with a smile and words of encouragement. I won’t forget that when I said, “I’d like to get my doctorate one day” you replied, “What are you waiting for?”
## TABLE OF CONTENTS

ABSTRACT OF THE DISSERTATION......................................................................................iii

ACKNOWLEDGEMENTS......................................................................................................iv

TABLE OF CONTENTS.........................................................................................................vi

LIST OF TABLES...................................................................................................................ix

LIST OF FIGURES...............................................................................................................xi

CHAPTERS

1. Introduction......................................................................................................................1

   Problem Statement.........................................................................................................1

   Purpose Statement.........................................................................................................6

   Research Questions.......................................................................................................7

   Significance of Study....................................................................................................7

2. Review of the Literature...............................................................................................9

   Complexity of Word Learning......................................................................................9

   Word Learning Through Direct Instruction...............................................................11

   Word Learning Through Wide Reading.......................................................................19

      Student Ability........................................................................................................20

      Text Features...........................................................................................................22

      Instructional Response.............................................................................................24

   Vocabulary Learning for Repeated Reading...............................................................26

   Content Area Vocabulary Instruction........................................................................30

   Vocabulary Assessment...............................................................................................33

   Theoretical Framework...............................................................................................39

      Theory of Automatic Information Processing.......................................................39

   Other Views of Automaticity.......................................................................................42
<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schema Theory</td>
<td>44</td>
</tr>
<tr>
<td>How the Two Theories Inform Vocabulary Instruction</td>
<td>46</td>
</tr>
<tr>
<td>3. Methodology</td>
<td>53</td>
</tr>
<tr>
<td>Participants</td>
<td>53</td>
</tr>
<tr>
<td>Materials</td>
<td>58</td>
</tr>
<tr>
<td>Assessment Materials</td>
<td>58</td>
</tr>
<tr>
<td>Intervention Materials</td>
<td>61</td>
</tr>
<tr>
<td>Intervention Design</td>
<td>68</td>
</tr>
<tr>
<td>Data Collection and Analysis</td>
<td>71</td>
</tr>
<tr>
<td>Quantitative Data</td>
<td>72</td>
</tr>
<tr>
<td>Qualitative Data</td>
<td>74</td>
</tr>
<tr>
<td>4. Findings</td>
<td>79</td>
</tr>
<tr>
<td>Quantitative Findings</td>
<td>79</td>
</tr>
<tr>
<td>Gates MacGinitie Vocabulary Subtest</td>
<td>80</td>
</tr>
<tr>
<td>Gates MacGinitie Comprehension Subtest</td>
<td>84</td>
</tr>
<tr>
<td>Productive Vocabulary Test</td>
<td>87</td>
</tr>
<tr>
<td>Maze Comprehension Test</td>
<td>91</td>
</tr>
<tr>
<td>Summary of Quantitative Data</td>
<td>94</td>
</tr>
<tr>
<td>Qualitative Findings</td>
<td>96</td>
</tr>
<tr>
<td>Consistency of Instructional Time</td>
<td>96</td>
</tr>
<tr>
<td>Qualitative Themes</td>
<td>98</td>
</tr>
<tr>
<td>Summary of Qualitative Data</td>
<td>104</td>
</tr>
<tr>
<td>5. Discussion</td>
<td>107</td>
</tr>
<tr>
<td>Significant Findings</td>
<td>108</td>
</tr>
<tr>
<td>High and Low Ability Readers</td>
<td>109</td>
</tr>
<tr>
<td>Gains from Post-test to Delayed Post-test</td>
<td>110</td>
</tr>
<tr>
<td>Type of Text Practice</td>
<td>113</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1. Participants in Study Groups.................................................................57
2. Target Vocabulary, Definitions, Living Word Vocabulary Data........64
3. Vocabulary Count for Books in the Study...............................................67
4. Design of the Study..............................................................................73
5. Quantitative Codes and Examples from the Study Data.................76
6. Number of Minutes Recorded by Lesson Components.......................80
7. Frequency of Coded Themes for Qualitative Data...............................87
8. Descriptive Statistics by Instructional Group and High/Low Factor for the Gates MacGinitie Vocabulary Subtest.................................90
9. Descriptive Statistics for Gates MacGinitie Vocabulary Subtest by Instructional Group across TestTime...................................................91
10. Descriptive Statistics for High and Low Groups on Gates MacGinitie Vocabulary Subtest across TestTime.............................................92
11. Descriptive Statistics by Instructional Group and High/Low Factor for Gates MacGinitie Comprehension Subtest.................................................94
12. Descriptive Statistics for Gates MacGinitie Comprehension Subtest by Instructional Group across TestTime...............................................96
13. Descriptive Statistics for High and Low Groups on Gates MacGinitie Comprehension Subtest across TestTime...........................................97
14. Descriptive Statistics by Instructional Group and High/Low Factor for the Productive Vocabulary Test

15. Descriptive Statistics for Productive Vocabulary Test by Instructional Group across Test Time

16. Descriptive Statistics for High and Low Groups on the Productive Vocabulary Test across Test Time

17. Descriptive Statistics by Instructional Group and High/Low Factor for the Maze Comprehension Test

18. Descriptive Statistics for High and Low Groups on the Maze Comprehension Test across Test Time
LIST OF FIGURES

1. Interaction Between Gates MacGinitie Vocabulary Test and High/Low Factor.................................................................................................................................................................93

2. Interaction Between Productive Vocabulary Test and Instructional Groups........................................................................................................................................................................................................100

3. Overall Means for All Instructional Groups on the Maze Comprehension Test...........................................................................................................................................................................................................103
CHAPTER 1
INTRODUCTION

Problem Statement

Long before the National Reading Panel (2000) identified vocabulary as one of the five essential components of reading instruction, vocabulary was recognized and studied as an important contributor to reading achievement (Baumann, Kame`enui, & Ash, 2003). The research interest dates back to early in the 20th century (Anderson & Nagy, 1992; Blachowicz, Fisher, Ogle, and Watts-Taffe, 2006;), but has recently been renewed by the National Reading Panel’s report (NICHD, 2000) and several research reviews, chapters, and books that have followed (Baumann, Kame enui, & Ash, 2003; Baumann & Kame ennui, 2004; Blachowicz, Fisher, Ogle, & Watts-Taffe, 2006; Hiebert & Kamil, 2005; Lehr, Osborn, and Heibert, 2004). Vocabulary’s importance is related to its strong connection to comprehension (Anderson & Freebody, 1981; NICHD, 2000; Ouellette, 2006). If the goal of reading is for students to understand what they read so they can make use of that knowledge, then any contributor to comprehension ability should be explored. The correlation between vocabulary knowledge and reading comprehension is well-documented in the literature (Mezynski, 1983; Stahl & Fairbanks, 1986). However, the relationship is rather difficult to unravel. There are several hypotheses that attempt to explain the correlation (Anderson & Freebody, 1981; Mezynski, 1983; Nagy, 2005; Sternberg & Powell, 1983).

The instrumentalist hypothesis describes the connection as a direct one (Nagy, 2005). Therefore, just knowing more words will
result in greater text comprehension. The instrumentalist hypothesis does not comment on where vocabulary knowledge originates, but only that it directly impacts comprehension ability.

The verbal aptitude hypothesis argues that the vocabulary-comprehension connection is due to an underlying third factor that can be described as general “verbal aptitude” (Sternberg & Powell, 1983). The verbal aptitude in this hypothesis has been conceptualized by different researchers as quick thinking ability, skill in inferring, or metalinguistic capacity (Anderson & Freebody, 1981; Mezynski, 1983; Sternberg & Powell, 1983; Nagy, 2005). This aptitude makes certain students better word learners, as well as better text comprehenders.

The access hypothesis explains that vocabulary knowledge is useful to comprehension when words can be accessed quickly and easily (Mezynski, 1983). This hypothesis argues the importance of depth of vocabulary knowledge, as well as breadth. Automaticity of word knowledge is very important in this explanation.

The fourth hypothesis, the knowledge hypothesis, emphasizes the role of a reader’s background knowledge in comprehension (Anderson & Freebody, 1981). In this hypothesis, knowing a word well implies that one knows other words and ideas related to the original word. This larger body of knowledge becomes crucial for understanding a text. Most authors do not see these hypotheses as mutually exclusive. Rather, all of the hypotheses may contribute to our understanding of how vocabulary knowledge and text comprehension are related.

Another reason for the long-standing research interest in vocabulary is born out of the theory that reading results in an
increase in knowledge and a subsequent increase in access to more knowledge through the written word (Cunningham & Stanovich, 1998; Stanovich & Cunningham, 1993). This relationship leads to what Stanovich has termed the “Matthew Effect” (Stanovich, 1986) where proficient readers read more and therefore increase their reading competency. These readers increase their vocabulary knowledge, which in turn helps them to comprehend other texts in the future. Vocabulary is an important component in this description of spiraling competency.

The extensive vocabulary research base has resulted in us knowing a good deal about vocabulary learning. Word knowledge is complex and multifaceted. Words can be “known” either receptively or expressively, in oral language or written language (Baumann, Kame’enui, & Ash, 2003). There are several aspects of the complexity of word knowledge recognized by researchers. Word learning is incremental in nature. Words often have more than one meaning and are multidimensional. Words are also interrelated in a network with other words, and understanding a word’s meaning can depend on what kind of word is being learned (Nagy & Scott, 2000). This complexity can make the word learning process a challenge.

If words are learned incrementally (Durso & Shore, 1991; Stahl, 2003), then word learning is not an all or nothing undertaking. Readers, in multiple exposures to words, increase their understanding by developing a more complete knowledge of a word’s decontextualized meaning. Research has shown that a combination of definitional and contextual information is most effective in teaching word meanings.
That means that vocabulary instruction should not consist of simply sending students to look up definitions and write sentences for words. Deep processing of meanings makes vocabulary learning more effective (Beck, McKeown, & Kucan, 2008; Stahl, 1999). Active, deep processing takes place when new information is combined in some way with old information. Students can be asked to discuss the meaning of the same word in different sentences, create scenarios using the word (Stahl, 1999), or answer silly questions that contrast two or more vocabulary words (Beck, Perfetti, & McKeown, 1982). These methods help students become actively engaged in the learning, which helps them retain more of what they have learned.

Despite the large vocabulary research base, there has been little change in classroom practice over the years (Blachowicz, Fisher, Ogle & Watts-Taffe, 2006). This may be due to the perceived enormity of the task, or confusion about what words to teach, how to structure vocabulary teaching, and how to promote transfer of learning (Berne & Blachowicz, 2008). All of these issues make it difficult for teachers to decide on the best way to integrate vocabulary teaching into their classrooms.

Early vocabulary research looked at either direct instruction of words and strategies (Beck, Perfetti & McKeown, 1982; McKeown, Beck, Omanson & Perfetti, 1983) or wide reading (Nagy, Anderson & Herman, 1987; Nagy, Herman & Anderson, 1985) as the most effective means of increasing vocabulary knowledge. While each camp still has its supporters, the consensus in the field is that an effective program
will contain both strategies (Kame’enui & Baumann, 2004; Nagy, 1988; Stahl, 1999). Therefore, a powerful vocabulary methodology would combine direct instruction of vocabulary and practice with target words in context. The question is: What form should that text practice take?

One of the agreed upon principles of vocabulary learning is that students need multiple exposures to words for the incremental nature of vocabulary learning to have an effect (Leung, 1992; NICHHD, 2000; Stahl, 2003). Therefore, one could assume that repeated reading of a single text would be beneficial as a means of text practice after vocabulary instruction. Repeated reading has been found effective in increasing the reading fluency of students; this increase in reading fluency is correlated with increases in comprehension (Dowhower, 1987; Herman, 1985; Pikulski & Chard, 2005). Other research has called into question whether the determining factor in fluency and comprehension increases is the repeated reading or just increased exposure to connected text that could take place within a wide reading format (Homan, Klesius, & Hite, 1993; Kuhn & Stahl, 2003). The context variability hypothesis (Bolger, Balass, Landen, & Perfetti, 2008) would help explain why contexts that vary promote word learning better than contexts that do not vary. However, these two methodologies have not been explored and compared as a follow-up strategy to vocabulary instruction. Recommendations to teachers about the most efficient kind of text practice to offer their students can only come as the result of research on vocabulary gains that take place in response to each of these conditions. This study is an attempt to clarify how
teachers can structure text practice as a means for increasing vocabulary learning following direct instruction of words.

Students are exposed to new vocabulary throughout the school day. The content areas provide an especially rich source of words for instruction in the elementary classroom. Content area vocabulary, and more specifically science vocabulary, is interrelated, providing strong connections among words and supporting concept development and retention. However, these words can be challenging to elementary students who may lack the necessary schema (relational structures) in which to place these new terms (McVee, Dunsmore, & Gavelek, 2005; Meyerson, Ford, Jones, & Ward, 1991). In addition, students need help in mastering science words since learning a word for a new concept is harder than learning a new word for an already known concept (Graves, 2009; Spencer & Guillaume, 2006).

Content area textbooks become an ever more central component of instruction as students move up through the grades (Stahl & Kapinus, 1991). Therefore, students who have well-developed science vocabularies will be able to more fully participate in and successfully comprehend science instruction. This knowledge can be a gateway to concepts that will be used throughout their schooling.

**Purpose Statement**

The purpose of this study is to examine how the type of text practice (repeated reading of one text versus reading of several related texts) influences content area vocabulary learning.
Research Questions

- What is the influence of two different types of text practice on the acquisition of content area vocabulary? Following vocabulary instruction, what are the effects of repeated reading with one text as compared to reading a variety of texts with the same target vocabulary?

- What is the difference in the acquisition of vocabulary knowledge for more proficient readers and less proficient readers?

- What is the difference in vocabulary growth between the repeated one-text group and the multiple text group after a four week intervention? Are the results consistent three months later?

Significance of Study

The significance of the study is its contribution to the growing body of literature that makes vocabulary teaching more accessible to classroom teachers (Beck, McKeown, & Kucan, 2002; Beck, McKeown, & Kucan, 2008; Graves, 2009; Stahl & Kapinus, 2001). There is very little written about how teachers can structure text reading as it pertains to vocabulary learning beyond encouraging students to engage in wide reading (Nagy, Anderson, & Herman, 1987). This study is an attempt to clarify if there are differences in the efficacy of repeated reading of one text or reading multiple related texts after direct content area vocabulary instruction.
This chapter gave a brief overview of the research that guided the design of this study. In Chapter 2, I elaborate on the theoretical and research foundation upon which the study was built. Chapter 3 describes the methodology used. In Chapter 4, I present the qualitative and quantitative findings of the study, and in Chapter 5 I discuss and present interpretations about the results and suggest implications for future research.
CHAPTER 2
REVIEW OF THE LITERATURE

This chapter will review the literature on vocabulary learning. To do so, a discussion regarding the complexity of word learning will open the chapter. After that the literature base surrounding word learning through direction instruction and through wide and repeated reading will be presented. Research on vocabulary instruction in the content areas and vocabulary assessment will follow. The chapter will close with a theoretical framework to explain the mechanisms underlying vocabulary learning.

The Complexity of Word Learning

The amount of vocabulary learning that takes place during a student’s school years is quite large and impressive. Estimates vary, but 3,000 words per year between grades 3-12 is an often quoted statistic and one that is a median figure within the range of estimates (Anderson & Nagy, 1992; Lehr, Osborn, & Hiebert, 2004; Nagy & Anderson, 1984; White, Slater, & Graves, 1990). The fact that all of this vocabulary growth takes place is made even more remarkable when one realizes how complex the word learning task actually is.

Words can be “known” either receptively or expressively, in oral language or written language (Baumann, Kame‘enui, & Ash, 2003). Expressive knowledge of a word requires a higher level of understanding than receptive knowledge. This can be seen in the fact that a person can understand a word in speech or text that he might
not be able to use correctly in his own speaking or writing. Written language, which is decontextualized, is more complex than spoken language, which retains the context that the speaker and the listener are in, as well as non-verbal cues, voice intonation, and facial expressions (Kamil & Hiebert, 2005). Learning words from written input is therefore a more complex task than learning words from oral input. However, the reason for the difficulty lies not only in the decontextualization of written text; it is also a consequence of the greater diversity and complexity of the language used in texts as compared to oral language. Speech is “lexically impoverished” compared to written language (Cunningham & Stanovich, 1998; Hayes & Ahrens, 1988). Therefore, the vocabulary demands of written text will be higher than the demands encountered in spoken interactions.

In addition to the expressive/receptive and oral/written dichotomies, there are five aspects of the complexity of word knowledge that are recognized by researchers (Nagy & Scott, 2000). Incrementality describes how word learning takes place in many steps. One encounter with a word will usually not be enough for learning to occur. This incremental view of word learning explains how a great deal of vocabulary knowledge can be gained incidentally from context over time even though individual encounters with words may not be that informative. Polysemy describes the fact that words often have more than one meaning and that learners need to be sensitive to the nuanced ways in which words are used. Multidimensionality explains the different dimensions of word knowledge a student might possess about a particular word (spoken form, written form, grammatical behavior,
conceptual meaning, associations with other words). Interrelatedness describes the fact that single words are part of a semantic network and learning a new word has an effect on the student’s understanding of the network. Conversely, more knowledge about the network impacts the student’s understanding of individual words. Heterogeneity explains that what it means to know a word depends on what kind of word is being learned.

Students expand their reading vocabularies either through direct instruction or through wide reading. There has been substantiation in the research for both methods of learning. The next two sections will review the research support for each method.

**Word Learning Through Direct Instruction**

The research base surrounding direct instruction of vocabulary has included an evolving discussion of principles of vocabulary instruction, as well as specific recommendations about classroom activities to promote reading vocabulary knowledge (McKeown & Beck, 2004). The argument in support of direct classroom instruction was recently summarized by Coyne, McCoach and Kapp (2007) as being generally effective and more beneficial for at-risk learners than inferring meaning from incidental exposure. For the past twenty-five years, research has been used to develop principles to describe effective vocabulary development.

Stahl (1986, 1999) has described three principles of effective vocabulary instruction that he culled from a review of the research. The first principle is that methods that provide a combination of
definitional and contextual information are more effective than methods providing only definitional information. In their meta-analysis of fifty-two studies, Stahl and Fairbanks (1986) found this to be so. By taking this two-pronged approach, investigators in the analyzed studies were able to significantly improve reading comprehension. The strategy of combining both definitional and contextual information has been echoed in more recent reviews of the literature (Baumann, Kame’enui & Ash, 2003; Graves, 2008).

Stahl’s second principle (Stahl, 1986; Stahl, 1999) states that children should be actively engaged in processing new word meanings. Active processing takes place when students are required to make associations between new information and previously known information. Semantic relatedness tasks, where words are categorized and relationships between words are explored, are powerful tools in helping students make connections between the known and unknown (Apthorp, 2006; Rupley & Nichols, 2005; Taylor, Mraz, Nichols, Rickelman & Wood, 2009). Deeper processing occurs in conjunction with the classroom discussion that transpires during vocabulary instruction. Stahl and Vancil (1986), in a study of semantic mapping and discussion as a vocabulary strategy, found that the more powerful predictor of vocabulary learning was the discussion, not the actual map. The authors hypothesized that discussion forced the students to process the word meanings more deeply and connect the new information to information they already possessed. The number of contributions an individual student made was not correlated to the test results, indicating that vocabulary learning was not related to oral
participation in the class discussion. The discussion was a benefit to discussants and non-discussants alike. Blachowicz and Lee (1991) also encouraged teachers to involve students in meaningful discussion about newly learned words as a means of promoting richer semantic connections.

Stahl and Fairbanks (1986) developed a scale of processing used for evaluating vocabulary instruction that takes into consideration three levels of processing of meanings: Association, Comprehension, and Generative. Association processing takes place when a student can link a new word to a specific definition or context; it involves rote memorization. A deeper level of processing is comprehension processing where a student can demonstrate the understanding of a word in a sentence or use definitional information about a word to find an antonym or synonym, or classify/sort the word in some way. The deepest level of processing is generative processing where a student can produce a novel response to a word, such as a new sentence or a definition in his own words (Baumann, Kame’enui, & Ash, 2003). This scale can be used to examine vocabulary instruction and determine the depth of processing that it requires.

Stahl’s third principle (Stahl, 1986; Stahl, 1999) states that students need multiple exposures to meaningful information about words. McKeown, Beck, Omanson, and Pople (1985) examined fourth grade students receiving one of three types of instruction at two different levels of exposure (four or twelve exposures to a word). More frequent exposures to the target words yielded better results on all measures. Jenkins, Matlock, and Slocum (1989) examined fifth grade
students who were presented with two different instructional methodologies implemented with low, medium, or high amounts of practice. They found that more practice on individual word meanings led to higher levels of mastery on the words that were taught. This principle of multiple exposures to words yielding higher word learning has been tested on a variety of student populations with consistent results (Apthorp, 2006; Baumann, Ware, & Edwards, 2007; Beck, McKeown, & Kucan, 2002; Lovelace & Stewart, 2009). One theory about why multiple encounters are important is that with increased exposure to a word a student increases his speed of access to that word (Beck, Perfetti, & McKeown, 1982; Stahl, 1991); more practice leads to faster access to the word’s meaning, making the meaning more available to the learner. To these authors, depth of knowledge is an important factor in word learning.

