

ALZAHRA NEY, TURKI SAMAH, Ph.D. The Effects of Context Clues for Deriving Word Meanings in Fourth-Grade Students with Adequate and Poor Vocabulary Knowledge. (2019).

Directed by Dr. Pamela Williamson. 226 pp.

The process of acquiring new words to use in daily life is called *vocabulary development*. The aim of vocabulary development is helping students learn the meanings of new words and concepts in various contexts and across all academic content areas. Research shows that there is a variety of ways to learn vocabulary including direct instruction, incidental learning, and context clues. Researchers such as Carlo, August, McLaughlin, and Snow (2004), Herman and Dole (1988), and Martin-Chang, Levy, and O'Neil (2007) pointed out that even though students can sometimes learn a new word when the definition is given, there are other times when they need strategies for using context to decipher unfamiliar words. Using context clue strategies helps students to understand the reading and improve their achievement. Context clues are words or hints found around an unfamiliar word that provides clues that reveal the meaning of the unknown word (Beck & McKeown, 1991). Research has shown positive effects of teaching the use of context clues on students' word learning (e.g., Baumann et al., 2002). When students possess the adequate knowledge of using a word-learning strategy (i.e., context clues), this will help them to become independent readers as well as serve their continued success in their lives after secondary school.

However, lack of research on such strategies to improve word knowledge appears to be one of the critical obstacles to enhanced vocabulary development for students, particularly students with reading difficulties, including those with learning disabilities

(LD; Jitendra, Edwards, Sacks, & Jacobson, 2004; Stahl & Nagy, 2006). In fact, quite a few studies have been conducted to investigate the utility of the context clue strategies as a teaching device to improve word knowledge, vocabulary acquisition, and comprehension with elementary, typically achieving students (e.g., Martin-Chang, et al., 2007; Nash & Snowling, 2006; Yuen, 2009). While most of the investigators obtained results that pointed to the superiority of the context clues as a teaching strategy, few of them could report results that were statistically significant. Furthermore, the findings of some studies (e.g., Goerss, Beck, & McKeown, 1999) showed that an instructional task based on the process of using context to derive word meaning information is a powerful model for one-on-one instruction. The question remains as to how useful the instructional intervention would be if it were adapted for small groups and classrooms. Additionally, to date, the question of whether instruction in context clue strategies can improve the ability of students with reading difficulties, including those with LD, to use context to derive the meanings of unfamiliar words has not been explored and remains an open question, one addressed in this study.

The purpose of this quasi-experimental research study was to examine the influence of vocabulary instruction that is based on a combination of a strategy and certain types of context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate (AVK) and poor vocabulary knowledge (PVK). Specifically, this study involved a comparison of two approaches: (a) *business as usual* instruction was used as a control condition, and (b) a nine-day vocabulary instructional intervention was used as a treatment condition. The

dependent variable in the study was a measure of the effects of the vocabulary instructional intervention (context clues strategy) on students' vocabulary acquisition. Two measurement instruments were used to measure the dependent variable in this study: (a) the Gates-MacGinitie Reading Tests (GMRT), Fourth Edition, and (b) an experimenter-constructed test (ECT) designed for this study. Five fourth-grade classrooms, with a total of 59 students, participated in the study. These classrooms came from four separate elementary schools, two public schools in the same school district and two private schools that exclusively serve students with diagnosed LD, located in three different urban cities in the southern United States.

Measures of vocabulary knowledge were administered to participants at three different moments in the study: (a) before providing the vocabulary intervention, (b) immediately after completing the intervention, and (c) three weeks later. Results revealed that after receiving the vocabulary instructional intervention, both groups of students (students with PVK and AVK) in the treatment condition significantly outperformed students in the control condition on both measures of vocabulary knowledge (GMRT and ECT). The changes in students' results on both measures of vocabulary knowledge across the two instructional conditions (treatment and control conditions) were consistent, which confirms the hypothesis that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

When examining the extent to which both groups of students in the treatment condition retained the learned context clue strategy three weeks post-intervention, the findings indicated that both groups of students performed significantly better in short-

term learning (on the immediate posttest) compared to their performance in long-term learning (on the delayed posttest). These findings provide evidence that the vocabulary instructional intervention was effective and suggest that teaching students how to use context clues while reading—even brief, direct, and explicit vocabulary interventions—improves their understanding and ability to derive the meanings of novel vocabulary words in new written contexts. Limitations, contributions, implications, and future directions are discussed.