While Stahl’s three principles have been widely accepted, other principles of effective instruction have also emerged from the literature. Nichols & Rupley (2004) and Taylor et al. (2009) have discussed how giving students the opportunity to compare and contrast new words with ones already known in a visual display can be beneficial to vocabulary learning. Teaching related groups of words (Beck, McKeown, & Kucan, 2002; Mezynski, 1983), whether content area vocabulary or semantically related vocabulary from a narrative, can help students make connections and put their learning in a context. Although selecting words from ones that students will encounter in their classroom reading (Rupley, Logan, & Nichols, 1999; Rupley & Nichols, 2005) makes a great deal of intuitive sense, this has not
always occurred in practice (Blachowicz, Fisher, Ogle, & Watts-Taffe, 2006). Finally, researchers acknowledge that different words need different instruction (Graves, 2006; Stahl, 2005). Words that represent new concepts for students need a different type of instruction than words that are more sophisticated synonyms for known concepts (Scott, Jamieson-Noel, & Asselin, 2003).

Graves (2008) reiterates that rich and powerful vocabulary instruction is instruction that activates prior knowledge, compares and contrasts word meanings, helps students make inferences, and includes frequent encounters with words. However, no one vocabulary instructional method has proved better than the others (Fisher & Blachowicz, 2005), so researchers and teachers have to make decisions regarding which methods to use by taking into account the principles just discussed. The following instructional methodologies would be supported by the principles of effective vocabulary instruction:

*Semantic Mapping* (Heimlich & Pittelman, 1986) has been used in classrooms across content areas with positive results. The teacher places a word on the board representing a central concept and asks students to work in groups to brainstorm a list of words related to the concept. The teacher then brings the whole class back together and lists all of the words generated. The teacher puts these words into categories as she is recording them and then asks the students to label the categories and perhaps suggest other categories. The discussion that takes place around the words and their placement in the categories has proven to be the most powerful feature of this technique (Stahl & Vancil, 1986).
Semantic Feature Analysis (Pittelman, Heimlich, Berglund, & French, 1991) is similar to Semantic Mapping in that it provides a visual display in which the students analyze and synthesize word meanings. Semantic Feature Analysis consists of a grid containing a set of related vocabulary down one side of the grid (e.g., different dinosaurs) and a list of possible features about the words across the top (e.g., carnivores, herbivores). In filling out the grid students make fine discriminations between related word meanings.

Possible Sentences (Stahl & Kapinus, 1991) is an activity in which the teacher chooses six to eight words that might be difficult for the students from a content area text, for example. An additional four to six words are chosen that are considered known by the students. The ten to twelve words are put on the board and a short definition of each word is given. Students are then asked to think of sentences that could include two of the words and may be found in the chapter. All suggestions are recorded without discussion about whether they are correct or not. Once finished, the students read the chapter or passage. Following the reading, the class returns to the sentences and discusses which ones would be true or not using the reading as a guide. The class then rewrites the incorrect sentences using this new knowledge.

Concept of Definition Maps (Schwartz & Raphael, 1985) are visual displays that help children understand the elements usually included in a definition: a category, defining characteristics, and examples. According to Fisher and Blachowicz (2005), the teacher can present the word and the students can use the text to complete the map by filling
Venn Diagrams (Nagy, 1988) help students compare and contrast two words. In the process they discover which characteristics are exclusive to each word and which characteristics are shared (Graves, 2009). Two intersecting circles are displayed. Each circle is labeled with one of the words. Shared characteristics are recorded in the intersection of the circles. Characteristics of each individual word are recorded in the parts of the circle that do not intersect. Venn diagrams can be constructed as a whole group lesson and later by small groups or pairs of students.

Linear Arrays (Nagy, 1988) are best constructed with groups of words that differ by degree (enraged, angry, mad, annoyed, furious). The differences among the words can be displayed by arranging them in a line in a justifiable order. Students can talk about why they put the words in the particular order as a means of enriching the knowledge of the related group of words.

Vocabulary Visits (Blachowicz & Obrochta, 2007) is a method of teaching content area vocabulary to elementary students based on the concept of a field trip. Vocabulary Visits prepares the students for the journey using a poster as a focal point. The teacher asks students to generate a list of words for things they see on the poster. Those words are added to the poster using post-its. Then a book about the topic is read and students give the teacher a “thumbs up” when they hear any of the words on the post-its. Other books are subsequently read and the post-it words may be re-organized as
meanings become clearer. Word games, sorting activities, and writing are also part of the Vocabulary Visits methodology. Evaluation takes place by having the students again generate a list of words for the topic and comparing the final list with the initial list.

Robust Instruction is a methodology developed and investigated by Beck and McKeown and their colleagues (Beck & McKeown, 2007; Beck, McKeown, & Kucan, 2002; Beck, Perfetti, & McKeown, 1982). It is used to give deep understanding to what Beck and McKeown call Tier Two words (words that are of high frequency for mature language users and are found across a variety of domains). Robust instruction begins with student-friendly definitions of several words. Each word is presented in several contexts and experienced multiple times. Students are engaged in activities in which they deal with various facets of a word’s meaning and its relationship to other words. Students are encouraged to generate their own sentences about the words and use and find the words in use outside of the classroom (Graves, 2009).

The preceding list of activities is not exhaustive but provides a sample of the kinds of direct instruction that follow the principles of effective vocabulary instruction. In the next section I will describe the research on wide reading and deriving meaning from context.
Word Learning Through Wide Reading

Taking into account Nagy and Scott’s (2000) contention that there are far too many words to learn to be taught through direct instruction, several researchers have explored the effectiveness of teaching students to use context to infer word meanings. Anderson and Nagy (1992) report that students learn about 5% of unknown words in text under natural reading conditions. Swanborn and de Glopper (1999), in a meta-analysis of 20 experiments, estimate the likelihood of incidental word learning during normal reading at closer to 15%. While either number appears small, the impact of these statistics is observed when the incremental nature of vocabulary development is considered. Students, through the cumulative effect of daily reading, will increase their vocabulary knowledge. Anderson and Nagy (1992) estimate that the median fifth grader spends approximately twenty-five minutes a day reading, which translates into about 2 million words of text per year. If, in a conservative estimate, 2 percent of the words encountered are unfamiliar, twenty thousand new words are encountered each year. If 5% of these are learned, that would account for one thousand new words learned each year. Reading volume, therefore, will have an impact on vocabulary growth if it is agreed that students can gain word meaning from context.

While there is agreement in the field that learning from context has a role in the vocabulary learning process (Cunningham, 2005; Graves, 2006; Scott, 2005), the extent of the role and the power of contextual learning have been some of the issues challenging researchers. Over the past 25 years, investigators have examined the
impact of several factors on the rate of vocabulary learning from context, including student ability, text factors, and instructional response.

**Student Ability**

Student reading ability has been seen as a factor impacting vocabulary learning from context. Sternberg and Powell’s (1983) theory of verbal comprehension is one way to examine the differential between vocabulary learning from context for students of high and low reading ability. They define verbal comprehension as a person’s ability to understand linguistic material, such as newspapers, magazines, textbooks, and lectures. Sternberg and Powell hypothesize that the close link between tests of verbal intelligence and vocabulary measures is due to the fact that they are both measuring a person’s ability to acquire new verbal knowledge. This underlying ability is what makes more skilled comprehenders better at learning vocabulary through context. In Sternberg and Powell’s view, the more skilled comprehenders are able to acquire more information through context, distinguish that information from unimportant information, and store and retrieve it as needed. The key for them is the ability to acquire new information, which leads to better reading comprehension results but also leads to more vocabulary learning from context. Nagy and Scott (2000) also include metalinguistic ability as a prerequisite for effective learning from context. Cain, Oakhill and colleagues (Cain, Lemmon, & Oakhill, 2004; Cain, Oakhill, & Ebro, 2003) contend that working memory is a factor that explains the
difference in how skilled and less skilled readers derive meaning from unknown words in text. Oral vocabulary (Hart & Risley, 1995) and familiarity with a particular topic (Nagy & Scott, 2000) have also been proposed to explain the differences in performance.

Several studies and meta-analyses have shown a difference in the ability to acquire vocabulary from context between high and low ability readers (Blackowicz & Zabroske, 1990; Herman, Anderson, Pearson, & Nagy, 1987; Kuhn & Stahl, 1998; McKeown, 1985; Stanley & Ginther, 1991; Swanborn & de Glopper, 2002; van Daalen-Kapteijns, Elshout-Mohr & de Glopper, 2001). McKeown (1985) and van Daalen-Kapteijns et al. (2001) suggest that low ability readers may have a more difficult time adjusting their hypotheses about words in the face of conflicting information. This would lend credence to Sternberg and Powells’ theory of lower skilled comprehenders not being able to distinguish between important and unimportant information when acquiring new verbal knowledge. The studies on deriving meaning from context found that students of all abilities benefitted from the interventions. However, the gaps between the two groups remained after all of the interventions. Swanborn and de Glopper (2002) argue that the fact that good readers learn more vocabulary incidentally than poor readers implies that the gap between the two ability groups will only get wider. They state that, therefore, wide reading without any further aid will not automatically lead to a significant gain in vocabulary size for low ability readers. Stanovich and Cunningham (1993) and Stanovich (1986) have shown that students differ in the independent reading that they do. Less skilled readers spend less
time reading and therefore have fewer opportunities to increase their skills in both reading and vocabulary acquisition from context.

**Text Features**

Text features include the many ways that text can help or hinder a student in acquiring meaning from context. Research on text features has yielded a number of influential factors. Nagy, Anderson, & Herman (1987) found that although students were able to derive meanings from context for narrative and expository text, morphological transparency was not a factor in whether or not a word’s meaning could be derived. Words including suffixes and words whose meanings could be broken down into recognizable parts were no more easily derived than other less transparent words. Conceptual difficulty was found to affect learning from context (Nagy, Anderson & Herman, 1987; Schwanenflugel, Stahl, & McFalls, 1997). Words that were more easily learned were those that typically had easier access to imagery and had greater accessibility to information stored in prior knowledge.

Beck, McKeown, and McCaslin (1983) developed a continuum by which contexts could be classified as misdirective, neutral, generally directive, or directive. The authors used adult subjects to test their hypothesis by having them read passages that the authors had classified using the continuum and fill in the blanks where the missing target words should be. The results of their study confirmed their hypothesis. However, neither Nagy, Anderson and Herman (1987) nor Schwanenflugel, Stahl, and McFalls (1997) found that contextual support had an impact on learning. The difference in results might be
due to the difference in assessment in the studies. Beck et al. (1983) used a modified cloze procedure and subjects were told that they should use the context to help them fill in the blanks. By contrast, Nagy, et al. (1987) and Schwanenflugel et al. (1997) gave students a more natural task in that they read intact texts and were not told anything about deriving meanings from context. At a later date (one week, and three days, respectively) they were given a multiple choice test to assess their knowledge of target words. This task was more difficult due to the time lapse and the lack of prompting.

Swanborn and de Glopper (1999), in the meta-analysis of 20 studies, found that the density of unknown words in a passage had an effect on word learning. They found that a low density of unknown words in a text produces a higher chance of word learning than a high density of unknown words. A text that holds a lot of unknown words will not give the reader much contextual aid and will make deriving meanings of unknown words much harder.

The proximity of target words to a relevant context clue was examined by Carnine, Kame’enui, and Coyle (1984); Cain, Oakhill, and Elbro (2003); and Cain, Lemmon, and Oakhill (2004). All three studies found that students were better able to determine the meanings of unfamiliar words when context clues were closer to the target words. Working memory and process demands were used as explanations for the difficulty caused by the separation of the word from its contextual information. Carnine, Kame’enui, and Coyle (1984) also looked at types of context clues and found that contextual information was
easier to use when it was in synonym form rather than in inference form.

**Instructional Response**

Several studies have been published over the past quarter century that looked at the results of interventions to teach students how to use context to increase vocabulary. Carnine, Kame‘enui, and Coyle (1984) found that rule-plus-systematic practice and systematic-practice-only conditions produced higher scores than a no-intervention condition. The absence of differences between the two intervention treatments suggests that practice in deriving meaning may be the determining factor in student success. Kuhn and Stahl (1998) in their review of 14 studies of word learning from context also found that practice is an important factor. In nearly all of the studies, treatments were effective compared to no-treatment controls. However, in the four studies where there was also a practice-only treatment, there were no significant differences between the strategy treatment and the practice-only groups. The authors recommended that future studies should include a practice only group to test the hypothesis that practice is the determining factor in increasing word learning from context.

In two studies (Blachowicz & Zabroske, 1990; Buikema & Graves, 1993), teachers used natural classroom environments to test instructional programs designed to teach middle school students to derive meanings from context. Both interventions showed skill growth in the natural setting. Teaching struggling readers has been the
focus of other studies (Goerss, Beck, & McKeown, 1999; van Daalen-Kapteijns, Elshout-Mohr, & de Glopper, 2001), where the goal was to have the students perform a sequence of activities (strategy) in order to give them a structure in which to process unknown words in text. Both studies reported success, but cautioned that a specific strategy was needed due to natural texts not providing strong enough contextual cues. They also acknowledged that the strategy did not help in closing the gap between students with high and low ability.

Fukkink and de Glopper (1998) in their meta-analysis of 21 studies reported that clue instruction was more effective than strategy or cloze instruction and did not agree that practice only would result in comparable results. Askov & Kamm (2001) also concluded that specific context clues can be taught, even to students as young as third graders. Finally, Baumann and colleagues (Baumann, Edwards, Boland, Olejnik, & Kame’enui, 2003; Baumann, Edwards, Font, Tereshinski, Kame’enui, & Olejnik, 2002) found that fifth grade students were able to learn to derive meaning from context though a combination of morphemic and contextual instruction. Support for instruction in morphemic analysis was somewhat more robust than support for instruction in contextual analysis, but the authors concluded that both would be beneficial.

The preceding studies differ in that some measured the ability to derive meaning from context, in which the subjects were prompted to attend to the text as a way of figuring out the meanings of unknown words. Other studies measured incidental word learning from text where subjects were not told what the task was beforehand and were left to
read in a more natural context. While these two tasks are obviously different and will have different results on assessments, both tasks require assessments that are sensitive to the incremental changes of vocabulary acquisition. Study conclusions showed that deriving meaning from context was more influenced by interventions. Incidental word learning continues to be an area requiring further investigation.

In the next section I will describe the research on repeated reading as it relates to vocabulary learning.

**Vocabulary Learning from Repeated Reading**

Research explaining the effectiveness of repeated reading on vocabulary growth has mainly taken place in the context of read-aloud interventions. However, this research can inform our understanding of how repeated reading done by students themselves can affect their vocabulary development.

Vocabulary is learned incrementally through repeated exposures to a word (Nagy & Scott, 2000). Carey (1978) first described how this phenomenon takes place in oral vocabulary development through “fast mapping”, a way of explaining the rapid vocabulary growth of young children. In fast mapping, a child can develop a quick, but incomplete meaning of a new word when first encountering it in text (Justice, Meier, & Walpole, 2005). This partial understanding of the sound, meaning, and use of the word, a “fast map”, is further developed with each successive encounter with the word until a more complete understanding is formed (Leung, 2008; Miller & Gildea, 1987).
In this “slow mapping” process, representations are refined with multiple exposures to the word (Curtis, 1987).

Research on read-alouds has looked at how repeated readings help children acquire and refine vocabulary knowledge. In one of the earliest studies, Elley (1989) demonstrated that repeating a reading three times helped children increase their vocabulary knowledge of unknown words by 15.4%. Seven-year-olds learned approximately three out of the twenty target words read to them in the stories. Leung (1992; 2008) demonstrated how young children used target words with more precision across three retellings of repeated storybook readings. Robbins and Ehri (1994) examined the effect of the number of exposures of target words on vocabulary learning. Kindergartners were read the same storybook twice, either two or four days apart. Some target words appeared twice in the story so some words were heard twice and others were heard four times. The authors determined that four exposures to a word were necessary, but not sufficient for higher rates of word learning.

Several researchers found that repeated exposure to text led to greater learning of target vocabulary (Elley, 1989; Leung, 1992; Robbins & Ehri, 1994). However, those researchers who added an explanation as part of their design found that even greater vocabulary growth could be attributed to teacher explanations. Senechal (1997) placed 30 three- and four-year-olds in one of three groups: a single reading of a storybook, a repeated reading (three times) of a storybook, or a repeated reading with a questioning condition. Children’s performance in the repeated reading conditions was better
than the single reading condition. In addition, children learned more vocabulary from answering questions during the three readings than from just listening to the three readings. Overall performance was better on receptive measures than on expressive measures. Penno, Wilkinson and Moore (2002) found that the 47 five to eight-year-olds who took part in their study benefitted from both frequency of exposure and teacher explanation of target vocabulary. Biemiller and Boote (2006) found that an average gain of 12% of word meanings was obtained through repeated reading. Including word explanations added a 10% gain for a total of 22%. In a follow-up study the authors found that in a delayed post-test word meanings were not lost and some new words were learned, possibly due to increased word consciousness. The effectiveness of teacher explanations was furthered explored by Coyne, McCoach, and Kapp (2007) who examined extended instruction. Extended instruction was characterized by explicit instruction using both contextual and definitional information, repeated exposures to words in varied contexts, and deep processing of meanings. They compared extended instruction to both incidental exposure and embedded instruction (which consisted of providing simple definitions within the context of the story). In a repeated reading methodology, extended instruction was more effective than either of the other models. The above findings may be due to the fact that listening to stories may be effective in developing receptive vocabulary, but more active participation is needed for growth in expressive vocabulary (Cunningham, 2005).
Some researchers have used the research on read-alouds along with the research on vocabulary learning to develop programs to make read-alouds a more powerful teaching tool for increasing vocabulary. Beck & McKeown (2001; 2007) hypothesized that the most important factor in vocabulary growth from read-alouds was getting students to talk about the texts. “Text Talk” was developed to enhance young children’s ability to construct meaning from decontextualized language. The interspersed open-ended questions in “Text Talk” encourage children to describe and explain text ideas, rather than recalling and retrieving words from the text. Studies on students in kindergarten and first grade showed increased vocabulary growth for those children who heard read-alouds using the “Text Talk” procedure. Blachowicz & Obrochta (2007) developed “Vocabulary Visits” to increase the effectiveness of read-alouds of informational text by reading thematic text sets. These sets allowed for repetition in multiple contexts while also adding opportunities for students to actively process the vocabulary in classroom conversations and activities. The authors found that the “Vocabulary Visits” group had significantly greater gains in words learned than the standard read-aloud group.

The preceding research demonstrates the power of read-alouds in vocabulary learning. It can also inform our understanding of increasing the vocabulary learning of students old enough and skilled enough to read independently. Repeated reading of text gives a student the opportunity for incremental learning to take place. A “fast map”, or partial understanding of a word’s meaning, may be gleaned from a single reading. However, a more sophisticated
understanding can only come with repeated exposures. Teacher instruction adds to the power of the read-aloud by focusing the students on pertinent information and encouraging deeper processing of the language to which they are exposed.

Content Area Vocabulary Instruction

Duke and Bennett-Armistead (2003) assert that the primary purpose of what they term “informational text” is to convey information about the natural or social world. This informational text makes up the bulk of the texts read by students in the content areas (math, science, social studies). It comes in the form of trade books and discipline area textbooks. Informational text differs from narrative text in its content, structure, organizational patterns, purpose, and technical vocabulary (Guastello, Beasley, & Sinatra, 2000; Harvey, 1998; Winters, 2001). The complexity of expository text increases when there is a large number of unknown words on the page (Ediger, 2000). Science reading, as well as the reading in other content areas, is a much slower process than the reading of narrative text in that one reads it to understand specific concepts, and therefore it is dense with terms and explanation of terms that are essential for understanding (Robertson, 2007). Despite these challenges, proficiency in content area reading is important for students to develop over the course of their school years (Bravo & Cervetti, 2008).

The ability to read and comprehend content area text can assist students in a variety of ways. Informational text is the primary
genre read by most adults in home and work environments, so an introduction to these texts in the primary grades will help prepare students for reading in adulthood and will help link home and school literacies (Duke, Bennett-Armistead & Roberts, 2003). Informational texts are also the texts used in upper grade content area instruction, so early exposure will give students the skills they need for dealing with content areas textbooks in later years (Moss, 2005).

Informational texts motivate children by providing them with answers to the questions they have about the world around them, which in turn builds background knowledge and schema (Misulis, 2009; Winters, 2001; Yopp & Yopp, 2006). For some children, informational text is a preferred genre (Duke & Bennett-Armistead, 2003). Content area texts also help build specialized vocabulary that aids students in concept development (Blachowicz & Fisher, 2000; Duke, Bennett-Armistead & Roberts, 2003; Harvey, 1998; Yopp & Yopp, 2006).

There is a scarcity of informational text in the elementary classroom (Moss, 2005). Yopp and Yopp (2000) state that only 14% of materials that primary grade teachers reported reading aloud on a given day were informational. Further investigation (Yopp & Yopp, 2006) led them to the conclusion that young children also have very little exposure to informational texts at home. Knowing the unique and significant contribution that informational texts can have on student learning and its scarcity in the elementary classroom, it is important for educators to expose young children to more expository text during the school day.
The study of vocabulary in informational or expository text is somewhat different than vocabulary study in narrative text in the reading/language arts classroom. Armbruster and Nagy (1992) identified three major differences between the two kinds of vocabulary. First, in content area instruction words are much more closely tied to the theme or purpose of the lesson. Second, vocabulary work in narrative text is often about learning a new word for a familiar concept; in content area lessons, the vocabulary work is often about learning an entirely new concept. Third, the vocabulary in content area texts has a higher degree of semantic relatedness. With these differences in mind, it becomes obvious how critical good content area vocabulary instruction is for text and subject area comprehension and learning.