THE EFFECTS OF CONTEXT CLUES FOR DERIVING WORD MEANINGS
IN FOURTH-GRADE STUDENTS WITH ADEQUATE AND
POOR VOCABULARY KNOWLEDGE

by

Turki Samah Alzahrany

A Dissertation Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Greensboro
2019

Approved by

Committee Chair

© 2019 Turki Samah Alzahrany

DEDICATION

*I dedicate this work to my parents, my wife, and my children,
without whom my educational and personal achievements
would not have been possible.*

APPROVAL PAGE

This dissertation written by Turki Samah Alzahraneey has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair _____
Pamela Williamson

Committee Members _____
Devdass Sunnassee

Stephanie Kurtts

Alan Kamhi

Date of Acceptance by Committee

Date of Final Oral Examination

ACKNOWLEDGMENTS

This dissertation journey has not only been a wonderful learning experience and academic writing piece, but also a process that has shown me how important it is to believe in myself, to continue to strive to be a better writer and researcher, to never give up, and to surround myself with wonderful people who support and care for me unconditionally. I wish to acknowledge my appreciation to several individuals for their personal and professional support throughout my doctoral program.

I would like to begin by expressing my sincere thanks to my committee chair and advisor, Dr. Pamela Williamson, for her continual guidance and endless answers to questions throughout my doctoral program. Her words of encouragement, as well as her expertise and sound advice are greatly appreciated. Without her unwavering support and incredible expertise, this project would not have been feasible. I am thankful that she has been part of this project from its inception.

I would also like to thank the rest of my wonderful and dedicated dissertation committee members, Drs. Stephanie Kurttz and Devdass Sunnassee and Prof. Alan Kamhi for all of their support and assistance through this project despite their own busy schedules and research. I also appreciate all the time they spent providing insightful and constructive feedback on my work throughout my doctoral program. A special thanks goes to Dr. Sunnassee, who has encouraged me and provided ongoing support any time I asked for it as I navigated my way through statistics. I am also grateful to Drs. Terry Ackerman and Melody Zoch, former dissertation committee members. It is a great honor

to have each of those respected professionals on my doctoral committee and a testament to the power of encouragement.

I am grateful to the fourth-grade teacher, Mrs. Tammy Launer, who welcomed me in her classroom and graciously agreed to participate in my pilot study. I would also like to thank the 21 fourth graders in her classroom who participated in the pilot study, and their parents. I am incredibly grateful for their willingness and support in shaping this project.

Many thanks also go out to the five awesome teachers who endured my presence for the last four months. The assistant superintendent, Ms. Catherine Berry, and the amazing private- and public-school principals welcomed me and were enthusiastic about my presence on their campuses. They were all outstanding. I obviously could not have finished this project without them, and of course, my thanks goes to the terrific fourth graders for their hard work and their parents who allowed me to spend time with their children as part of my present study. Thanks to all of them for helping me and for the role they played in making this possible.

Nobody has been more important to me in the pursuit of this project than the members of my family. I would like to thank my parents for their prayers, love, and endless support. Their love, wise counsel, and guidance are with me in whatever I pursue. Most importantly, I wish to express my unqualified thanks to my loving and supportive wife, Samiyah, and my two wonderful children, Albaraa and Lana, who provide unending inspiration. I also wish to thank my brothers and sister for supporting me spiritually throughout writing this dissertation and my life in general.

Last, but certainly not least, thanks to the University of Jeddah and the government of my country, the Kingdom of Saudi Arabia. I could not have gone through the doctoral program overseas without their financial support. I would like to thank everyone who helped to make this project possible and bring it to a successful conclusion. Without all of the support, I would never have gotten this far.

TABLE OF CONTENTS

	Page
LIST OF TABLES	xi
LIST OF FIGURES	xiii
CHAPTER	
I. INTRODUCTION	1
Context Clues.....	3
The Fertility/Futility Debate: A Conceptual Framework for the Study	5
Statement of the Problem.....	7
Significance of the Study	11
Research Purpose and Questions	13
Subquestion 1A.....	14
Subquestion 1B.....	14
Subquestion 1C.....	14
Subquestion 1D.....	15
Subquestion 2A.....	15
Subquestion 2B.....	15
Subquestion 2C.....	16
Definitions of Key Terms	16
Summary.....	18
II. REVIEW OF RELEVANT LITERATURE	21
Five Components of Reading.....	21
Learning Disabilities	23
Prevalence	25
Characteristics.....	26
Vocabulary Knowledge and Development	28
Importance of Vocabulary Knowledge and Development.....	29
Learning-to-read phase	30
Reading-to-learn phase	30
Reading Comprehension and Vocabulary Connections	31
Importance of Vocabulary Development for Students with Learning Disabilities.....	34
Factors that Influence the Vocabulary Development.....	36
Vocabulary Instruction.....	41
What is Vocabulary Instruction?	42