Many of the words in content area vocabulary instruction are low frequency and technical words (Duke & Bennett-Armistead, 2003; Taylor, Mraz, Nichols, Rickelman & Wood, 2009). Despite the challenge that these words might present, students will need to learn them to understand the text and also to develop the concepts necessary for further content area learning (Blachowicz & Obrochta, 2005). This greater concept understanding will allow students to learn more vocabulary from context when it is encountered (Harmon, Wood & Hendrick, 2008). They will also have better text comprehension as a result of increased vocabulary learning. Young (2005), and Espin and Foegen (1996) found that the level of understanding of content area vocabulary was an excellent predictor of a student’s ability to understand the text. Further, Spycher (2009) found that young
children who knew more science vocabulary were able to express their understanding of science concepts more effectively.

Students need a deep and elaborated understanding of vocabulary concepts in the content areas because these words are the keys to conceptual knowledge. A superficial knowledge of a meaning, comprised of either contextual or definitional knowledge alone, will not suffice (Spencer & Guillaume, 2006). Students need to have new vocabulary linked to old knowledge in order to make use of pre-existing schemata (Winters, 2001). They also need clarification on words that have multiple meanings in different content areas (Carlisle, Fleming, Gudbrandsen, 2000).

Harmon, Wood and Hendricks (2008) contend that there are four key principles about vocabulary instruction in the content areas: 1) Vocabulary knowledge is critical for comprehending informational texts, 2) Vocabulary instructional strategies used in language arts classrooms are applicable in the content areas, 3) Informational text features affect vocabulary learning, and 4) Classroom instructional time is needed for vocabulary learning. These four points and the preceding discussion remind us of the importance of vocabulary teaching to student understanding of not only content area vocabulary, but content area texts and concepts.

**Vocabulary Assessment**

Determining a student’s knowledge and understanding of a particular word depends a great deal on how the word is assessed (Baumann, Kame‘enui, & Ash, 2003). Graves (1986) points out that
traditionally word knowledge has been assessed using a multiple-choice format that measures vocabulary knowledge through the recognition of a synonym or a definitional phrase. This method is used widely because of both its convenience and its ability to differentiate between students with no knowledge of a word and those with some knowledge. However, one drawback to this type of assessment is its lack of sensitivity to partial and increasing word knowledge. In addition, multiple choice vocabulary assessments have not always been able to empirically demonstrate the vocabulary-comprehension connection that correlational studies have shown exists. Blachowicz (1999) explains that knowing words well enough to select an appropriate synonym on a multiple choice test may not necessarily be sufficient knowledge for comprehension; there is a distinction between a definition and full conceptual knowledge.

While multiple choice tests can be easier to construct and score, they appear to only measure a certain depth of knowledge. In order to understand what vocabulary tests measure it is important to analyze what students know about the word meanings they correctly and incorrectly identify on tests (Curtis, 1987). Dale (1965) discussed four stages of knowing a word: Stage 1—“I never saw it before”, Stage 2—“I’ve heard of it, but I don’t know what it means”, Stage 3—“I recognize it in context. It has something to do with…”, and Stage 4—“I know it.” Multiple choice vocabulary tests seem to be most effective at distinguishing students who have no knowledge and those who have knowledge beyond the Stage 2 level. However, full Stage 4 knowledge is not required to correctly answer multiple choice items.
Over the last 25 years attempts have been made to increase the contextualization of vocabulary assessment (Pearson, Hiebert, & Kamil, 2007) as well as to design multiple choice assessments that are sensitive to degrees of word knowledge and the incremental nature of word learning. One example of this enhanced multiple choice format involved each target word appearing in the test twice, at two different levels of difficulty (Herman, Anderson, Pearson, & Nagy, 1987). The levels of difficulty were manipulated by the types of distractors provided for the words. Easier items had distracters that were either semantically different or from a different part of speech than the target word; more difficult items had distracters that were semantically close to the target item and were from the same part of speech. However, multiple choice tests, in whatever form they take, assess the receptive aspect of vocabulary knowledge. In order to discriminate between fine gradations of incremental word learning and to measure transferability of vocabulary knowledge to reading comprehension, measurement needs to take place in a productive domain. In other words, can the student use the word in some way? Swanborn and de Glopper (2002) created such an assessment by using a four-point scale to assess answers on a definition task where students were asked to write down as much as they could about the meaning of the target words. The four-point scale allowed for scoring of partial word knowledge.

Another issue in the construction of vocabulary assessments is deciding on what words to test (Biemiller, 2004; Pearson, Hiebert, and Kamil, 2007). Graves (2009) gives criteria for classroom teachers to
use when constructing teacher-made multiple choice vocabulary tests. Biemiller (2004) explains that part of the difficulty in deciding what should be tested is that we lack a comprehensive, accurate listing of most word meanings known by children at certain ages. He cites Dale and O’Rourke’s (1981) *Living Word Vocabulary* as the best available source on when word meanings are likely to be learned. The *Living Word Vocabulary* contains entries on 44,000 words identifying the grade at which each word is known by 67% or more of the children or adults in that grade. Some standardized tests use a “dictionary sampling” method to select words (Curtis, 1987). In this method, words selected for assessment are representative of the percentages of words in each of the word classes included in the dictionary. Vocabulary tests include a large proportion of high frequency words. However, high frequency words cannot be assumed to always be easy items (e.g., fire meaning to let go).

For most classroom teachers decisions about word choice will be driven by the materials they cover in literacy or content area instruction. Using what Beck, McKeown, and Kucan (2002) describe as Tier Two words is a way of teaching vocabulary that mature readers and writers use. By contrast, content area teachers may want to focus on the vocabulary that will allow their students to better understand important concepts in the subject area (Blackowicz & Obrochta, 2005). The National Reading Panel (NICHD, 2000) concluded, through their review of the research, that standardized vocabulary measures did not seem to be sufficiently sensitive for use as the only dependent measure in research. They suggested that teacher-generated or
experimenter-generated measures be used as one component of the evaluation. They further suggested the use of more than a single measure of vocabulary as a sound methodology.

One method of teacher and/or experimenter-generated assessment used to evaluate competency in both vocabulary and comprehension skill development is the cloze method (Reutzel & Cooter, 1999; Steinman, 2002). Cloze is a procedure in which the nth word of a passage is deleted and the reader is required to provide the missing words. It is different from a “fill in the blank” activity in that it is applied to an entire passage, not a single sentence, and is therefore contextualized. A maze assessment is a modification of the cloze activity with the typical blank space of a cloze activity being replaced by three choices, the target word and two distracters (Espin & Foegen, 1996; Gillingham & Garner, 1992). The difficulty of the task can be manipulated by the choice of the distracters. While cloze and maze assessments require the student to have an understanding of the vocabulary needed to complete the passage, they are actually more valuable assessments of comprehension. In order to choose the correct missing word in a passage, the student needs to have an understanding of the passage’s content and how a particular word fits into the overall development of the passage. Cloze and maze assessments have been used experimentally to assess intrasentence and intersentence comprehension, as well as passage comprehension (Gillingham and Garner, 1992).

Finally, Read’s analysis of vocabulary assessment, while first conceptualized for ESL learners, can inform the study of vocabulary
assessment for all learners (Read, 2000; Read and Chappelle, 2001). Read and his colleagues identified three continua to evaluate assessments. The Discrete-Embedded continuum describes whether vocabulary is regarded as a separate construct distinct from comprehension (Discrete) or forms part of the assessment of some other, larger construct (Embedded). The Selective-Comprehensive continuum refers to the relationship between the target items on the test and the hypothetical population of vocabulary that the items represent. In a Selective measure specific vocabulary items are the focus of the assessment (e.g., words in a particular content area chapter); a Comprehensive measure takes into account the entire vocabulary content of the material, including reading and listening tasks and the student response of writing and speaking. The Context-Dependent-Context-Independent continuum refers to the degrees that textual context is required to determine the meaning of a word. A vocabulary measure that is Context-Independent is one in which the student can provide a correct response without referring to any context; one that is Context-Dependent is one that measures the student’s ability to take into account contextual information in order to provide a correct response.

In the next section I will outline a theoretical framework that takes into account both direct instruction and text reading as avenues for learning new vocabulary.
Theoretical Framework

In order for a theoretical framework to be comprehensive enough to explain vocabulary acquisition, it would have to account for learning through direct instruction, as well as through reading. I propose that the contributions of two theories can do just that for our understanding of vocabulary development. The separate yet complimentary influences of automatic information processing theory and schema theory can inform our understanding about how learning takes place under both conditions. The incremental nature of word learning and the inter-relatedness of words will prove to be important factors in understanding how these two theories can be applied to vocabulary development (Stahl, 1999).

Theory of Automatic Information Processing

During the 1970s and 1980s several models of information processing were described by a number of different theorists in the field of educational psychology (McInerney, 2005; Wood, 1998). These theories sought to explain complex cognitive processes using models. One theory in particular has been used in the last three decades to explain the acquisition of higher level reading abilities, including fluency and comprehension competency. The LaBerge-Samuels model of automatic information processing is an information processing theory that places attention and attentional resources at the heart of the model (LaBerge & Samuels, 1974). The authors of the original model describe attention as having two components: internal and external. External attention is manifested in orienting behavior. A person
directs his eyes and ears to the source of informational input. External attention is an important prerequisite for learning in general and learning to read in particular.

Internal attention is far more difficult to observe, but can be described by what has become known as the “cocktail party phenomenon.” Cocktail parties are frequently crowded gatherings where many conversations are taking place simultaneously. You may be talking to a friend when you hear a conversation of interest behind you. Without turning your head or giving any other outward sign to either your friend or the person behind you, you begin to switch attention back and forth to take in as much of both conversations as you can. In that case, your internal attentional resources shift back and forth between the two information sources in which you are interested (Samuels, 2004).

Internal attention has three characteristics: alertness, selectivity, and limited capacity (Massaro & Cowan, 1993; Samuels, 2004). Alertness can be thought of as vigilance or the active attempt to come in contact with the source(s) of information. Selectivity is the ability to select what to attend to among many competing sources of information. Limited capacity describes the finite nature of our attentional resources. When we are learning a complicated skill, the demands of learning the skill use up our attentional resources, and we can only attend to one task at a time. Novice drivers experience this phenomenon while more experienced drivers are able to perform driving tasks more fluently while simultaneously listening to music or talking to a passenger (Logan, 1997; Thurlow & van den Broek, 1997).
In automatic information processing theory, automaticity is described as the ability to perform a task with little attention. Automatic processing is fast, effortless, autonomous (in that it begins and runs to completion without intention), and is not available to consciousness (Logan, 1997; Thurlow & van den Broek). When certain tasks become automatic, attentional resources are freed up to deal with other higher level tasks (Samuels & Flor, 1997).

Beginning readers must use attention in order to decode. Therefore, if a beginning reader also wants to comprehend what is read, he has to switch his attention from decoding to comprehension (Blum & Koskinen, 1991; Schwanenflugel, et al, 2006). This process is slow, laborious, and sometimes frustrating to the reader. In contrast, experienced readers can decode with automaticity and thereby have attentional resources free to work on comprehending the text.

In the LaBerge-Samuels model (1974) there are several components or processing stages, including visual memory, phonological memory, episodic memory, and semantic memory. The LaBerge Samuels model is not a linear model; there are feedback loops at several points. What is most important to take into account is that with practice and increased skill, the processing systems can become more accurate and then eventually automatic, which will free up attentional resources for the construction of meaning from text. There is an important distinction to be made between accuracy and automaticity. A student can be accurate without being automatic. In general, if a student is automatic, there is a high level of accuracy combined with speed.
Other Views on Automaticity

While the theory of automaticity has held up to the scrutiny of the research community, there have been challenges to the centrality of attentional resources to the model. Logan (1988a, 1988b) and Stanovich (1990) have suggested that automaticity may be viewed as a memory phenomenon.

The theory assumes that novices begin with a general algorithm that is sufficient to perform the task. As they gain experience, they learn specific solutions to specific problems, which they retrieve when they encounter the same problem again. Then, they can respond with the solution retrieved from memory or the one computed by the algorithm. At some point, they may gain enough experience to respond with a solution from memory on every trial and abandon the algorithm entirely. At that point, their performance is automatic. Automatization reflects a transition from algorithm-based performance to memory-based performance.

(Logan, 1988b, p.501).

To return to our beginning reader example, the beginning reader who laboriously sounds out a word is using an algorithm, whereas the reader who has had extensive practice can recognize the word quickly and automatically as a memory phenomenon.

The consequences of automaticity are upheld if you view the underlying mechanism controlling automaticity as attentional resources or a memory phenomenon. The important factor in either case is that
enough exposure to and practice with text allows for automaticity to
develop. Only then will the reader be able to “unglue from print” and
work on comprehension (Chall, 1996).

Empirical work in the area of automatic information processing
theory has been focused on increasing reading fluency as a factor
leading to improved reading comprehension. Kuhn and Stahl’s fluency
research review (2003) concluded that it is unclear whether the
mechanism behind fluency’s influence on comprehension is automatic
word recognition or the contribution of prosody and phrasal
boundaries, but clearly the issue of practice and repeated exposure to
text is important in understanding fluency’s contribution to
comprehension.

Several vocabulary researchers (Beck, Perfetti, & McKeown, 1982;
McKeown, Beck, Omanson, & Perfetti, 1983; McKeown, Beck, Omanson, &
Pople, 1985; Stahl & Fairbanks, 1986) have looked at the impact of
multiple exposures of target words on the learning of vocabulary, and
its impact on comprehension. They have found that multiple exposures
allowed for more automatic processing of the words in text and greater
growth in measures of vocabulary knowledge and comprehension.

When students are encouraged to read widely, they have an
increased exposure to text and practice with the reading task. With
increased exposure and practice, automaticity develops (Kuhn,
Schwanenflugel, & Meisinger, 2010). When reading is automatic,
cognitive resources are freed up for comprehension, as well as
attention to the meanings of words in the passage. Therefore,
vocabulary development takes place as a consequence of a more
“mindful” reading of the text that can only happen when automaticity is achieved. While automatic information processing theory is one of the information processing theories that informs our understanding of vocabulary development and learning, another information processing theory that has much to tell us about vocabulary is schema theory.

**Schema Theory**

Schema theory is an information processing model that explains the comprehension of text using prior knowledge (Pressley, 2000). A central premise of the theory is that much of our knowledge is stored in complex relational structures (or schemata) that help people understand events more easily. Schema theorists believe that comprehension is a matter of activating or constructing a schema that provides a coherent explanation of objects and events in a passage (McVee, Dunsmore & Gavelek, 2005). Connectionist theory, as presented by Adams, (1990) is an extension of the theory to include a representational microstructure and learning mechanism that some theorists felt was missing from the original discussion on schema.

According to schema theory, reading involves a simultaneous analysis at many different levels. The levels include graphophonemic, morphemic, semantic, syntactic, pragmatic, and interpretive (Anderson, 2004). Reading is an interactive process where analysis does not proceed in a “bottom up” fashion from print to text interpretation. Instead, a reader’s schema helps inform the interpretation the reader brings to the text as it is being read. When a piece of text is ambiguous or invites inference, a schema will help the reader make
choices that can maintain comprehension in moments of confusion. For instance, in a paragraph about wrestling, the reader’s “wrestling schema” will lead him to interpret the word “lock” as a hold placed on one wrestler by another. The word “lock” would be interpreted quite differently if the paragraph were about a merchant closing up his shop for the night. The hypothesis the reader has already formulated about the text calls up a certain schema that then tips the scales in the direction of one of the two possible meanings for “lock” (Anderson, 2004; McVee, Dunsmore, & Gavelek, 2005).

The reader’s schema affects both learning and recall of information and ideas in the text. Researchers have discussed six functions of schemata. They provide the basis for: 1) assimilating text information, 2) making inferential elaborations that fill in the gaps in messages, 3) allocating attention to important text elements, 4) searching memory in an orderly fashion, 5) formulating a summary of information, and 6) making inferences that can enable one to reconstruct an original message despite having forgotten some of the details (Bransford, 2004). These functions illustrate why the knowledge possessed by the learner, in interaction with the text, has pervasive effects on performance. The schema helps the reader interpret the text within a structure that aids understanding and facilitates comprehension.

One implication of the theory is that some children may appear to have poor comprehension or memory skills, but in actuality they may lack or fail to activate the schema necessary for successful text comprehension (Derry, 1996). Children can begin developing schematic
representations for recurring events in their lives from an early age (Pressley, 2000). This information helps them draw inferences about stories told or read to them. The construction or elaboration of schemata continues throughout life as students are exposed to new experiences, oral language, and text. The more elaborate and complete a reader’s schema, the faster and more automatically the reader can process the text. This is shown by the fact that readers spend less time reading text on topics that are familiar to them (Anderson, 2004).

There has been much empirical work on comprehension development that uses schema theory as its theoretical foundation (Anderson & Pearson, 1984; Prince & Mancus, 1987; Spires, Gallini, & Riggsbee, 1992). This research supports the fact that increased understanding of text comes with the development, activation, and elaboration of schema. Vocabulary development, which is a natural correlate of comprehension, has not been equally explored using this theoretical framework.

**How the Two Theories Inform Vocabulary Instruction**

Although both automatic information processing theory and schema theory have been used to explain comprehension processes, they can also be useful in describing vocabulary processes. There is a longstanding agreement in the field of reading that there is a vocabulary comprehension connection (Anderson & Freebody, 1981; Mezynski, 1983). This connection has been the basis for a good deal of vocabulary research (Beck & McKeown, 1991; Stahl & Fairbanks,
Theories that are developed to explain comprehension processes will necessarily have an influence on our understanding of vocabulary processes. The two competencies are linked in a symbiotic relationship. It is important that a reader understand the meanings of enough of the words in a text to make sense of the entire passage. Conversely, an understanding of the overall concepts in a text will help the reader develop and refine his knowledge about individual word meanings.

As presented earlier in this chapter, students increase and expand their reading vocabularies by two means: direct instruction and wide reading (Graves, 2006). Automatic information processing theory and schema theory inform our understanding of those two types of vocabulary learning, although there has been little discussion in the literature to this point that uses these theories to explain vocabulary learning.

Direct vocabulary instruction, especially when it takes the form of rich vocabulary instruction as proposed by McKeown & Beck (2004), helps students both build and elaborate schemata. When teachers help children explore the relationships between words and understand nuances in word meanings, they are in essence teaching for schema development. With our understanding that words are interrelated comes the knowledge that the more intense and richly connected the word work is, the more elaborate, flexible, and useful the schema becomes.

During wide reading, the increased exposure to a variety of texts helps students to elaborate and refine pre-existing schema (Scott, 2005). Those enriched schemata aid students in their ability to
derive meanings from context (Stahl, 1999). Schema building and vocabulary development are additive processes and do not take place from one exposure to text (Nagy, 2005). Students need to see words repeatedly in texts before they can be said to have "learned" a word. The incremental nature of word learning (Nagy & Scott, 2000) accounts for how these repeated exposures help students refine their knowledge of words.

Each theory helps inform our understanding of how reading vocabulary develops. However, the interaction of the two theories is a more potent explanation that takes into account the two sources for word learning: direct instruction and wide reading (Graves, 2006).

Early vocabulary research looked at either direct instruction of words and strategies (Beck, Perfetti, & McKeown, 1982; McKeown, Beck, Omanson & Perfetti, 1983) or wide reading (Nagy, Anderson & Herman, 1987; Nagy, Herman & Anderson, 1985). Presently, the consensus in the field is that an effective program will contain both strategies (Kame'enui & Baumann, 2004; Nagy, 1988; Stahl, 1999). Only by taking into account the LaBerge Samuels model of automatic information processing theory and schema theory do we have the theoretical underpinnings to explain vocabulary learning through both direct instruction and wide reading. When we only look at one theory we limit our understanding of the vocabulary learning task in all of its richness (Thurlow & van den Broek, 1997). The two theories do indeed have some parallels, but it is important to acknowledge that they have unique emphases that privilege certain central issues. Automatic information processing theory and the construct of automaticity have
more to tell us about why wide reading is important to vocabulary
development (Samuels & Flor, 1997). Schema theory can help us
understand why direct instruction has its place in vocabulary
learning, as well as why wide reading makes for richer and more
sophisticated vocabulary knowledge (Rumelhart, 1981). The interaction
of direct instruction and wide reading helps readers develop,
maintain, and enrich their vocabularies by the most efficient means.
It is important to understand how the interaction of the two theories
has a place in the explanation. When automaticity is achieved, schema
knowledge can be more efficiently accessed and applied. Automatic
readers will have cognitive resources in reserve to call up schemata
to use in making sense of words in the text (Logan, 1997). These
elaborated schemata make subsequent reading on a topic more efficient
and automatic. The reader can anticipate what will come next using his
particular schema. This makes for a more fluent reading of the text
(Anderson, 2004).

Wide reading and vocabulary development are connected in a loop
that helps the reader develop greater automaticity, as well as
vocabulary knowledge and comprehension competency (Nagy & Scott,
1990). This is what Stanovich (1986) referred to as the Matthew
Effect. The more reading a student does, the more competency he
develops. This competency (and the confidence that accompanies it)
leads to more reading, and the increase in reading again leads to
greater competency in a “rich get richer” spiral. This competency can
be overall comprehension competency, but for the purposes of this
discussion the focus remains on vocabulary development. The more we
read the more chances we have to refine our knowledge of word meanings and discern the nuances of contextualized vocabulary.

It is the interaction of the increased cognitive resources of the automatic reader combined with the reader’s elaborated schema use that best explains the phenomenon of the increase in vocabulary knowledge to the extent that it develops during the school years. However, this theoretical explanation has not been discussed in the vocabulary learning research and represents a gap in the research that the proposed study will attempt to address.

A theoretical explanation of vocabulary learning using automatic information processing and schema theories will privilege certain instructional strategies over others. These theories help us understand the importance of fluent reading and support instructional methodologies that promote reading as one means of vocabulary development. They remind us that word knowledge should be rich, interrelated, and contextualized to be most useful, and therefore so should our direct instruction methodologies. Therefore, this theoretical framework recommends a method of instruction that includes a rich, elaborated form of direct instruction followed by opportunities to repeatedly be exposed to those words in connected text. One way of repeatedly exposing students to vocabulary in connected text is through repeated reading. Another way is to read a variety of texts that contain the same target words. The theoretical underpinning for the utility of repeated reading can be seen in the earlier presentation of automatic information processing theory. However, the extended hypothesis that reading multiple texts
containing the same target vocabulary will be more beneficial for vocabulary growth than repeated reading of the same text comes from the work of Perfetti and colleagues (Bolger, Balass, Landen, & Perfetti, 2008; Perfetti, Wlotko, & Hart, 2005; Reichle & Perfetti, 2003) on context variability theory.

Context variability theory uses an instance-based theoretical framework (Logan, 1997) to account for the influence of context and definitions on vocabulary learning. In the theory, acquisition of meaning is described as incremental with each incidental exposure to a word in context. While an initial experience with a word in context can result in vocabulary learning, these initial meanings are somewhat fragile and malleable. Vocabulary knowledge gained from exposure to context is rarely full and complete, and is much more likely to be partial. This partial knowledge may overemphasize the situational properties of a word’s meaning as opposed to the more abstract, decontextualized knowledge of a word’s meaning. However, the authors (Bolger et al, 2008) hypothesize that contexts can activate related words through a passive resonance process. This resonance mechanism would cause words already in the reader’s corpus that are related to the context to be activated. These words, along with the words in the context, would become associated with the new word. What is actually learned is a weak pattern of association that would become part of the word’s associative meaning. Since each instance (encounter with a word) sets down a contextualized episode, the learner’s understanding of a word’s meaning will depend on his degraded memory over a series of context-specific instances (Reichle & Perfetti, 2003). More varied
contexts will supply more traces for this resonance. Each individual trace may be weak since it has not been repeated, but the variety of contexts will provide more traces for the resonance process. A single context that is repeated will have stronger but fewer memory traces, leading to a more contextualize, less abstract understanding of the meaning of the word. In multiple encounters with varied texts, abstraction occurs over instances as the context portion of the memory fades and the meaning features shared over contexts become central. If a context variability hypothesis is applied, one would expect children to have greater vocabulary growth from readings of the target vocabulary in different contexts than from repeated reading of the target vocabulary in the same context.

Based on a review of the preceding literature, it is clear that there has been a great deal of research in the area of vocabulary. The theoretical framework explores how the contributions of two theories can contribute to our understanding of the vocabulary learning process. However, an identified gap in the literature is in understanding how to combine direct instruction of vocabulary and the reading of that vocabulary in connected text to make the best use of limited instructional time. The purpose of this study is to explore what the most efficient method of that follow-up reading would be in order to increase vocabulary knowledge. In the next chapter I will delineate just how this study will attempt to do that.
CHAPTER 3

METHODOLOGY

The focus of this research is to compare the effectiveness of two different kinds of text practice following content area vocabulary instruction. In order to achieve this goal an experimental study was undertaken to teach content area vocabulary to three groups of students and then compare three conditions: a repeated reading condition using a single text, a condition where multiple related texts were read, and a control condition where texts were read that were unrelated to the target vocabulary. In addition, qualitative data were also collected and analyzed as a means of contextualizing the quantitative findings.

Participants

A total of 36 third grade students participated in the study (12 students in each of three groups). The study was conducted in a public elementary school (Kindergarten through third grade) located in a suburban community in the northeast. The school has a diverse population of over 700 students. The ethnic makeup of the school is 15.2% white, 7.6% African American, 5.1% Hispanic, and 72.1% Asian/Indian. Average class size is 20.1 students (NJ Department of Education, 2009). There are eight third grade classrooms in the school.

A review of the research shows that while children as young as preschoolers can benefit from vocabulary instruction (Robbins & Ehri, 1994; Senechal, 1997), most of the very early vocabulary growth comes
in an oral /aural format. When children are first learning to read (early primary grades) they are learning to decode words that are already in their oral vocabularies (Graves, 2006). During the third and fourth grade years, and continuing throughout their schooling and into adulthood, students primarily increase their vocabulary through the reading that they do. Therefore, third grade students were chosen for this study as a way of looking at students who are beginning to make the shift to learning vocabulary from text. Students in English as a Second Language classes were not included in the study since their vocabulary knowledge is dependent on their status as second language learners. Students classified as Eligible for Special Education and Related Services were not included in the study since the present research is not focused on students with diagnosed learning disabilities, and many of these children do not presently receive their reading instruction in the mainstream classroom.

The sampling procedure took place in the following manner. Six third grade teachers were recruited to participate from the eight presently teaching in the school. A verbal and written explanation of the study was made to all of the third grade teachers. Then a request for participation was sent through email and through flyers in the teachers’ mailboxes at school. Four teachers immediately volunteered to participate. All other teachers were then approached to gauge interest. Two of the four remaining teachers who were approached agreed to participate; the other two declined.

Permissions were sought from parents to have their children pre-tested using the Gates MacGinitie Reading Tests, Fourth Edition
Consent forms were sent home and I was available to speak about the study to interested parents during the third grade Back-to-School Night. The consent form made it clear that participation in the Gates MacGinitie Pre-test did not guarantee participation in the actual intervention and that only six students from each classroom would be chosen to participate in the study. All consent forms are included in Appendices A-C.

While the student permissions were being sought, the six participating teachers were asked to list half hour blocks of time when their students would be able to participate. In submitting their lists, teachers had to take into account such scheduling factors as lunch and special area classes (gym, art, music), as well as periods when some of their students might be involved in Basic Skills reading, Basic Skills math, bilingual Chinese, English as a Second Language, Speech or Occupational Therapy. All of these different areas of the curriculum put restrictions on individual teacher schedules that had to be considered. While all classes in the school are grouped heterogeneously, the teachers were also asked to generally describe their classes as high, medium, or low in order that no condition would be formed from two high or two low groups. Scheduling restraints and availability, and teacher ratings of their classes as high, medium, or low were used to form three groups. Each group consisted of twelve students (six from one class, and six from a second class). Two classrooms were used to supply students for each condition in an attempt to limit the influence of teacher interest, style, skill, or curriculum emphasis on the vocabulary knowledge of students in any
particular condition. A determination to use 12 participants in each condition was made for two reasons: 1) it was felt that a total of six students taken from each classroom would not be unduly disruptive to the teacher’s program, 2) the space constraints in the school make a group of 12 students the maximum that could be comfortably accommodated and managed in the third grade wing during instructional time.

Students for whom permission was obtained were administered the Gates MacGinitie Vocabulary subtest. The scores from eligible students in each particular class were ranked. The list of ranked scores for each class was divided into two groupings (High and Low scorers) at the median class score. From each classroom’s list, the top three scorers and the bottom three scorers were chosen to participate. Therefore, each condition included six high scorers (three from each of two classes) and six low scorers (three from each of two classes) for a total of 12 subjects for the particular condition. Table 1 shows how the groups were formed.
Table 1

Participants in Study Groups

<table>
<thead>
<tr>
<th>Group A</th>
<th>Group B</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 students from 2 assigned classrooms (6 students from classroom 1 and 6 from classroom 2). Of the 6 students from each classroom, the top 3 scorers and the bottom 3 scorers were chosen to participate.</td>
<td>12 students from 2 assigned classrooms (6 students from classroom 3 and 6 from classroom 4). Of the 6 students from each classroom, the top 3 scorers and the bottom 3 scorers were chosen to participate.</td>
<td>12 students from 2 assigned classrooms (6 students from classroom 5 and 6 from classroom 6). Of the 6 students from each classroom, the top 3 scorers and the bottom 3 scorers were chosen to participate.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Classroom 1</th>
<th>Classroom 2</th>
<th>Classroom 3</th>
<th>Classroom 4</th>
<th>Classroom 5</th>
<th>Classroom 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1A</td>
<td>High</td>
<td>S1B</td>
<td>High</td>
<td>S1C</td>
<td>High</td>
</tr>
<tr>
<td>S2A</td>
<td>High</td>
<td>S2B</td>
<td>High</td>
<td>S2C</td>
<td>High</td>
</tr>
<tr>
<td>S3A</td>
<td>High</td>
<td>S3B</td>
<td>High</td>
<td>S3C</td>
<td>High</td>
</tr>
<tr>
<td>S4A</td>
<td>Low</td>
<td>S4B</td>
<td>Low</td>
<td>S4C</td>
<td>Low</td>
</tr>
<tr>
<td>S5A</td>
<td>Low</td>
<td>S5B</td>
<td>Low</td>
<td>S5C</td>
<td>Low</td>
</tr>
<tr>
<td>S6A</td>
<td>Low</td>
<td>S6B</td>
<td>Low</td>
<td>S6C</td>
<td>Low</td>
</tr>
<tr>
<td>S7A</td>
<td>High</td>
<td>S7B</td>
<td>High</td>
<td>S7C</td>
<td>High</td>
</tr>
<tr>
<td>S8A</td>
<td>High</td>
<td>S8B</td>
<td>High</td>
<td>S8C</td>
<td>High</td>
</tr>
<tr>
<td>S9A</td>
<td>High</td>
<td>S9B</td>
<td>High</td>
<td>S9C</td>
<td>High</td>
</tr>
<tr>
<td>S10A</td>
<td>Low</td>
<td>S10B</td>
<td>Low</td>
<td>S10C</td>
<td>Low</td>
</tr>
<tr>
<td>S11A</td>
<td>Low</td>
<td>S11B</td>
<td>Low</td>
<td>S11C</td>
<td>Low</td>
</tr>
<tr>
<td>S12A</td>
<td>Low</td>
<td>S12B</td>
<td>Low</td>
<td>S12C</td>
<td>Low</td>
</tr>
</tbody>
</table>

After reviewing each teacher’s available times and considering the teacher’s overall class ranking (high, medium, or low), Group A, Group B and the Control Group were formed. One of the groups could
meet in the morning and the other two groups could meet in the afternoon. In order that the two experimental groups remained as equivalent as possible in all factors, I chose to make the two afternoon groups the experimental groups (Group A and Group B) and the morning group the Control group. An ANOVA was run on the Gates Vocabulary Subtest scores to determine that there were no significant differences between the three groups before the start of the intervention. No significant differences were found among the groups at that time.

**Materials**

Materials used during this study consisted of both Assessment and Intervention materials. Each group of materials will be discussed separately in the sections that follow.

**Assessment Materials**

Assessment instruments were chosen in order to measure student performance in both a receptive and a productive format using both standardized and curriculum-based methods. The use of multiple measures is in keeping with the findings and recommendations of the National Reading Panel (NICHD, 2000). In this way students were assessed on the target vocabulary that they were taught, as well as on a more global sample of vocabulary. In order to achieve this goal, the vocabulary and comprehension subtests of the Gates MacGinitie Reading tests were used as the standardized measures, and two
curriculum-based measures (Productive Vocabulary test and Maze Comprehension test) were developed for this study.

The Gates MacGinitie Reading Test (MacGinitie, MacGinitie, Maria, & Dreyer, 2000) is a group administered survey test that assesses student achievement in reading. The third grade level test consists of two subtests. The vocabulary subtest measures reading vocabulary by presenting each word in a brief context intended to suggest a part of speech but not provide clues to meaning. Students select a word or phrase that is closest to the target word from four given choices. The comprehension subtest measures students’ abilities to read and understand different types of prose. All of the passages are taken from published books and periodicals. Some questions require a literal understanding of the passage; others require students to make inferences and draw conclusions. The derived scores that can be obtained from the Gates MacGinitie include National Percentile Ranks, Normal Curve Equivalents, National Stanines, Extended Scale Scores, and Grade Equivalents. Reliability coefficients indicate strong total test and subtest internal consistency levels with values at or above .90. Alternate form correlations for the total test scores were at or above .90. The authors presented evidence for validity based on high correlations with the Third Edition of the Gates-MacGinitie, which has strong validity indicators, as well as piloting and careful procedures for the development of the Fourth Edition (Spies and Plake, 2005).

The Productive Vocabulary test was developed to assess vocabulary knowledge in a productive format. Students were given the target words with two lines following each word. Students were then required
to write a definition or a meaningful sentence for each word. The 20 target words were taken from the texts that were used during the vocabulary intervention. There were five target words for each of the four science topics used for the intervention. The test was scored using a four-point scale that allowed me to examine partial word knowledge in keeping with a theory of the incremental nature of word learning (Swanborn & de Glopper, 2002). The scale was developed by Swanborn and de Glopper (2002) for a study where they measured incidental word learning from context in response to different reading purposes. In that study, two raters independently scored all assessments, and this resulted in an inter-rater reliability of Cohen’s kappa .94. To use the scale, a score of 0 was given for a wrong answer; a score of 1 was given when there was some association with the target word’s meaning; a 2 was given for reasonably complete word knowledge (might be constrained by context and not show that the word could function in other contexts as well); a 3 was given to decontextualized, full word knowledge.

All of the productive vocabulary tests were scored by an additional scorer (primary grade teacher) to assess inter-rater reliability and resolve any ambiguities in the scoring procedures. A coefficient of .948 was obtained for the pre-test ratings, a coefficient of .968 was obtained for the post-test ratings, and a coefficient of .971 was obtained for the delayed post-test ratings.

The Maze Comprehension Test was developed by writing four informative passages using the target vocabulary for each of the four
topics. Third grade teachers reviewed these passages to determine if their readability and content were appropriate for a third grade reader. Maze is a curriculum-based measure that has been found to be a valid measure of reading comprehension (Parker, Hasbrouck, & Tindal, 1992; Wayman, Wallace, Wiley, Ticha, & Espin, 2007; Wiley & Deno, 2005). The Maze test for this study was created by leaving the first sentence intact and then beginning with the second sentence dropping out words using a lexical deletion pattern, whereby nouns, verbs, and adjectives are deleted instead of using a fixed ratio deletion where every nth word is deleted. The lexical deletion pattern was not shown to have any consistent differences in difficulty, reliability, or validity from the fixed ratio deletion (Parker, Hasbrouck, & Tindal, 1992). Assessing the ability to choose lexically important words made for a more valid measure of content understanding for the present study. Scores were determined by the number of combined correct words chosen in all four passages.

**Intervention Materials**

Literacy instruction in the school chosen for this study takes place in the framework of a literacy block that includes guided reading groups, read-alouds, independent reading, shared writing, and writer’s workshop. Science instruction consists of kit-based units that are supplemented with related literature. All teachers follow the same curriculum and use the same materials. The science topics chosen for the study were four topics that were not covered in the third grade curriculum for the district. They were chosen so that
students would not be influenced by prior classroom discussion and instruction, and would be formally exposed to the content area vocabulary for the first time during the study. Three books were chosen for each of the four science topics to be used for instruction. The topics and book titles were the following:

**Volcanoes**

*I Can Read About Earthquakes and Volcanoes* by Deborah Merrians**

*Volcanoes* by Lily Wood

*Volcanoes* by Seymour Simon

**Whales**

*Whales* by Kevin J. Holmes**

*Baby Whales Drink Milk* by Barbara Juster Esbensen

*Splash! A Book About Whales and Dolphins* by Melvin and Gilda Berger

**Moon**

*(First Starts) The Moon* by Lesley Sims**

*Where Does the Moon Go?* by Sidney Rosen

*The Moon Book* by Gail Gibbons

**Hurricanes**

*Can You Believe? Hurricanes* by Sandra Markle**

*Wild Weather: Hurricanes!* by Lorraine Jean Hopping

*Hurricanes* by Seymour Simon

   Among the three books selected for each topic, the book chosen for the Repeated Reading condition (the starred text in the list) was one that I felt would hold students’ interest over three readings. Books for Repeated Reading also dealt with the vocabulary concepts in a motivating and accessible manner. All books were chosen in
consultation with the third grade teachers in the building. The books were either already in their classroom libraries as books that could be read during free time or ones that teachers thought would be appropriate for the grade level. During the intervention all duplicates of study books were taken out of the classroom teachers’ libraries in an attempt to control the exposure to texts for the different groups.

The following words were chosen as target words for each of the four science topics in the intervention and are found in all three books related to each topic:

<table>
<thead>
<tr>
<th>Volcanoes</th>
<th>Whales</th>
<th>Moon</th>
<th>Hurricane</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash</td>
<td>Mammals</td>
<td>Craters</td>
<td>Equator</td>
</tr>
<tr>
<td>Crust</td>
<td>Calf/Calves</td>
<td>Satellite</td>
<td>Eye</td>
</tr>
<tr>
<td>Magma</td>
<td>Blubber</td>
<td>Orbits</td>
<td>Cyclone</td>
</tr>
<tr>
<td>Lava</td>
<td>Blowhole</td>
<td>Gravity</td>
<td>Storm Surge</td>
</tr>
<tr>
<td>Erupt/Eruption</td>
<td>Baleen</td>
<td>Phases</td>
<td>Tropical/Tropics</td>
</tr>
</tbody>
</table>

The vocabulary words used in the intervention were chosen because they were common to all of the books for the particular topic and they were words that were essential to understanding the important concepts embedded in the books. Most of the student-friendly definitions used in the study came from the Collins Cobuild Student’s Dictionary (2005). When the Cobuild dictionary did not include a particular word, I wrote a student-friendly definition using the
principles of the Cobuild dictionary (complete sentences, written in plain language).

All chosen words were then cross-referenced with Dale and Rourke’s *Living Word Vocabulary* (1981). The *Living Word Vocabulary* contains 44,000 word meanings and the grade level at which the word was first known by 67% or more of children or adults. The grade range reported in the LWV is from 4th grade to 16+ since the information for this resource was collected using written tests. However, Biemiller (2004) states that a grade level 2 could be formed by using words known by 81% or more of the children in grade 4. Using that criterion, the many level 4 and level 6 ratings for the vocabulary chosen would suggest that those words would be appropriate for use with third graders in independent reading. There were only five words chosen that were not at a 4th or 6th grade level according to the LWV. Those five words and definitions were reviewed and kept in the study due to their necessity in understanding the grade level texts and concepts presented in the texts, and the fact that student-friendly definitions were written that were not beyond a third grader’s conceptual understanding. In Table 2 the words are listed with the student-friendly definitions and data from the Living Word Vocabulary.

Table 2

*Target Vocabulary, Definitions, and Living Word Vocabulary Data*

<table>
<thead>
<tr>
<th>Word</th>
<th>LWV Grade Level</th>
<th>LWV Score</th>
<th>Definition from COBUILD or written for the study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash</td>
<td>4</td>
<td>80%</td>
<td>Ash is the grey powder-like substance that is left after</td>
</tr>
</tbody>
</table>
something is burned.

<table>
<thead>
<tr>
<th>Term</th>
<th>Frequency</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crust</td>
<td>6</td>
<td>76%</td>
</tr>
<tr>
<td>Magma</td>
<td>13</td>
<td>67%</td>
</tr>
<tr>
<td>Lava</td>
<td>4</td>
<td>67%</td>
</tr>
<tr>
<td>Erupt</td>
<td>6</td>
<td>77%</td>
</tr>
<tr>
<td>Mammals</td>
<td>8</td>
<td>73%</td>
</tr>
<tr>
<td>Calf</td>
<td>4</td>
<td>80%</td>
</tr>
<tr>
<td>Blubber</td>
<td>6</td>
<td>85%</td>
</tr>
<tr>
<td>Blowhole</td>
<td>13</td>
<td>42%</td>
</tr>
<tr>
<td>Baleen</td>
<td>16</td>
<td>44%</td>
</tr>
<tr>
<td>Craters</td>
<td>4</td>
<td>71%</td>
</tr>
<tr>
<td>Satellite</td>
<td>6</td>
<td>65%</td>
</tr>
<tr>
<td>Orbit</td>
<td>4</td>
<td>68%</td>
</tr>
</tbody>
</table>

The earth’s crust is its outer layer.
Magma is the melted rock under the earth’s surface that erupts from volcanoes, making lava, cinder, or ash.
Lava is the very hot liquid rock that comes out of a volcano.
When a volcano erupts, it throws out a lot of lava, ash, and steam.
Mammals are animals that give birth to babies rather than laying eggs, and feed their young with milk.
The young of some animals, such as elephants, giraffes, and whales, are called calves.
Blubber is a thick layer of fat between the skin and the muscle layer of whales.
A blowhole is an opening for breathing found on the top of a whale’s head.
Baleen are fringed plates that grow from the upper jaw of some whales. They are used to strain food from the water.
A crater is a large hole in the ground, which has been caused by something hitting it or by an explosion.
A satellite is a natural object in space that moves around a planet or star.
An orbit is the curved path followed by an object going around a planet, a moon, or the sun.
<table>
<thead>
<tr>
<th>Word</th>
<th>Count</th>
<th>Percentage</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gravity</td>
<td>4</td>
<td>67%</td>
<td>Gravity is the force that makes objects attract each other.</td>
</tr>
<tr>
<td>Phases</td>
<td>10</td>
<td>74%</td>
<td>The monthly change in the appearance of the moon as seen from Earth.</td>
</tr>
<tr>
<td>Equator</td>
<td>4</td>
<td>78%</td>
<td>The equator is an imaginary line around the middle of the earth, halfway between the North and South poles.</td>
</tr>
<tr>
<td>Eye</td>
<td>4</td>
<td>86%</td>
<td>The eye of a hurricane is the fairly calm center of the storm.</td>
</tr>
<tr>
<td>Cyclone</td>
<td>4</td>
<td>76%</td>
<td>A cyclone is a violent tropical storm.</td>
</tr>
<tr>
<td>Storm surge</td>
<td></td>
<td></td>
<td>A storm surge is a rush of water onto the shore that is caused by a big storm.</td>
</tr>
<tr>
<td>Tropical</td>
<td>4</td>
<td>73%</td>
<td>Tropical means belonging to the tropics; the tropics are the hottest part of the earth, near the equator.</td>
</tr>
</tbody>
</table>

An analysis of the target vocabulary found in each book showed that the majority of the books had multiple instances of the target words. Four of the books contained a target word that only appeared once. Since this is a study of readings from the literature found in elementary classrooms, variability in the number of times the target words appear in a particular book would be expected. The following table shows the number of times each target word appears in each text.
Table 3

Vocabulary Count for Books in the Study

<table>
<thead>
<tr>
<th>Target Vocabulary</th>
<th>Earthquakes and Volcanoes by Deborah Merrians</th>
<th>Volcanoes by Lily Wood</th>
<th>Volcanoes by Seymour Simon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash</td>
<td>4</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>Crust</td>
<td>13</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Magma</td>
<td>8</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Lava</td>
<td>11</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Erupt/Eruption</td>
<td>9</td>
<td>30</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td><strong>Whales by Kevin J. Holmes</strong></td>
<td><strong>Baby Whales Drink Milk by Barbara Juster Esbensen</strong></td>
<td><strong>Splash! By Melvin and Gilda Berger</strong></td>
</tr>
<tr>
<td>Mammals</td>
<td>5</td>
<td>9</td>
<td>2</td>
</tr>
<tr>
<td>Calf/Calves</td>
<td>11</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Blubber</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Blowhole</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Baleen</td>
<td>17</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td><strong>First Starts: The Moon by Lesley Sims</strong></td>
<td><strong>Where Does the Moon Go by Sidney Rosen</strong></td>
<td><strong>The Moon Book by Gail Gibbons</strong></td>
</tr>
<tr>
<td>Craters</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Satellite</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Orbits</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Gravity</td>
<td>6</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Phases</td>
<td>2</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>
Can You Believe? Hurricanes by Sandra Markle | Wild Weather: Hurricanes! By Lorraine Jean Hopping | Hurricanes by Seymour Simon
---|---|---
Equator | 3 | 1 | 2
Eye | 15 | 16 | 12
Cyclone | 2 | 2 | 7
Storm Surge | 4 | 4 | 1
Tropics/Tropical | 4 | 10 | 11

**Intervention Design**

The intervention was designed for implementation over 12 sessions to take place three times a week over a four week period. Additional sessions before and after the intervention were necessary for pre- and post-testing. Delayed post-test sessions took place three months after the intervention. I acted as teacher/researcher for this study. As a former teacher (both general education and special education) I had the experience and knowledge to guide the students through the lessons. With one person working with all three groups, I could control the language of instruction and maintain a uniformity of presentation. However, since different groups of learners come to instruction with a variety of factors that influence the learning, instruction that is consistently delivered is not always consistently received (Lave & Wenger, 1991). The qualitative data helped describe both the consistency of instruction and how that instruction was received by the three different groups of students.

Lessons were designed in accordance with vocabulary research recommendations reviewed in Chapter 2. Lessons combined definitional
and contextual information, encouraged active processing of word meanings, and provided multiple exposures to the target vocabulary (Stahl, 1986; Stahl, 1999). Three readings of a book (or group of books) were chosen for the study to replicate the most common number of readings in the literature regarding vocabulary growth after repeated reading (Coyne, McCoach, & Kapp, 2007; Elley, 1989; Leung, 1992, 2008; Senechal, 1997). Research on repeated reading notes the greatest amount of growth on a variety of factors taking place between the third and the fifth readings (Dowhower, 1989).

Each week/unit began with a lesson where target vocabulary was introduced using an adaptation of Blachowicz and Obrochta’s Vocabulary Visits methodology (2005; 2007). Vocabulary Visits creates virtual field trips for the classroom using texts, interactive learning with visuals, and other activities. On the first day of the unit, a poster was presented that illustrated the particular science theme to stimulate discussion. Students were asked to talk briefly about some of the things they knew about the science topic. As the students named words related to the topic, I recorded the words on post-it notes and put them on the poster. I added any target words to the post-it notes that were not generated by the students. Next, I reviewed all of the words listed and provided the students with student-friendly definitions for the words. After the definitions were given, the first book was introduced. Students worked in dyads to partner read the book. Both Group A-Repeated Reading and Group B-Wide Reading read the same book on the first day of instruction.
During the second session of each week, the students participated in a review of the target words that included an activity such as Semantic Mapping (Rupley & Nichols, 2005; Stahl, 2005) or the Connect Two Game (Graves, 2009). They then worked in dyads to read the same book a second time (Group A-Repeated Reading) or read another book containing the same target vocabulary (Group B-Wide Reading). A list of words and definitions written on chart paper was available for all students to see as they read. On the third day students worked in dyads to read the same book a third time (Group A-Repeated Reading) or read another book containing the same target vocabulary (Group B-Wide Reading). After this reading the students worked in pairs on writing a paragraph using as many target words as they could. They were encouraged to use the book and vocabulary definition chart as resources. A combination of group work and reading in dyads was used to foster the active processing of vocabulary that comes when students discuss their ideas with others (Blachowicz & Lee, 1991; Stahl, 1999; Stahl & Vancil, 1986). Dyad groupings were changed over the course of the lessons to limit the influence of any particular dyad (weaker student/stronger student, two weaker students, etc.) on study results.

Group C, the control group, was exposed to the same introductory lesson as the other two groups. Those students also had the vocabulary definition chart on display during the lessons, and worked on the same activities on the second and third days of each week. However, the texts that they read were unrelated to the chosen science topic. For instance, if the topic was volcanoes, they were exposed to
the original volcano lesson but read books about dinosaurs, rivers, etc.

For the first two weeks the students were encouraged to post-it any pages where they came across a vocabulary word that had been introduced. In reviewing the analytic notes of the lessons, it became apparent that while the placing of post-its on vocabulary words was motivating to the students, it tended to distract them from actually reading connected text in the study books. Therefore, students were no longer instructed to post-it vocabulary in the books after the second week of the study.

During the intervention, I worked with each group three days a week for sessions lasting approximately 30 minutes. All lessons took place at the time originally scheduled with the exception of the last lesson for the Wide Reading Group (Group B) that was rescheduled for earlier in the day because of a classroom party. Each group participated in three lessons on each of the four topics. Detailed lesson plans for each unit are included in Appendix D. All lessons took place in a small classroom that was being used as an auxiliary teacher’s room on the second floor of the building down the hall from the third grade classrooms.

**Data Collection and Analysis**

Both qualitative and quantitative data were collected for this study. The study was originally designed to look at quantitative data. In addition, qualitative data were collected as a way to
further understand the quantitative results. Each data source will be presented separately in the following sections.

Quantitative Data

Test data were collected at three different points for this study: pre-test, post-test, and delayed post-test. The design for the analysis was a three factor mixed design with two fixed variables and one repeated variable. The fixed variables were the three conditions (Group A - repeated reading, Group B - wide reading, Group C - control group) as one factor and the pre-intervention performance of students on the Gates MacGinitie (High or Low scoring students) as the other factor. The repeated measure was the scores on the assessment at three different times (pre-test, post-test, and delayed post-test). A separate repeated measures analysis was run for each of the four assessments (Gates Vocabulary subtest, Gates Comprehension subtest, Productive Vocabulary Test, Maze Comprehension Test). Raw scores were used as the analyzed score for each of the assessments. Data were analyzed using PASW Statistics Grad Pack 18 (SPSS, Inc., 2009). Table 4 illustrates the design of the study.
Table 4

**Design of the Study**

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Group B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Control Group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Scorers</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After the pre-test was completed, an overall ANOVA was run on the Gates MacGinitie Vocabulary subtest scores to confirm the equivalency of the groups in the three conditions at the outset of the intervention. No significant differences were found among the groups at the outset. However, as anticipated, differences were found between the high and low groups since they were separated at the inception of the study precisely so that particular factor could be examined.

The null hypothesis ($H_0$) for the overall three factor mixed design ANOVA for each of the assessments was that there were no differences between conditions. In addition, another null hypothesis was that there was no interaction between the condition and the high/low scoring groups.
Qualitative Data

Qualitative data were used as a way of providing a richer explanation of the quantitative findings. Qualitative data were derived from three data sources: observational notes taken as students were engaged in reading partnerships, observations of the lessons caught on videotape, and analytic memos written as soon after the lesson as possible. The multiple qualitative data sources allowed me to triangulate my findings and better explain the quantitative results.

Collection of data. Observational notes on reading partnerships were made as the reading took place. Analytic memos were made as soon after the completion of the lessons as possible, often right after the lesson but always within an hour of the end of the lesson. These notes contained comments about overall lesson effectiveness, as well as successes and challenges taking place during the lessons. All notes and analytic memos were later typed. Lesson videotapes were reviewed and a narrative transcription was written describing the activity taking place, the time frame of each lesson component, and what individual students were doing throughout the lesson. The videotapes and memos were used to confirm the validity of implementation of the model, to confirm the consistency of lesson presentation among the three conditions, and to note student affect and responses during the lessons.

Due to some technical difficulties with equipment at the beginning of the study, the first four days of lessons were not captured by the digital recorder. All other lessons were captured
with the exception of the eighth day’s lesson for the Repeated Reading group (Group A). In addition, the footage of the fourth day’s lesson for that same group was cut off after 7:01 minutes.

Observational notes and anecdotal memos were taken during and immediately after lessons. The digitally recorded lessons were not reviewed and transcribed until several weeks after the conclusion of the lessons. The data were collected at different times in an attempt to view the lessons from different perspectives. By collecting some data immediately (observational notes and anecdotal memos), my observations and recollections were very fresh, but quite possibly influenced by my concerns and impressions regarding how lessons were proceeding at the time. By contrast, taking other data some weeks after the conclusion of study lessons (transcripts from digitally recorded lessons) there was some “emotional distance” from the lessons and they could be viewed with the objectivity that time can provide.

One example of this change in perspective from the two data sources was evident in my changing view of two students from one of the groups. Both boys in this particular group were behaviorally challenging. My notes after lessons often mentioned one of these students who was particularly non-compliant to teacher demands. My conjecture from memory and reading over the anecdotal notes was that this student was the most disruptive to the learning of others. When the recordings were reviewed it became apparent that while this student was the most openly defiant to me, the other of the two boys was actually more disruptive to student learning by his call outs and general activity level. This example showed me how the multiple
sources of data could help me build a fuller picture of what was actually going on and keep me from letting preconceived notions and assumptions dominate my interpretations.

Coding of data. Coding of qualitative data is “the formal representation of analytical thinking” (Marshall & Rossman, 2006, p.160). It helps the researcher to develop themes by which the data can be analyzed. All qualitative data were read and re-read multiple times in preparation for coding. During this repeated reading of data certain themes related to student behavior and reaction to instruction emerged. Those themes and examples from the study data are included in Table 5.

Table 5

Qualitative Codes and Examples from the Study Data

<table>
<thead>
<tr>
<th>Code</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty with Decoding/ Comprehension</td>
<td>Tithi is struggling to read (silently). She looks miserable. When I go over to read with her for a moment it is apparent that she is struggling to decode many of the words that the other children are reading with ease.</td>
</tr>
<tr>
<td>Distractions/Interruptions from Instruction</td>
<td>Copier making noise for the first two minutes of the lesson.</td>
</tr>
<tr>
<td>Motivated</td>
<td>As I started putting up the words, some of the kids caught on to the fact that the words were being</td>
</tr>
</tbody>
</table>
placed in categories. They were very excited to help figure out the categories and started waving their hands and bouncing up and down until I called on them.

<table>
<thead>
<tr>
<th>Not Motivated</th>
<th>Eric was less cooperative today. He didn’t want to sit on the floor.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused</td>
<td>Hrishi and Anjali are very into finding vocabulary and finding new facts. They are hunched over their book and call out excitedly when either finds a word or fact.</td>
</tr>
<tr>
<td>Not Focused</td>
<td>Christina is not paying attention as Craig gives a long explanation. Now she is playing with her necklace.</td>
</tr>
</tbody>
</table>

After codes were established all data were re-read and coded. Data were read separately for each code that was chosen. As coding progressed it became apparent that many observations that were coded “Motivated” were also coded “Focused.” Therefore, all codings of “Motivated” and “Focused” were reviewed to check that these two themes were not describing the same phenomenon. Following this analysis it was decided that although many data points could be described as showing instances of “Motivated” and “Focused” behavior, they were two different factors that co-occurred less than a third of the time. Consequently, both codings were retained for analysis. An example of an observation coded Focused, but not coded Motivated is, “Ariana is
more focused when she’s working in a dyad than during the group lesson.” An example of an observation coded Motivated but not coded Focused is, “Sara and Anjali are very enthusiastic when reading. They read loudly, jump up, and sometimes act out what they read.”

As will be further explicated in Chapter 4, the qualitative data collected helped me to comment on student engagement and motivation, as well as the consistency of lesson implementation across the three groups. The qualitative data can be viewed as a source of contextualization for the quantitative findings. In Chapter 5 the qualitative data also helps to clarify requirements in furthering this line of research beyond this study.
CHAPTER 4

FINDINGS

Two different kinds of data were collected during this study. The study was designed to analyze the quantitative results obtained before and after a four week intervention looking at text practice following vocabulary instruction. In addition, qualitative data were collected as a means of contextualizing the quantitative results. In this chapter I will present the quantitative findings first. The qualitative findings that follow that presentation help provide context in which to more fully interpret the quantitative findings.

Quantitative Findings

For each of the four dependent variables discussed below, there was a within-subjects component to the analysis and a between-subjects component. For the within-subjects component there were four hypotheses tested (TestTime, TestTime*Group, TestTime*High/Low, TestTime*Group*High/Low). For the Between-Subjects component there were three hypotheses tested (Group, High/Low, Group*High/Low). There were three levels of TestTime: Pre-test, Post-test, and Delayed Post-test. Three groups were examined: Repeated Reading, Wide Reading, and Control. The High/Low factor contained two levels: High and Low. In the sections that follow, descriptive statistics and findings will be presented for each of the four dependent variables (Gates Vocabulary Subtest, Gates Comprehension Subtest, Productive Vocabulary Test, Maze Comprehension Test).
Gates MacGinitie Vocabulary Subtest

The Gates MacGinitie Vocabulary Subtest (2000) requires students to read words embedded in a brief context and choose the closest synonym from four choices. The means and standard deviations for this measure are shown in Table 8.

Table 8

Descriptive Statistics by Group and High/Low Factor for Gates MacGinite Vocabulary Subtest

<table>
<thead>
<tr>
<th></th>
<th>Control (n=12)</th>
<th>Wide Reading (n=12)</th>
<th>Repeated Reading (n=12)</th>
<th>Hi/Low Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>36.42 6.67</td>
<td>33.83 9.62</td>
<td>33.92 9.74</td>
<td>34.72 8.62</td>
</tr>
<tr>
<td>High</td>
<td>41.83 .41</td>
<td>39.33 4.27</td>
<td>41.33 2.42</td>
<td>40.83 2.90</td>
</tr>
<tr>
<td>Low</td>
<td>31.00 5.22</td>
<td>28.33 10.61</td>
<td>26.50 8.41</td>
<td>28.61 8.10</td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>37.42 6.40</td>
<td>36.67 7.92</td>
<td>35.42 9.85</td>
<td>36.50 7.99</td>
</tr>
<tr>
<td>High</td>
<td>41.00 3.57</td>
<td>41.67 2.66</td>
<td>41.67 2.07</td>
<td>41.44 2.68</td>
</tr>
<tr>
<td>Low</td>
<td>33.38 6.82</td>
<td>31.67 8.43</td>
<td>29.17 10.74</td>
<td>31.56 8.51</td>
</tr>
</tbody>
</table>
The main effect for the within-subjects TestTime factor was significant \([F(2,60)=34.602, p=.000]\). The mean scores for the Pre-test, Post-test, and Delayed Post-test were 34.72, 36.50, and 38.25 respectively. Scores increased linearly over TestTime with the highest mean score occurring on the Delayed Post-test. Therefore, student mean scores increased each time they encountered this assessment measure. There was no decrease in score for the Delayed Post-test as would have been expected after a three month hiatus from instruction. The mean and standard deviations are displayed in Table 9.

**Table 9**

*Descriptive Statistics for Gates MacGinitie Vocabulary Subtest by Group across TestTime*

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>40.83</td>
<td>41.44</td>
<td>42.44</td>
</tr>
<tr>
<td></td>
<td>2.90</td>
<td>2.68</td>
<td>2.43</td>
</tr>
<tr>
<td>Wide Reading</td>
<td>28.61</td>
<td>31.56</td>
<td>34.06</td>
</tr>
<tr>
<td></td>
<td>8.90</td>
<td>8.51</td>
<td>8.56</td>
</tr>
<tr>
<td>Repeated Reading</td>
<td>33.97</td>
<td>33.42</td>
<td>37.58</td>
</tr>
<tr>
<td></td>
<td>9.74</td>
<td>9.85</td>
<td>8.97</td>
</tr>
</tbody>
</table>
The main effect for the within-subjects interaction between TestTime and the Hi/Low factor was also found to be significant [F(2,60)=10.375, p=.000]. The mean and standard deviations are presented in Table 10 and Figure 1 displays the ordinal interaction.

**Table 10**

*Descriptive Statistics for High and Low Groups on Gates MacGinitite Vocabulary Subtest across TestTime*

<table>
<thead>
<tr>
<th></th>
<th>Pre-test</th>
<th></th>
<th>Post-test</th>
<th></th>
<th>Delayed Post-test</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>High/Low Groups</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High</td>
<td>40.83</td>
<td>2.90</td>
<td>41.44</td>
<td>2.68</td>
<td>42.44</td>
<td>2.43</td>
</tr>
<tr>
<td>Low</td>
<td>28.61</td>
<td>8.10</td>
<td>31.56</td>
<td>8.51</td>
<td>34.06</td>
<td>8.56</td>
</tr>
</tbody>
</table>
There was a large difference between the mean scores of the high and low groups at the Pre-test. While the scores of both the high and the low groups increased across TestTime, Figure 1 shows that the trend for the high group was not as steep as the trend for the low group.

There were no significant results for the interaction between TestTime and Group [$F(4,60)=1.324, p=.271$] nor for the interaction
between TestTime, Group, and the High/Low factor \[F(4,60)=.781, p=.542]\).

In examining the between-subjects effects there was a significant High/Low difference as anticipated \[F(1,30)=23.867, p=.000\], but no Group effect was found \[F(2,30)=.493, p=.616\]. In addition, the interaction between Group and the High/Low factor was not found to be significant \[F(2,30)=.472, p=.628\].

**Gates MacGinitie Comprehension Subtest**

The Gates MacGinitie Comprehension Subtest requires students to read a series of passages and answer multiple choice questions about the passages. Various questions require either a literal or inferential understanding of what was read. The means and standard deviations for this measure are shown in Table 11.

**Table 11**

*Descriptive Statistics by Group and High/Low Factor for Gates Comprehension Subtest*

<table>
<thead>
<tr>
<th></th>
<th>Control (n=12)</th>
<th>Wide Reading (n=12)</th>
<th>Repeated Reading (n=12)</th>
<th>Hi/Low Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td><strong>Pre-test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>34.58 6.56</td>
<td>33.75 7.63</td>
<td>34.17 9.16</td>
<td>34.17 7.64</td>
</tr>
<tr>
<td>High</td>
<td>37.33 3.50</td>
<td>38.50 3.73</td>
<td>39.33 3.98</td>
<td>38.39 3.62</td>
</tr>
<tr>
<td>Low</td>
<td>31.83 8.01</td>
<td>29.00 7.75</td>
<td>29.00 10.24</td>
<td>29.94 8.32</td>
</tr>
<tr>
<td></td>
<td>Post-test</td>
<td>Delayed Post-test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------------------------------</td>
<td>--------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>38.92 7.96 35.75 9.98 38.83 7.99 37.83 8.57</td>
<td>43.50 3.45 38.25 10.71 39.58 7.40 40.44 7.88</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>43.00 1.67 41.17 5.04 43.17 3.19 42.44 3.48</td>
<td>45.33 1.21 42.50 3.62 43.17 2.71 43.67 2.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low</strong></td>
<td>34.83 9.83 30.33 11.11 34.50 9.22 33.22 9.70</td>
<td>41.67 4.08 34.00 14.00 36.00 9.07 37.22 9.90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The main effect for the within-subjects TestTime factor was significant \([F(2,60)=32.315, p=.000]\). The mean scores for the Pre-test, Post-test, and Delayed Post-test were 34.17, 37.83, and 40.44 respectively. Again, as with the Gates MacGinitie Vocabulary Subtest, scores increased linearly over TestTime with the highest mean score occurring on the Delayed Post-test.
Table 12

Descriptive Statistics for Gates MacGinitie Comprehension Subtest by Group across TestTime

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Control</td>
<td>34.58</td>
<td>6.56</td>
<td>38.92</td>
</tr>
<tr>
<td>Wide Reading</td>
<td>33.75</td>
<td>7.63</td>
<td>35.75</td>
</tr>
<tr>
<td>Repeated Reading</td>
<td>34.17</td>
<td>9.16</td>
<td>38.83</td>
</tr>
<tr>
<td>Total</td>
<td>34.17</td>
<td>7.64</td>
<td>37.83</td>
</tr>
</tbody>
</table>

There were no significant results for the interaction between TestTime and Group [F(4,60)=2.027, p=.102], the interaction between TestTime and High/Low factor [F(2,60)=1.668, p=.197], nor the interaction between TestTime, Group, and High/Low factor [F(4,60)=.384, p=.819].

In examining the between-subjects effects, once again as expected the High/Low factor was significant [F(1,30)=13.105, p=.001] (see Table 13).
Table 13

*Descriptive Statistics for High and Low Groups on Gates MacGinitie Comprehension Subtest across TestTime*

<table>
<thead>
<tr>
<th>High/Low Groups</th>
<th>Pre-test Mean</th>
<th>Pre-test SD</th>
<th>Post-test Mean</th>
<th>Post-test SD</th>
<th>Delayed Post-test Mean</th>
<th>Delayed Post-test SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>38.39</td>
<td>3.62</td>
<td>42.44</td>
<td>3.48</td>
<td>43.67</td>
<td>2.83</td>
</tr>
<tr>
<td>Low</td>
<td>29.94</td>
<td>8.32</td>
<td>33.22</td>
<td>9.70</td>
<td>37.22</td>
<td>9.90</td>
</tr>
</tbody>
</table>

There were no significant between-subjects effects for Group [F(2,30)=.643, p=.533] nor for the interaction between Group and the High/Low factor [F(2,30)=.272, p=.764].

**Productive Vocabulary Test**

The Productive Vocabulary Test measures target vocabulary knowledge in a productive format. Students read each of the 20 target words and write a definition or meaningful sentence for each. Means and standard deviations for this test are shown in Table 14.
Table 14

*Descriptive Statistics by Group and High/Low Factor for the Productive Vocabulary Test*

<table>
<thead>
<tr>
<th></th>
<th>Control (n=12)</th>
<th>Wide Reading (n=12)</th>
<th>Repeated Reading (n=12)</th>
<th>Hi/Low Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
<td>Mean  SD</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10.83  4.02</td>
<td>9.33  3.89</td>
<td>9.75  4.00</td>
<td>9.97  3.91</td>
</tr>
<tr>
<td>High</td>
<td>12.83  2.79</td>
<td>11.00  3.58</td>
<td>12.00  3.29</td>
<td>11.94  3.13</td>
</tr>
<tr>
<td>Low</td>
<td>8.83  4.26</td>
<td>7.67  3.72</td>
<td>7.50  3.51</td>
<td>8.00  3.66</td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>36.75  6.80</td>
<td>33.58  9.95</td>
<td>25.33  9.28</td>
<td>31.89  9.82</td>
</tr>
<tr>
<td>High</td>
<td>38.00  5.93</td>
<td>37.17  6.65</td>
<td>30.33  5.89</td>
<td>35.17  6.78</td>
</tr>
<tr>
<td>Low</td>
<td>35.50  7.92</td>
<td>30.00  11.95</td>
<td>20.33  9.73</td>
<td>28.61  11.40</td>
</tr>
<tr>
<td>Delayed Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>29.00  5.43</td>
<td>27.17  9.09</td>
<td>22.25  5.28</td>
<td>26.14  7.24</td>
</tr>
<tr>
<td>High</td>
<td>30.50  5.50</td>
<td>30.67  8.80</td>
<td>25.00  1.67</td>
<td>28.72  6.31</td>
</tr>
<tr>
<td>Low</td>
<td>27.50  5.39</td>
<td>23.67  8.66</td>
<td>19.50  6.35</td>
<td>23.56  7.33</td>
</tr>
</tbody>
</table>

The main effect for the within-subjects TestTime factor on the Productive Vocabulary Test was significant \[F(2,60)=295.774, p=.000\].
The mean scores for the Pre-test, Post-test, and Delayed Post-test were 9.97, 31.89, 26.14 respectively. Scores increased from the Pre-test to the Post-test. The Delayed Post-test scores demonstrate expected regression in performance on a test given three months after the instruction. The mean and standard deviations are displayed in Table 15.

Table 15

Descriptive Statistics for Productive Vocabulary Test by Group across TestTime

<table>
<thead>
<tr>
<th>Group</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Control</td>
<td>10.83</td>
<td>4.02</td>
<td>36.75</td>
</tr>
<tr>
<td>Wide Reading</td>
<td>9.33</td>
<td>3.89</td>
<td>33.58</td>
</tr>
<tr>
<td>Repeated Reading</td>
<td>9.75</td>
<td>4.00</td>
<td>25.33</td>
</tr>
<tr>
<td>Total</td>
<td>9.97</td>
<td>3.91</td>
<td>31.89</td>
</tr>
</tbody>
</table>

There was a significant within-subjects interaction between TestTime and the Group \([F (4,60)= 5.938, p=.000]\). Figure 2 demonstrates the interaction. The trend for growth between the Pre-test and the Post-test was the sharpest for the control group. However, the decline between the Post-test and the Delayed Post-test was also the sharpest for the Control group. The Repeated Reading group had the flattest profile of the three Groups.
There were no significant results for the interaction between TestTime and the High/Low factor \( [F(2,60)=.977, \ p=.382] \), nor for the interaction between TestTime, Group, and the High/Low factor \( [F(4,60)=.834, \ p=.509] \).

In examining the between-subjects effects there was a significant High/Low difference as anticipated \( [F(1,30)=7.941, \ p=.008] \) (see Table 16 for means and standard deviations).
Table 16

*Descriptive Statistics for High and Low Groups on Productive Vocabulary Test across Test Time*

<table>
<thead>
<tr>
<th>High/Low Groups</th>
<th>Pre-test Mean</th>
<th>Pre-test SD</th>
<th>Post-test Mean</th>
<th>Post-test SD</th>
<th>Delayed Post-test Mean</th>
<th>Delayed Post-test SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>11.94</td>
<td>3.13</td>
<td>35.17</td>
<td>6.78</td>
<td>28.72</td>
<td>6.31</td>
</tr>
<tr>
<td>Low</td>
<td>8.00</td>
<td>3.66</td>
<td>28.61</td>
<td>11.40</td>
<td>23.56</td>
<td>7.33</td>
</tr>
</tbody>
</table>

There was also a significant effect for Group \([F(2,30)=4.137, p=.026]\). The Control Group scored significantly higher than the Repeated Reading group on this measure. There was not a significant effect for the interaction between the Group and the High/Low factor \([F(2,30)=.324, p=.725]\).

**Maze Comprehension Test**

The Maze Comprehension Test requires students to choose words (from a choice of three) to fill in blanks in content related paragraphs to measure comprehension. Means and standard deviations for this test are shown in Table 17.
Table 17

Descriptive Statistics by Group and High/Low Factor for the Maze Comprehension Test

<table>
<thead>
<tr>
<th></th>
<th>Control (n=12)</th>
<th>Wide Reading (n=12)</th>
<th>Repeated Reading (n=12)</th>
<th>Hi/Low Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>41.08</td>
<td>3.03</td>
<td>38.92</td>
<td>6.71</td>
</tr>
<tr>
<td>High</td>
<td>42.67</td>
<td>2.80</td>
<td>41.83</td>
<td>4.92</td>
</tr>
<tr>
<td>Low</td>
<td>39.50</td>
<td>2.51</td>
<td>36.00</td>
<td>7.38</td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47.25</td>
<td>2.38</td>
<td>45.17</td>
<td>4.82</td>
</tr>
<tr>
<td>High</td>
<td>48.67</td>
<td>1.21</td>
<td>47.50</td>
<td>1.98</td>
</tr>
<tr>
<td>Low</td>
<td>45.83</td>
<td>2.48</td>
<td>42.83</td>
<td>5.85</td>
</tr>
<tr>
<td>Delayed Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>47.42</td>
<td>1.62</td>
<td>45.58</td>
<td>5.02</td>
</tr>
<tr>
<td>High</td>
<td>48.33</td>
<td>.52</td>
<td>47.33</td>
<td>2.66</td>
</tr>
<tr>
<td>Low</td>
<td>46.50</td>
<td>1.87</td>
<td>43.83</td>
<td>6.40</td>
</tr>
</tbody>
</table>

The main effect for the within-subjects TestTime factor on the Maze Comprehension Test was significant [F(2,60)=90.506, p=.000]. The
mean scores for the Pre-test, Post-test, and Delayed Post-test were 39.58, 46.14, 46.17 respectively. Scores increased from the Pre-test to the Post-test and remained stable from the Post-test to the Delayed Post-test. Therefore, students made gains between the Pre-test and the Post-test, but did not have a significant loss in scores between the Post-test and the Delayed Post-test. Figure 3 demonstrates this growth pattern.

**Figure 3**

*Overall Means for All Instructional Groups on the Maze Comprehension Test*
There were no significant within-subject effects for the interaction between TestTime and Group \([F(4,60)=.222, p=.925]\), nor between TestTime and High/Low factor \([F(2,60)=1.316, p=.276]\), nor between TestTime, Group, and High/Low Factor \([F(4,60)=.498, p=.738]\).

In examining the between-subjects effects there was a significant High/Low difference as anticipated \([F(1,30)=11.046, p=.002]\) (see Table 18 for means and standard deviations).

**Table 18**

*Descriptive Statistics for High and Low Groups on Maze Comprehension Test across TestTime*

<table>
<thead>
<tr>
<th>High/Low Groups</th>
<th>Pre-test</th>
<th>Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>High</td>
<td>42.28</td>
<td>3.61</td>
<td>48.06</td>
</tr>
<tr>
<td>Low</td>
<td>36.89</td>
<td>5.40</td>
<td>44.22</td>
</tr>
</tbody>
</table>

There were no significant between-subject effects for Group \([F(2,30)=.982, p=.386]\) nor for the interaction between Group and the High/Low factor \([F(2,30)=.491, p=.617]\).

**Summary of Quantitative Data**

Specific quantitative findings differed depending on the assessment that was being examined. On the Gates MacGinitie Vocabulary subtest the within-subjects TestTime factor was significant and demonstrated that students made gains from Pre-test to Post-test,
and continued to make gains from Post-test to Delayed post-test. In addition, the within-subjects interaction between TestTime and the High/Low factor demonstrated that while all of the groups’ scores showed an increase from Pre-test to Post-test to Delayed Post-test measures, the trend for the Low group was steeper.

For the Gates MacGinitie Comprehension subtest, again there was a significant within-subjects effect for TestTime with scores increasing linearly from Pre-test to Post-test to Delayed Post-test. There was additionally a significant between-subjects effect for the High/Low factor with the High group scoring better than the Low group as expected.

For the Productive Vocabulary test there was a more expected pattern seen in the significant TestTime effect, where an increase in scores was seen between the Pre-test and the Post-test, but then a decrease in scores was seen between the Post-test and the Delayed Post-test. This decline in the Delayed Post-test scores would be expected on an assessment given three months after instruction. There was also a significant within-subject interaction between TestTime and Group on the Productive Vocabulary Test. For this interaction the Control Group demonstrated a steeper increase in scores from Pre-test to Post-test and a steeper decline in scores from Post-test to Delayed Post-test than either of the other groups. The Repeated Reading Group had the flattest profile of the three groups.

There were two significant between-subjects effects for the Productive Vocabulary test. The first was the expected High/Low differences with the High group scoring better. The second was a
Group effect. Here the Control Group scored better than the Repeated Reading Group.

For the Maze Comprehension Test, the only significant within-subjects effect was TestTime. There was an increase in test scores from the Pre-test to the Post-test. However, the anticipated drop in scores from the Post-test to the Delayed Post-test was not seen; scores essentially stayed the same between the two test administrations.

Students in all three Groups made gains between Pre-test and Post-test measures for all of the assessments. Therefore, instruction was successful in teaching all students new vocabulary and increasing their comprehension scores. However, the text practice that followed the instruction did not significantly affect performance in this study.

**Qualitative Findings**

The qualitative findings will be presented through an examination of the consistency of instructional time among the three groups and through a discussion of the six qualitative themes that emerged during the repeated readings of the qualitative data.

**Consistency of Instructional Time**

One of the important functions of the qualitative data was to help document the consistency of the lessons. The qualitative data showed a general consistency in timing for the lessons that were digitally recorded.
### Table 6

*Number of Minutes Recorded by Lesson Components and Group*

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Activity</th>
<th>Control</th>
<th>Wide</th>
<th>Repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whale Day 2</td>
<td>Review Target Words</td>
<td>3:27</td>
<td>4:38</td>
<td>4:03</td>
</tr>
<tr>
<td></td>
<td>Connect Two Game</td>
<td>6:22</td>
<td>6:17</td>
<td>Recording</td>
</tr>
<tr>
<td></td>
<td>Partner Read</td>
<td>12:44</td>
<td>12:42</td>
<td>Ends at</td>
</tr>
<tr>
<td></td>
<td>Sharing Facts/carpet</td>
<td>4:45</td>
<td>3:53</td>
<td>7:01??</td>
</tr>
<tr>
<td>Whale Day 3</td>
<td>Review Target Words</td>
<td>4:15</td>
<td>4:38</td>
<td>4:31</td>
</tr>
<tr>
<td></td>
<td>Partnership Read/Write</td>
<td>20:32</td>
<td>25:14</td>
<td>21:00</td>
</tr>
<tr>
<td>Moon Day 1</td>
<td>Intro/poster/definitions</td>
<td>9:50</td>
<td>12:52</td>
<td>11:23</td>
</tr>
<tr>
<td></td>
<td>Partner Read</td>
<td>9:38</td>
<td>11:42</td>
<td>11:21</td>
</tr>
<tr>
<td>Moon Day 2</td>
<td>Review Target Words</td>
<td>3:56</td>
<td>3:13</td>
<td>4:25</td>
</tr>
<tr>
<td></td>
<td>Possible Sentences</td>
<td>5:40</td>
<td>6:49</td>
<td>8:29</td>
</tr>
<tr>
<td></td>
<td>Partner Read</td>
<td>12:15</td>
<td>16:51</td>
<td>12:24</td>
</tr>
<tr>
<td>Moon Day 3</td>
<td>Review Target Words</td>
<td>3:26</td>
<td>5:22</td>
<td>3:46</td>
</tr>
<tr>
<td></td>
<td>Partnership Read/Write</td>
<td>21:05</td>
<td>19:52</td>
<td>20:14</td>
</tr>
<tr>
<td>Hurricane Day 1</td>
<td>Intro/poster/definitions</td>
<td>10:34</td>
<td>11:38</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Partner Read</td>
<td>8:34</td>
<td>14:48</td>
<td>Recording</td>
</tr>
<tr>
<td>Hurricane Day 2</td>
<td>Review Target Words</td>
<td>2:20</td>
<td>2:39</td>
<td>4:16</td>
</tr>
<tr>
<td></td>
<td>Four Square Activity</td>
<td>10:21</td>
<td>8:14</td>
<td>8:37</td>
</tr>
<tr>
<td></td>
<td>Partner Read</td>
<td>10:28</td>
<td>12:13</td>
<td>11:30</td>
</tr>
<tr>
<td>Hurricane Day 3</td>
<td>Review Target Words</td>
<td>3:16</td>
<td>2:28</td>
<td>2:42</td>
</tr>
<tr>
<td></td>
<td>Partner Read/Write</td>
<td>20:03</td>
<td>19:59</td>
<td>18:13</td>
</tr>
</tbody>
</table>
The length of the discussions varied due to two identified factors. One factor was the complexity of the conversation and the interest in the topic. Some groups became very interested in a topic and that led to more elaboration during the conversation. An example of that was seen during the Hurricane Day 2 lesson with the Control Group. An extra word was explored using the Four Square Activity because the students were so enthusiastic about giving their ideas. By contrast, some discussions went longer because behavior had to be addressed. In these conversations, students called out or otherwise interrupted one another and those behaviors had to be dealt with on the spot. The Wide Reading Group’s time interval for the Introduction of the poster and vocabulary during the Moon Day 1 lesson was longer than the other two groups for just those reasons. However, overall instructional time was fairly consistent across groups.

Qualitative Themes

Several themes began to emerge as the qualitative data were repeatedly read. In the next section the six themes are discussed separately. A table that represents the instances of coding for each of the themes follows the discussion.

Difficulty with decoding or comprehension. Throughout the course of the study lessons it became apparent that several students struggled to decode and/or comprehend grade level texts. These students coped with this academic frustration in different ways. Some
students retreated into themselves and shied away from reading aloud. This was the case with one of the students in the Wide Reading Group. An example from the Hurricane Day 1 lesson was, “When Tithi gets the book she stays silent and then says, “I’m reading.” Some students began acting out when it was time to read. Another example from the Wide Reading group was, “Jose starts to get loud and bossy as Tithi hesitates.”

The Wide Reading Group had substantially more incidences of coding for this category. Of the 13 instances coded for the Wide Reading Group, 11 of the codings mentioned one of two students who were struggling with the text. These two students were consistently low scorers on all pre-, post-, and delayed post-test assessments. Of the only two instances noted for the Control Group, one of the notes was written about a student who I thought might have been struggling. It was written early in the study and the student turned out to be a competent reader. This may have been a case of me forming an opinion too hastily.

**Distractions and interruptions from instruction.** The Control Group had fewer distractions from instruction than the other two groups. The distractions to this group were mostly due to noise from the copier in the room. There was also an interruption from instruction during one lesson when I didn’t have all of the materials ready. Both the Wide Reading and Repeated Reading groups were instructed in the afternoon. There were more instances of interruptions coded during those sessions. One reason was that there
were more teachers coming in to use the copier in the afternoon. Afternoons appeared to be a more popular time for making copies. When teachers came in to use the copier, they appeared to be as quiet as possible so as not to interrupt the lesson. However, the noise from the copier and the movement around the copier were hard for some of the students to ignore.

In addition, one of the teachers in the Wide Reading group was less organized than the other teachers in the study. There were several occasions when I went to pick up students and this teacher’s students were not ready. In fact, on a couple of occasions I had to wait by the door with the six students from the other class as this teacher finished her math lesson. Furthermore, on one of the study days I found that this particular teacher was out for the day. Unfortunately, she left the substitute a fun video to be shown during the time I would be instructing her students. The students from this group were very resistant to leaving that “break time” to participate in my study on that particular day and the lesson, therefore, started late and included several very unhappy students.

**Motivated.** In reviewing the codings for Motivated, the Control Group looked more motivated than the other two groups even taking into account the missing footage from the Repeated Reading Group. Evidence of this motivation was seen in notes such as, “They were very excited to figure out the categories,” and “They both read enthusiastically and wanted to know if we were going to have time to get back together as a group and report the interesting facts.” While the other two
groups demonstrated fewer instances of “Motivated” behavior, they still displayed it. For instance, for the Wide Reading Group I noted, “I review the directions for writing with an excited Christine and Anoushka. They are motivated to get writing.” For the Repeated Reading Group, “Anjali is excited about the book. She asks me a question.”

At the beginning of the study, all students were instructed to use post-its to mark target vocabulary words during their partnership reading time. When reading through my anecdotal notes during the course of the study, it became obvious to me that although the students found the act of “post-it-ing” motivating, it appeared that it was actually distracting them from focusing on the reading of connected text since they were over-focused on “finding” certain words. Therefore, instead of attending to the reading of the passages and incidentally finding vocabulary along the way, the focus became on searching through the text to just find a particular word without concentrating on the content of the reading. To remedy this situation, I instructed the students to read without the post-it notes after Whale Lesson 1. Some of the motivation that was evident earlier in the study (and coded during the reading of data) may have decreased as a consequence.

Not motivated. There were very few instances coded as Not Motivated for the Control Group. The two instances that were coded included a comment written about one student who came to the lesson in a bad mood and was very quiet. She recovered later in the lesson.
The other coding referred to a comment I wrote about how hard it was to stay motivated working in a triad as opposed to a partnership. By contrast, the Wide Reading Group had 28 instances documented. Sixty one percent of those codings were attributable to one student. That student was very resistant to instruction. Since the completion of the study that student has been referred to and classified by the Child Study Team due to difficulties in the areas of reading and behavioral control. If one other Wide Reading Group student is added to the count, 75% of the codings for Not Motivated for this group can be attributed to these two students. An example of field notes documenting behavior coded as Not Motivated included,

“After watching Jose fool around and avoid reading I went over to speak with him privately. I told him that he hadn’t been doing his best and that I needed him to do a better job. He put his head down and looked defeated. A couple of minutes later he started reading. Jose definitely does not like to be corrected. First he is very silly. However, when reprimanded he gets whiney.”

An example of data coded Not Motivated from the Repeated Reading Group includes, “Isabella is listening as Eric reads. She doesn’t seem that interested, but she’s listening.”

**Focused.** As discussed previously, all items coded as Motivated and as Focused were reviewed to make sure that these two themes were not redundant. It was determined that while Motivated and Focused described different characteristics, they appeared to be correlated.
For each of the groups the number of occurrences of Motivated and Focused was: Control Group 59, Wide Reading Group 30, Repeated Reading Group 22. These numbers are in keeping with the ratios in each of the groups for Motivated and Focused separately. An example of focused behavior from the Control Group was, “Riddhi was attentive today. She was quiet in the group, but working hard in her dyad.” For the Wide Reading Group it was noted on the first day of lessons, “They seemed very involved once they got started.” An example of focused behavior from the Repeated Reading Group included, “Katherine and Hrishi are working nicely together and are focused on the book.” The Control Group appeared to be the most focused of the three groups. That and the fact that the Control Group also seemed the most motivated may help to explain some of the quantitative findings.

**Not focused.** There are far fewer incidences of Not Focused codings for the Control Group than the other two groups. However, there were some instances noted. “Arjun looking around a bit during partner reading,” and “Arushi is a little bit distracted by her pencil as Jacob reads, but gets right to her own reading at her turn.” Examples of instances coded Not Focused for the Repeated Reading Group included “Katherine was distracted when I picked the camera off of the tripod. She gets distracted with whatever is in her hand or on the table. Meghana seems to lose some of her focus when others are reading.” For the Wide Reading Group, the following instances were noted, “Varun can get very silly when he’s reading. He doesn’t pay too much attention when Hadeel reads,” and “Once Josh started on the
writing assignment, Craig did not help and started talking to other kids.”

Taking into account the one and a half days of missing footage for the Repeated Reading Group, the number of instances coded Not Focused is fairly equivalent between the Wide Reading and Repeated Reading Groups. In the Wide Reading Group 19% of the occurrences were attributable to one student and another 21% were attributable to one other student. In the Repeated Reading Group one student was named in 24% of the codings for Not Focused.

**Table 7**

*Frequencies of Coded Themes for Qualitative Data*

<table>
<thead>
<tr>
<th>Theme</th>
<th>Control</th>
<th>Wide</th>
<th>Repeated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty with decoding/comprehension</td>
<td>2</td>
<td>13</td>
<td>7</td>
</tr>
<tr>
<td>Distraction/Interruptions from Instruction</td>
<td>5</td>
<td>14</td>
<td>9</td>
</tr>
<tr>
<td>Motivated</td>
<td>91</td>
<td>50</td>
<td>38</td>
</tr>
<tr>
<td>Not Motivated</td>
<td>2</td>
<td>28</td>
<td>18</td>
</tr>
<tr>
<td>Focused</td>
<td>104</td>
<td>62</td>
<td>41</td>
</tr>
<tr>
<td>Not Focused</td>
<td>15</td>
<td>84</td>
<td>70</td>
</tr>
</tbody>
</table>

**Summary of Qualitative Data**

The coding of data helped to qualitatively “paint a picture” of the three groups. The Control Group experienced the fewest interruptions to instruction, displayed the fewest difficulties with decoding and comprehension, and appeared more motivated and focused.
than the other two groups. The overall impression of the Control Group garnered from the data was that of an attentive, motivated group of students who got along well with each other, participated actively in class discussions, and were easily placed in changing reading dyads. Surprisingly, they did not question the disconnect between the vocabulary instruction/activities and the topics of the follow up books. They remained enthusiastic about sharing information throughout the study. Their teachers were very responsible and respectful of time restraints. These students were always ready when I went to pick them up for the lessons.

The Wide Reading Group had a couple of students who had some real difficulties in decoding and comprehension when reading the study texts. There were also a few students with some challenging behaviors. This set a different tone to our time together. One student had many avoidance behaviors (sneaking to the bathroom) and could show a great deal of resistance when challenged. Another student in the group was motivated to participate but extremely loud and disruptive to the learning of others. In addition, one of the teachers sending students to this group was somewhat disorganized herself and therefore her students were not always ready when I came to pick them up for lessons. This group was challenging to work with since time was needed for behavioral management and I did not have the luxury of time to deal with those needs in this small study. Therefore, there were several instances where I just “soldiered on” in an attempt to get through the lesson. By contrast, if this had been my actual class, I might have stopped the lesson, dealt with the
behavior for a longer period of time, and later gotten back to the content. These behavioral issues may have impacted the quantitative results.

The Repeated Reading Group was less resistant and defiant than the Wide Reading Group, but as a group their behavior was less mature. It is interesting to note that this lack of maturity was displayed in both high and low scorers in the group. One of the high scorers was seen sucking her thumb during one of the sessions. In some instances, this lack of maturity looked like lack of focus. While this group did not grumble or challenge my authority as much as the Wide Reading Group, they weren’t as motivated or focused as the Control Group. They were also instructed during the last period of the day, and many of the students seemed tired at that time.

The unique attributes of the three groups led to a different experience for all three groups despite a consistency of lesson presentation that could be documented in the data. These differences were not controlled for in the design of the experiment but may well have impacted the quantitative results.
CHAPTER 5

DISCUSSION

This chapter will discuss the results presented in Chapter 4 within the context of the previous literature review, delineate the limitations of the study, and make recommendations for classroom practice and further research.

The primary purpose of the study was to examine how text practice following vocabulary instruction influences content area vocabulary learning. In order to do this, students were placed in one of three conditions: a Repeated Reading group, a Wide Reading group, or a Control group. There were twelve students in each group, six high and six low scorers as determined by performance on the Gates MacGinitie Vocabulary subtest prior to instruction. The groups were formed using participants from two different classrooms. Each group was exposed to the same vocabulary lessons based on principles of effective vocabulary instruction (Stahl 1986, 1999). These lessons presented the students with a combination of definitional and contextual information, had them actively process new word meanings, and gave them multiple exposures to the words.

Three lessons were presented for each of the science units. One group, the Repeated Reading group, read a single text over three lessons. The second group, the Wide Reading group, read three different texts that all included the target vocabulary. The third group, the Control group, read three unrelated books that did not contain the target vocabulary. The intervention took place over four weeks; each week presented a new science topic and new books to read.
as follow-up. Pre-test, Post-test, and Delayed Post-test results on two standardized measures and two researcher-developed measures were used to run a series of repeated measure ANOVAs. Qualitative data was also collected as a way of contextualizing the quantitative results. The research questions that guided this study were:

- What is the influence of two different types of text practice on the acquisition of content area vocabulary? Following vocabulary instruction, what are the effects of repeated reading with one text as compared to reading a variety of texts with the same target vocabulary?
- What is the difference in the acquisition of vocabulary knowledge for more proficient readers and less proficient readers?
- What is the difference in vocabulary growth between the repeated one-text group and the multiple-text group after a four week intervention? Are the results consistent three months later?

**Significant Findings**

The significant findings of the study are presented under several categories in order to better understand the role of certain factors on the results. First, the results for readers based on their pre-test ability (High/Low readers) are discussed. Then, group gains between Post-test and Delayed Post-test are discussed. The overall effect for students engaged in the different types of text practice is then presented. The influence of the type of assessment used is followed by a discussion about how group dynamics impacted on the findings.
High and Low Ability Readers

There was a significant High/Low between-subjects effect on all four of the dependent variables. High scorers continued to score better than low scorers throughout the testing sequence. This supports the research that argues that direct instruction does not close the gap for less able learners (Swanborn & de Glopper, 2002; van Daalen-Kapteijns, Elshout-Mohr & de Glopper, 2001). However, there was a surprising result in that the within-subjects interaction between TestTime and High/Low scores for the Gates MacGinitie Vocabulary subtest was significant. In other words, while there continued to be a gap between High and Low scorers in the Pre-test, Post-test and Delayed Post-test scores, the Low scorers had a steeper trend line than the High scorers.

One possible explanation for this finding could be found in the fact that the Low scorers presumably had less word consciousness than the High scorers at the beginning of the study (Anderson & Nagy, 1992; Scott & Nagy, 2004). While the intervention would theoretically increase word consciousness for all students, its significance in this small study might only be seen in these low scorers who had less word consciousness to start, making any gains look significant. This steeper trend for the Low group can be seen on the Gates MacGinitie Vocabulary test because it is a multiple choice test where the answer is on the page and has to be chosen out of an array. On the Productive Vocabulary test, where the student has to produce a definition without any other support, the task may have been too
challenging for this modest increase in word consciousness to be detected.

While there was only a modest increase in the trend line of the Low scorers, it is an interesting finding in that it argues against Stern and Powell’s theory of verbal comprehension (1983). Their theory contends that verbal comprehension describes a person’s ability to understand linguistic material (such as books and lectures), and that more skilled comprehenders are better at learning vocabulary from context because they can distinguish important information from unimportant information and store and retrieve it more efficiently. If Sternberg and Powell’s theory is subscribed, there should be no trend differential between the high and low scorers where the low scorers appear to begin to close the gap. However, the difference in this study is that students weren’t just reading the words in context. They were exposed to rich and direct vocabulary instruction before the reading, and the combination of the two (as was presented in the theoretical framework) may have benefitted the low scorers enough to detect this ordinal interaction in the data.

**Gains from Post-test to Delayed Post-test**

On the two standardized assessments (Gates Vocabulary subtest and Gates Comprehension subtest) students continued to make gains from the Post-test to the Delayed Post-test. This was a surprising finding in that one would expect a decrease in performance after a three month break from instruction. Instead, mean scores increased on both of these measures at the Delayed Post-test administration. This
phenomenon could also be attributed to the additional word consciousness that the four week intervention produced in the students (Anderson & Nagy, 1992; Scott & Nagy, 2004). However, scores went up on the Gates Comprehension subtest as well as the Gates Vocabulary subtest. This could be a result of the strong vocabulary comprehension connection discussed in Chapter 1 (Anderson & Freebody, 1981; NICHD, 2000; Mezynski, 1983; Ouellette, 2006; Stahl & Fairbanks, 1986). If that were the case, the increases in vocabulary may have stimulated and helped maintain the comprehension increases.

Additionally, schema theory could have played a role in this unexpected increase (Anderson, 2004; Anderson & Pearson, 1984; Bransford, 2004). The contextualization of the vocabulary instruction presented and the opportunities for rich conversations during instruction may have led to the construction and elaboration of schemata that the students would not have had otherwise. These more fully elaborated schemata may have led to increased word and content knowledge that helped students better interpret the Gates comprehension passages presented and fueled the increase in their scores long after instruction ended.

This “Schema Theory Effect” could also explain how Post-test and Delayed Post-test scores for the Maze Comprehension test were nearly identical. Students gained schema knowledge through instruction and conversation during lessons and were able to use those elaborated schemata to successfully complete the maze activities on the Post-test administration. According to Bransford (2004), schemata are helpful in making inferential elaborations that fill in the gaps in messages,
which would be especially helpful in completing a maze activity. While that explains the growth from the Pre-test to the Post-test, how can the retention of scores after a three month absence from instruction be explained? It may be that these rich and elaborated schemata were more resistant to being forgotten over time than an individual definition or fact. In instance-based theoretical frameworks, such as Context Variability Theory, the individual encounters with content aides in the acquisition of meaning by providing the learner with multiple exposures to the learning (Bolger et al, 2008; Logan, 1997). These multiple exposures activate other knowledge (in this case words and conceptual knowledge already acquired) to strengthen the pattern of associations that become part of the word or concept’s associative meaning. In this way the conceptual knowledge gained was more resistant to loss from Post-test to Delayed Post-test. Even if there was some loss of individual traces, the majority of the network of traces remained. Therefore, this gain in conceptual knowledge was maintained from the Post-test to the Delayed Post-test.

The Productive Vocabulary Test was the only measure that showed the more typical decrease in scores from the Post-test to the Delayed Post-test. This may be explained by the nature of the task. The Productive Vocabulary Test required the student to recall and produce a definition or a meaningful sentence using the word. Producing a definition or using a word in a novel context are two of the most challenging ways in which to demonstrate word knowledge; they require a high level of expertise about the word. When one considers the
incremental nature of word learning, this ability comes at the later stages (Durso & Shore, 1991; Stahl, 2003). Students may have been able to remember definitions immediately following instruction, but in order to continue to produce those definitions (or use the target vocabulary in an original sentence) they would need to have moved further along the continuum of incremental learning. Therefore, this decrease in the Delayed Post-test scores may be a result of early, incomplete word learning. However, even when consideration is given to the decrease in scores at the Delayed Post-test, much of the gains made between the Pre-test and Post-test were maintained at the Delayed Post-test.

Type of Text Practice

As a result of the research reviewed in Chapter 2, I considered it likely that the Wide Reading Group would score higher on the assessments than either of the other two groups. Based on the research, I felt that the multiple and varied encounters with the vocabulary would lead to a greater number of connections to the words (Bolger et al, 2008; Logan, 1997). Those connections would then translate into enhanced performance on targeted vocabulary and comprehension measures. Those results were not found in this study. Instead, there were no group differences found in three out of the four assessment measures. For the Gates Vocabulary subtest, the Gates Comprehension subtest, and the Maze Comprehension subtest there were no within-subject effects for the interactions between condition and TestTime, nor were there any significant between-subjects effects for
Group. In other words none of the Groups scored significantly higher or lower than the others on these assessments. This is quite possibly due to a small sample size making it hard to detect any difference. The small n may have negatively impacted on the ability to identify statistical significance on several measures. In that case the likelihood of making a Type II error and failing to reject the null hypothesis even though it was false was increased. Therefore, there may have been findings that were missed because of the sample size that would have been detected with a larger sample of the population.

There was a significant between-subjects Group effect for the Productive Vocabulary test. On this test the Control Group scored significantly higher than the Repeated Reading group. This was quite a surprising finding since the Control group did not read any related follow-up texts and the Repeated Reading group had three opportunities to read target vocabulary in connected text following the lessons. Further, there was a significant within-subjects effect for the interaction between TestTime and Group with the Control group having a steeper upward trend in scores from the Pre-test to the Post-test and a steeper downward trend in scores from the Post-test to the Delayed Post-test than the Repeated Reading Group. These results may be attributable to group characteristics not related to the differences in text practice following the instruction. Specific group differences will be discussed later in this chapter. However, the fact that the Control group demonstrated both the sharpest increase between Pre-test and Post-test and then the sharpest decline between Post-test and Delayed Post-test might be explained by how each of the
groups benefitted from instruction. In order to produce definitions and sentences for the Productive Vocabulary test that would gain a higher score on the rating scale, a student needed to have a fuller and more decontextualized understanding of the word (Baumann, Kame’enui, & Ash, 2003; Nagy & Scott, 2000; Stahl & Fairbanks, 1986). If the Control group had the higher jump from Pre-test to Post-test, that would suggest that at least for a time they had command over those decontextualized meanings. However, those higher level definitions may have been harder to maintain over time without more exposures to the vocabulary; those meanings might not yet have been “solidified” for the student. Therefore, there would be a sharper decrease from the Post-test to the Delayed Post-test. By contrast, if the Repeated Reading group had not yet gained command of the fuller decontextualized definitions from the Pre-test to the Post-test, their rating scale scores on the Post-test would be lower and they would actually have less to lose in the three month interim. Therefore, while their test scores declined, the decline was not as sharp as that of the Control Group.

**Impact of Type of Assessment on Results**

As first presented in Chapter 2, determining a student’s knowledge and understanding of a particular word depends a great deal on how the word is assessed (Baumann, Kame’enui, & Ash, 2003). Multiple choice formats have been used traditionally because they are easier to construct and score, and because they differentiate between students with no knowledge of a word and those with some knowledge
(Graves, 1986). However, this type of testing lacks the sensitivity to detect partial and increasing word knowledge. Therefore, it is not surprising that the Productive Vocabulary test was more effective at discriminating between the groups than the Gates Vocabulary test. Its strength may have been in discriminating between degrees of partial word knowledge. On the Gates Vocabulary test a student may have been able to pick a correct answer with a partial associative understanding of the word’s meaning. This is what Dale (1965) termed Stage 3 knowledge (“I recognize it in context. It has something to do with…”). Full Stage 4 knowledge (“I know it”) is not needed to correctly answer multiple choice items. By contrast, the scoring for the Productive Vocabulary test used a four point scale adapted from Swanborn and de Glopper (2002). This scale was designed precisely to allow for measuring partial word knowledge. This gave the Productive Vocabulary test added capacity for discriminating between levels of vocabulary understanding. Surprisingly, what the Productive Vocabulary test was able to detect was that the Control group and the Repeated Reading group’s scores were significantly different. Although, having the Control group outscore the Repeated Reading group was not an expected outcome, it does not diminish the Productive Vocabulary Test’s ability to measure partial word knowledge, and therefore, more closely detect word knowledge growth.
Classrooms are complex societies where students and teachers interact with each other (Ratcliff, Jones, Costner, Savage-Davis, & Hunt, 2010). In classroom settings, engagement is particularly important because it functions as a behavioral conduit through which students’ motivational processes contribute to subsequent learning (Jang, Deci, & Reeve, 2010). Students need to be engaged to be successful in school and groups produce more ideas when members have high rather than low motivation (Bechtoldt, De Dreu, Nijstad, & Choi, 2010; Skinner, Furrer, Marchand, & Kindermann, 2008).

The three groups in this study, although chosen through a process that fostered randomization and group equivalency, proved to be qualitatively different in the way they responded to instruction and group work. All three groups had a mixture of high and low students. All three groups had student participants from two different classrooms. In an attempt to keep the experimental groups as equivalent as possible, the control group was given the one morning slot available, and the two experimental groups were given the two afternoon slots. Despite the planning and forethought put into designing the study, the three groups presented quite differently, and the qualitative data supports that they were, indeed, quite distinctive from one another.

As presented in Chapter 4, the scheduling of the three groups may have impacted on group performance. The room used for instruction was quieter and experienced fewer interruptions during the morning period with the Control group than during either of the afternoon periods.
with the two experimental groups. While this was not the only
difference among the groups, it was one that impacted on the
“classroom climate” by helping to set a calmer and less disjointed
tone to the lessons.

Teacher promptness may have factored into the results. For one
group in particular (Wide Reading Group), one of the two teachers
supplying students for the group was late in having her students ready
on several occasions. This left the six students from the other
teacher waiting with me in the hall for several minutes. When we
finally got down the hall and started the lesson, we were already
behind schedule and several of the children had become restless as we
waited.

Beyond scheduling and teacher promptness, there were several
factors that demonstrated group differences as documented in the
qualitative data. As a group the Control Group was more motivated and
focused, and demonstrated fewer decoding and comprehension problems
while reading in partnerships. Stahl and Vancil’s (1986) research
showed that rich discussion benefitted discussants and non-discussants
alike. However, there is a presumption that non-discussants are
focused and motivated to take in what is being discussed. If students
aren’t focused or motivated to participate in some way, there will be
an impact on their performance (Skinner, Furrer, Marchand, &
Kindermann, 2008). Since the Control Group demonstrated more
instances of motivated and focused behavior and fewer instances
characterized as “not motivated” and “not focused,” it is not
surprising that their performance looked stronger on the one measure
that was able to discriminate between the three groups. The other two groups, with their higher instances of “not focused” and “not motivated” behavior, may have not been able to benefit to the same degree from the instruction provided.

Additionally, students who start off as confident in their abilities engage with the learning task in ways that lead to more success. This success reinforces their initial optimism and continues to encourage engaged learning. By contrast, students with a lack of confidence tend to avoid challenges or engage in tasks with a lack of enthusiasm leading to decreased success, thereby solidifying their original self-doubts (Skinner, Furrer, Marchand, & Kindermann, 2008). If the Control Group were more engaged and focused than the other two groups from the start, increased success would be an expected outcome according to the research on motivation and learning (Guay, Ratelle, & Chanal, 2008; Patall, Cooper, & Wynn, 2007).

**Limitations of the Study**

The most profound limitation to the study would have to be the small sample size. Although a large population of study participants would have been ideal, the restrictions put on me by the availability of space and the needs of the teachers providing students for the study conspired to make an n of 12 the most feasible number for this study. The fact that one group (control) was taught in the morning when the two other (experimental) groups were taught in the afternoon may have had an impact on the results. While there is no definitive proof that that factor had an effect, I wish I could have eliminated
it as a nuisance variable by scheduling all groups in either the morning or the afternoon. Unfortunately, teacher schedules would not allow for this so the time of day for instruction became a possible factor to take into account.

The length of the intervention was another possible limitation to the study. The students were instructed for four weeks. Each week covered a different science topic. The students may have needed a longer exposure to instruction than the three topic lessons afforded them. More exposures to the vocabulary may have fostered additional incremental learning. The one-week duration on each topic may have only given students the opportunity to acquire early, partial knowledge of a word’s meaning. With increased exposure to the word over a longer intervention, quite possibly greater increases in word knowledge might have taken place. Those greater increases may have been more detectable through assessment measures than the partial learning gained through one week’s worth of lessons and readings.

Early in the study a decision was made to take the three highest scorers and the three lowest scorers on the Gates MacGinitie Vocabulary Pre-test from each class. This was done in order to get a good separation between the high and low groups despite the small n. This decision may have allowed some low scorers into the study who, by virtue of their academic frustration prior to the study, affected the groups in a qualitatively negative way. This quite possibly changed the group dynamics for those groups. These particular students had reading difficulties and presented as somewhat disaffected as learners. One has been referred and classified for special education
since the study ended. Two others have been referred to the school’s Intervention and Referral Services Committee so the committee can brainstorm ideas and strategies for their teachers to use in the classroom. A randomized high and low group from each class may not have recruited quite so many high risk students.

**Classroom Implications**

Despite the fact that the data obtained did not make it possible to reject the null hypotheses, there are some classroom implications that can be gleaned from this study. Unfortunately, a recommendation for a particular type of text practice cannot be made as a result of this research. However, this study does support the research regarding the power of direct, rich vocabulary instruction. All of the students benefitted from the direct instruction regardless of what they read immediately following that instruction. Teachers should encourage and facilitate rich conversations in which students deeply process target words and have multiple exposures to those words. In addition, the results of the study would recommend multi-faceted vocabulary assessment to obtain a fuller picture of vocabulary learning. This should include a measure with a productive format to better measure the incremental increase in word knowledge.

**Recommendations for Future Research**

The findings (and lack of findings) of this research raise some questions for future research in the area of vocabulary learning. The issue of the small sample size for this study may have concealed
findings that could possibly be uncovered in a larger study. The issue of text practice following vocabulary instruction in the content areas, therefore, has not been resolved. Another study comparing wide reading and repeated reading using a larger sample may help to shed light on the best practice.

Researchers should continue to use multiple measures for assessing vocabulary growth as recommended by the National Reading Panel (2000). In this way a fuller and more nuanced picture of incremental growth will be available for study. In addition, more research into rating scales such as the one developed by Swanborn and de Glopper (2002) would increase the precision by which researchers and eventually teachers could evaluate vocabulary growth.

Finally, since group dynamics appeared to have an impact on quantitative results, future research should take those factors into account. Teachers are unable to control for certain group dynamics in their classrooms and researchers need to be sensitive to that fact. However, for quantitative research, the more confounding variables that can be eliminated the more confident we can be in the findings obtained.
REFERENCES


Curtis, M. E. (1987). Vocabulary testing and vocabulary instruction. In M. G. McKeown, & M. E. Curtis (Eds.), *The nature of vocabulary


   *Journal of Instructional Psychology, 29*(4), 312-316.


Graves, M. F. (2008). Instruction on individual words: One size does not fit all. In A. E. Farstrup, & S. J. Samuels (Eds.), *What research has to say about vocabulary instruction* (pp. 56-79). Newark, DE.: International Reading Association.


instruction (pp. 150-181). Newark, DE.: International Reading Association.


research to practice (pp. 1-23). Mahwah, NJ: Lawrence Erlbaum Associates.


Children’s Books Cited


APPENDICES

Appendix A          Permission Letter
Appendix B          Video Addendum
Appendix C          Assent Form
Appendix D          Intervention Lesson Plans
Appendix E          Productive Vocabulary Test
Appendix F          Maze Comprehension Test
Appendix G          Examples of Text Used in Study
Dear Parents,

I am a doctoral candidate in the Graduate School of Education at Rutgers University. I also work at Town Center as a Learning Consultant. The school district has allowed me to contact you to ask permission to have your child participate in a research study. My research is on Vocabulary Instruction and as part of that research I will be giving students a reading pre-test first. After the pre-test is given, I will randomly pick 6 students from your child's class to take part in vocabulary lessons with students from other classrooms. There will be a total of 36 students in the study. The lessons will take place for three half hour periods each week for four weeks and will concentrate on strategies for teaching young children vocabulary. The students will participate in the lessons during a half hour of their classroom literacy time and may miss either a read-aloud, mini-lesson, or partner reading time. Post-testing will take place immediately after the lessons are completed and again three months later. Your child will be praised for his or her performance while we are working together and will be able to participate in motivating and educational activities where he/she will learn content area vocabulary that will be used throughout his/her schooling. There are no known risks to your child from participating in the study. His/her grades and classroom placement will not be affected in any way.

You do not have to allow your child to take part in the study. Participation is strictly voluntary. You may withdraw your permission at any time with no penalty to your child. The research is confidential. Confidential means that research records will include some information about your child, such as gender. I will keep this information confidential by storing it in a secure location. The Institutional Review Board and my professors at Rutgers University are the only parties allowed to see the data, except as may be required by law.

If you have any questions about the research or your child's rights as a research subject, you may contact me at lorell.levy@ww-p.org or at (609)716-8330 ext. 6505, or you may contact the Sponsored Programs Administration at:

Rutgers University Institutional Review Board for the Protection of Human Subjects
Office of Research and Sponsored Programs
3 Rutgers Plaza
New Brunswick, NJ 08901-8559
Telephone: (732) 932-0150 ext. 2104
Email: humansubjects@orsp.rutgers.edu

Sincerely,

Lorell Levy

I agree to allow my child (Name)____________________to participate in the research study Vocabulary Instruction.

Signature of Parent or Guardian____________________________ Date___________________
APPENDIX B

Videotape Addendum to Consent Form

You have already agreed to allow your child to participate in my research study of Vocabulary Instruction. I am asking your permission to videotape the lessons as I teach them. The purpose of the videotape is to have a record of me teaching the lesson to the various groups. It will be used to help me judge the consistency with which I present the lessons to the different groups. The camera will be focused on me and not on any individual child. However, during the course of taping, the images and voices of some children may be recorded. Therefore, I am asking for your permission to videotape.

Please note that you do not have to agree to the videotaping in order to have your child participate in the study. If you do not agree to the videotaping, your child will be seated away from the video camera and no images of your child will be recorded.

The recordings will be stored in a locked file in my home office. Only the principal investigator and my professors at Rutgers will have access to the recordings. After the study is completed, the videotapes may be shared with other educators as a training tool.

Your signature on this form grants the investigator (Lorell Levy) permission to videotape lessons that your child will participate in during the above-referenced study. The investigator will not use the video recordings for any other reason than those stated in the consent form.

I give Lorell Levy permission to videotape the lessons she will be teaching to my child

______________________________for the study: Vocabulary Instruction.

Signature of Parent or Guardian:_____________________________________________________________
You are invited to take part in a research study about vocabulary. You will be learning some new science words and read some interesting science books. This study is being conducted by me (Lorell Levy). I work here at Town Center School as a Learning Consultant, but I am also a student at Rutgers University. I am doing this study as part of my dissertation.

If you agree to participate, I will work with you and other third graders three times a week for four weeks. I will also be giving you some vocabulary and reading tests before and after the four weeks of lessons.

This won’t affect your grades or your schoolwork. It will be a fun way to learn some new words. You can stop at any time and you won’t be penalized for not participating. One of your parents will also have to give permission for you to participate in the study. Your parents will be given my phone number in case you and your parents have any questions about the research. They will also have the number for the Office of Research and Sponsored Programs at Rutgers University in case there are any questions about your rights as a research subject. You will be given a copy of this form to keep.

Do you have any questions? You do not have to be in the study. If you agree to participate, please sign below:

Student Signature:____________________________________________  Date: __________________________

Investigator Signature:_________________________________________ Date:____________________________
APPENDIX D

Intervention Lesson Plans

Day One- Volcano Unit

Groups A, B, and C

Procedure:

10 minutes

- Teacher presents poster of volcanoes and students talk about what they know about volcanoes
- Teacher records volcano vocabulary that is generated on post-it notes and places the notes on the poster (any target words that are not named by the students will be placed on the poster by the teacher)
- Teacher reviews all words listed and provides student-friendly definitions for target vocabulary

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read *I Can Read About Earthquakes and Volcanoes*
  - Group B students will read *I Can Read About Earthquakes and Volcanoes*
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be looking to post-it any pages where they find a target word; students in group C will be asked to post-it interesting words about their topic

5 minutes

- The group reconvenes for post reading discussion
  - Did you understand the words in the book?
  - Were you confused about any words?
  - Did you have trouble understanding any part of the book?
  - What was your favorite part? Why?
Day Two- Volcano Unit

Groups A, B, and C

Procedure:

15 minutes

- Teacher will review the target words from the previous day and call the students attention to a chart containing all of the target words with their definitions written out. This resource will be available to the students during this session and the next session
- Using a semantic mapping methodology, the teacher will put the word “volcano” in the middle of a piece of chart paper
- Students will work in groups of four to brainstorm as many words as they can think of that are related to volcano
- The students reconvene in the larger group to share the words they came up with
- As words are named the teacher writes them on the chart paper grouped in broad categories
- Students then discuss how to label those broad categories and why they chose those labels
- The teacher then extends the discussion by encouraging students to notice interrelationships between the words

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read *I Can Read About Earthquakes and Volcanoes*
  - Group B students will read *Volcanoes* (by Lily Wood)
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be encouraged to give the thumbs up sign when they find a target word in the text; students in group C will be asked to look for interesting words about their topic
Day Three- Volcano Unit

Groups A, B, and C

Procedure:

5 minutes

- Teacher briefly reviews the target vocabulary and definitions from the chart on display

15 minutes

- Students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read *I Can Read About Earthquakes and Volcanoes*
  - Group B students will read *Volcanoes* (by Seymour Simon)
  - Group C students will read a book on an unrelated science topic

10 minutes

- Students work in dyads to write a paragraph using as many target words as they can
- The book they have just read and the vocabulary definition chart will be available as resources
Day One- Whale Unit

Groups A, B, and C

Procedure:

10 minutes

- Teacher presents poster of whales and students talk about what they know about whales
- Teacher records whale vocabulary that is generated on post-it notes and places the notes on the poster (any target words that are not named by the students will be placed on the poster by the teacher)
- Teacher reviews all words listed and provides student-friendly definitions for target vocabulary

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read Whales (by Kevin Holmes)
  - Group B students will read Whales (by Kevin Holmes)
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be looking to post-it any pages where they find a target word; students in group C will be asked to post-it interesting words about their topic

5 minutes

- The group reconvenes for post reading discussion
  - Did you understand the words in the book?
  - Were you confused about any words?
  - Did you have trouble understanding any part of the book?
  - What was your favorite part? Why?
Day Two- Whale Unit

Groups A, B, and C

Procedure:

15 minutes

- Teacher will review the target words from the previous day and call the students’ attention to a chart containing all of the target words with their definitions written out. This resource will be available to the students during this session and the next session
- The teacher will lead the students in playing the Connect Two Game by listing the target vocabulary and five other words found in the first day’s whale reading in two columns

<table>
<thead>
<tr>
<th>Mammals</th>
<th>Milk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ocean</td>
<td>Teeth</td>
</tr>
<tr>
<td>Dive</td>
<td>Blubber</td>
</tr>
<tr>
<td>Blowhole</td>
<td>Shrimp</td>
</tr>
<tr>
<td>Baleen</td>
<td>Calf</td>
</tr>
</tbody>
</table>

- Students will then be encouraged to identify similarities or other relationships between any word in the first column and any word in the second column
- Any two words can be accepted as long as the student can justify the relationship
- Students can also make up sentences using the two words
- This activity will take place orally in the whole group

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read *Whales* (by Kevin Holmes)
  - Group B students will read *Baby Whales Drink Milk* (by Barbara Juster Esbensen)
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be encouraged to give the thumbs up sign when they find a target word in the text; students in group C will be asked to look for interesting words about their topic
Day Three- Whale Unit

Groups A, B, and C

Procedure:

5 minutes

• Teacher briefly reviews the target vocabulary and definitions from the chart on display

15 minutes

• Students work in dyads and partner read
• During the reading portion of the lesson:
  o Group A students will read *Whales* (by Kevin J. Holmes)
  o Group B students will read *Splash! A Book About Whales and Dolphins* (by Melvin and Gilda Berger)
  o Group C students will read a book on an unrelated science topic

10 minutes

• Students work in dyads to write a paragraph using as many target words as they can
• The book they have just read and the vocabulary definition chart will be available as resources
Day One- Moon Unit

Groups A, B, and C

Procedure:

10 minutes

• Teacher presents poster of the moon and students talk about what they know about the moon
• Teacher records moon vocabulary that is generated on post-it notes and places the notes on the poster (any target words that are not named by the students will be placed on the poster by the teacher)
• Teacher reviews all words listed and provides student-friendly definitions for target vocabulary

15 minutes

• Book is presented and students work in dyads and partner read
• During the reading portion of the lesson:
  o Group A students will read First Start The Moon (by Lesley Sims)
  o Group B students will read First Start The Moon (by Lesley Sims)
  o Group C students will read a book on an unrelated science topic
• While reading, students in Groups A and B will be looking to post-it any pages where they find a target word; students in group C will be asked to post-it interesting words about their topic

5 minutes

• The group reconvenes for post reading discussion
  o Did you understand the words in the book?
  o Were you confused about any words?
  o Did you have trouble understanding any part of the book?
  o What was your favorite part? Why?
Day Two- Moon Unit

Groups A, B, and C

Procedure:

10 minutes

- Teacher will review the target words from the previous day and call the students’ attention to a chart containing all of the target words with their definitions written out. This resource will be available to the students during this session and the next session.
- The teacher will introduce the Possible Sentences activity by listing the target vocabulary and other words found in the first day’s reading about the moon:
  - Craters  Earth  Moon
  - Satellite  Phases  Pulls
  - Astronaut  Orbits  Small
  - Travels  Gravity  Night
- The teacher will provide short definitions for any words not already discussed.
- Students in Groups A and B will be directed to think of sentences containing at least two of the words; sentences that might be in the book they are going to read today. Students in group C will be directed to think of sentences containing at least two of the words and ones that make sense according to what they know about the moon.
- Student contributions are then written on the chart paper.
- Both accurate and inaccurate sentences are included and not addressed at this time.

15 minutes

- Students work in dyads and partner read.
- During the reading portion of the lesson:
  - Group A students will read *First Starts The Moon* (by Lesley Sims).
  - Group B students will read *Where Does the Moon Go?* (by Sidney Rosen).
  - Group C students will read a book on an unrelated science topic.

5 minutes

- After the reading, the teacher returns to the sentences on the chart and the group discusses whether each sentence could or could not be true based on the reading.
Day Three- Moon Unit

Groups A, B, and C

Procedure:

5 minutes

- Teacher briefly reviews the target vocabulary and definitions from the chart on display

15 minutes

- Students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read First Starts The Moon (by Lesley Sims)
  - Group B students will read The Moon Book (by Gail Gibbons)
  - Group C students will read a book on an unrelated science topic

10 minutes

- Students work in dyads to write a paragraph using as many target words as they can
- The book they have just read and the vocabulary definition chart will be available as resources
Day One- Hurricane Unit

Groups A, B, and C

Procedure:

10 minutes

- Teacher presents poster of a hurricane scene and students talk about what they know about hurricanes
- Teacher records hurricane vocabulary that is generated on post-it notes and places the notes on the poster (any target words that are not named by the students will be placed on the poster by the teacher)
- Teacher reviews all words listed and provides student-friendly definitions for target vocabulary

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read Can You Believe? Hurricanes by Sandra Markle
  - Group B students will read Can You Believe? Hurricanes by Sandra Markle
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be looking to post-it any pages where they find a target word; students in group C will be asked to post-it interesting words about their topic

5 minutes

- The group reconvenes for post reading discussion
  - Did you understand the words in the book?
  - Were you confused about any words?
  - Did you have trouble understanding any part of the book?
  - What was your favorite part? Why?
Day Two- Hurricane Unit

Groups A, B, and C

Procedure:

15 minutes

- Teacher will review the target words from the previous day and call the students’ attention to a chart containing all of the target words with their definitions written out. This resource will be available to the students during this session and the next session
- Students will fold a piece of paper into four quadrants
- In the upper left box, each student will write the first target word
- Have students name examples of the concepts or ways of describing the concept; several of those words are written in the upper right box
- Non-examples are listed and written in the lower right box
- Students write a definition in their own words for the lower left box
- All target vocabulary is reviewed and discussed through this method

15 minutes

- Book is presented and students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read Can You Believe? Hurricanes by Sandra Markle
  - Group B students will read Wild Weather: Hurricanes! By Lorraine Jean Hopping
  - Group C students will read a book on an unrelated science topic
- While reading, students in Groups A and B will be encouraged to give the thumbs up sign when they find a target word in the text; students in group C will be asked to look for interesting words about their topic
Day Three- Hurricane Unit

Groups A, B, and C

Procedure:

5 minutes

- Teacher briefly reviews the target vocabulary and definitions from the chart on display

15 minutes

- Students work in dyads and partner read
- During the reading portion of the lesson:
  - Group A students will read *Can You Believe? Hurricanes* (by Sandra Markle)
  - Group B students will read *Hurricanes* (by Seymour Simon)
  - Group C students will read a book on an unrelated science topic

10 minutes

- Students work in dyads to write a paragraph using as many target words as they can
- The book they have just read and the vocabulary definition chart will be available as resources
In the spaces below, write a definition for each word. If you cannot think of the definition, use the word in a meaningful sentence.

mammals

gravity

crust

blubber

cyclone

phases

ash
craters
tropical
calf
baleen
equator
magma
blowhole
satellite
lava___________________________________________
______________________________________________
eye____________________________________________
______________________________________________
orbit__________________________________________
______________________________________________
erupt__________________________________________
______________________________________________
storm surge____________________________________
One of the scariest sights on earth is the sight of an active volcano. To see the force of a volcano lets you know how powerful really is! Magma is hot, melted rock. You can’t see the magma because it is under the . However, when that magma pushes up through the earth’s and starts to create a volcano, you can’t really what happens next. First, you may see a lot of steam and coming out of the volcano. That may be your idea to get out of the way. Sometime later lava will pour out of the volcano.

When magma comes up and out of a volcano it is called lava. It can form several different kinds of when it cools. An
eruption can last hours, days, weeks, or even years!

Scientists work hard to predict when volcanoes will erupt. As they get better at predicting eruptions, lava, science, people will have more time to get out of the way of these scary happenings.
We had such a great time at the zoo! We got to see a lot of different animals. My favorites were the whales because they were so cute and beautiful swimming around in their large tanks. We found out that even though whales look like fish, they are not. They are mammals. That means that they are born alive. They don’t come from eggs. The baby can weigh as much as 400 pounds when it is born.

That’s one big baby! Since whales are mammals they breathe air like we do. But a whale doesn’t have a heart like you or me. It breathes through a blowhole on the top of his head. To keep warm in the very cold water of the ocean, whales have baleen blankets. I wish I had some of that in the winter. Some whales have teeth, but others have baleen bladders that are used to strain little, tiny creatures that they like to eat. I’m glad I’m not that tiny!
The old astronaut told his amazing story to the reporter. He was one of the first men to walk on the moon! He talked about the craters that he and the other astronauts would see into when they were taking their walks. He also talked about how weird it was to be somewhere that did not have any rain, how they would float up and down like they were as light as marshmallows. He explained that the moon was a satellite of the earth and therefore it circled around and around the earth all year long. That circling around the earth means it orbits the earth. That’s why we on earth see the moon in different phases. That’s the different parts we can see of the moon as it makes its journey around the earth.

Even though the reporter didn’t understand all of the science, he was fascinated by the astronaut’s stories.
Hurricanes are gigantic, spinning storms. They develop in areas around the equator where the water is warm. People long ago used to think that an evil god made these storms because they were so destructive. Today we know that hurricanes develop in areas by gathering energy through contact with the ocean waters. Hurricanes are one of three kinds of cyclones and they are the strongest kind. Hurricanes have steady winds of more than 74 miles per hour. In the middle of a hurricane is a calm center that is called the eye. When the hurricane moves across the ocean and gets to the land there is a storm surge where a huge dome of comes up and over the coastline. This can be very dangerous as it can cause flooding to the towns near the coastline.
Samples of content area text used for the study:

“The moon appears to change shape, but it doesn’t. The different shapes we see are called phases of the moon. We are seeing different amounts of light being reflected on the moon. How much light we see depends on the positions of Earth, moon and sun.” (Gibbons, 1997)

“Baleen whales eat krill and small fish. Krill are tiny shrimp. Baleen whales eat by taking large mouthfuls of krill or fish. The animals become trapped in the whales’ baleen. Baleen whales are then able to swallow the krill and fish.” (Holmes, 1998)

“Hurricanes are huge spinning storms that develop in warm areas around the equator. Hurricanes bring strong winds, heavy rains, storm surges, flooding, and sometimes even tornadoes. Coastal areas and islands are in the most danger during a hurricane, but even inland areas are at risk.” (Simon, 2003)

“The mantle is below the crust of the earth. It is 1,800 miles (3,000 kilometers) wide and contains rock material and magma. Hot liquid magma also contains gas and steam.” (Merrians, 1996)