What Effect does Vocabulary Instruction have on Student Reading Achievement?	43
What Makes an Effective Vocabulary Instructor?	44
What are Effective Methods of Implementing Vocabulary Instruction?	46
Vocabulary instruction for students with and without learning disabilities.....	47
Types of vocabulary interventions.....	48
Non-generative vocabulary interventions	49
Generative vocabulary interventions	51
Contextual analysis strategies	52
Review of the context clues literature.....	54
Meta-analysis studies on teaching context clues	56
Experimental studies on teaching context clues	59
Context clue instruction	59
General strategy instruction	64
Combination of strategy and context clues.....	73
Summary	76
III. METHODOLOGY	78
Research Questions and Hypotheses	78
Subquestion 1A.....	79
Subquestion 1B.....	79
Subquestion 1C.....	79
Subquestion 1D.....	79
Subquestion 2A.....	80
Subquestion 2B.....	80
Subquestion 2C.....	80
Research Designs and Data Analysis.....	81
Participants and Setting.....	84
Data Collection and Study Procedures	86
Procedures for Recruiting Participants	86
Costs and payments to the participants.....	87
Potential risks to participants	88
Procedures for maintaining confidentiality of the data collected.	88

Procedures for protecting the privacy of potential subjects during recruitment.....	89
Teacher intervention training.....	90
Testing time	91
Study Procedures	91
Step 1: Vocabulary pretests	91
Step 2: Instructional conditions	92
Step 3: Immediate vocabulary posttests.....	94
Step 4: Delayed posttests	94
Instrumentation and Other Documents	95
Instructional intervention	95
Dependent measures	99
The standardized test.....	99
Validity and reliability of the Gates- MacGinitie Reading Tests.....	102
How and why the Gates-MacGinitie Reading Test was used for this study?	103
The experimenter-constructed test.....	105
Validity and reliability of the experimenter- constructed test.....	106
Multiple-choice recognition instrument.....	106
Meaning-recall instrument.....	110
Fidelity of implementation.....	113
Inter-scorer reliability	115
IV. RESULTS	117
Research Question 1	123
Subquestion 1A.....	123
Subquestion 1B.....	127
Subquestion 1C.....	130
Subquestion 1D.....	133
Research Question 2	138
Subquestion 2A.....	138
Subquestion 2B.....	141
Subquestion 2C.....	145
Research Question 3	149
Interobserver Agreement	156
Inter-Scorer Reliability	157
V. DISCUSSION.....	158
Overview of the Results.....	161
Research Question 1	161

Subquestion 1A.....	162
Subquestion 1B.....	162
Subquestion 1C.....	162
Subquestion 1D.....	162
Research Question 2.....	165
Subquestion 2A.....	165
Subquestion 2B.....	165
Subquestion 2C.....	166
Research Question 3.....	167
Uniqueness of the Present Study.....	168
Interpretation of the Findings.....	172
Contribution to the Broader Research.....	174
Limitations and Strengths of the Present Study.....	177
Recommendations and Implications for Future Research.....	180
Conclusion.....	189
 REFERENCES.....	 191
 APPENDIX A. OVERVIEW OF THE PILOT STUDY RESULTS, EFFECT SIZE, AND REQUIRED SAMPLE SIZE.....	 216
 APPENDIX B. AN EXAMPLE OF THE LESSON PLANS USED IN THE VOCABULARY INSTRUCTIONAL INTERVENTION.....	 220
 APPENDIX C. EXAMPLES OF THE QUESTION ITEMS INCLUDED IN BOTH EXPERIMENTER-CONSTRUCTED INSTRUMENTS.....	 222
 APPENDIX D. TEACHER FIDELITY CHECKLIST.....	 224

LIST OF TABLES

	Page
Table 1.1 Seven Common Types of Context Clues and Related Signal Words.....	4
Table 3.1 The Three-Way Mixed ANOVA Design Table.....	82
Table 3.2 The Pretest-Posttest, Nonequivalent Group Design Table	83
Table 3.3 Test Forms and Times.....	100
Table 4.1 Participants Table	118
Table 4.2 Results of the Two-Way ANOVA for the Effects of Condition and Group at the ESS Pretest Scores.....	120
Table 4.3 The Unweighted Marginal Means of the ESS Pretest Scores for Condition and Group.....	121
Table 4.4 Pairwise Comparisons for the ECT Scores by Condition x Time	126
Table 4.5 Pairwise Comparisons for the ECT Immediate and Delayed Posttest Scores by Condition for the PVK Group	128
Table 4.6 Descriptive Statistics for the ECT Scores by Condition x Group at the Immediate and Delayed Posttests	130
Table 4.7 Pairwise Comparisons for the ECT Immediate and Delayed Posttest Scores by Condition for the AVK Group	132
Table 4.8 Pairwise Comparisons for the ECT Scores from Immediate Posttest to Delayed Posttest.....	135
Table 4.9 Descriptive Statistics for the ECT Scores by Group for the Immediate and Delayed Posttests	136
Table 4.10 Results of the Univariate Tests for the GMRT Scores by Condition at the Pretest and Delayed Posttest	140
Table 4.11 Pairwise Comparisons for the GMRT Scores by Condition at the Pretests and Delayed Posttests	141

Table 4.12 Pairwise Comparisons for the GMRT Pretest and Delayed Posttest Scores by Condition for the PVK Group.....	143
Table 4.13 Descriptive Statistics for the GMRT Scores by Condition x Group at the Pretest and Delayed Posttests.....	144
Table 4.14 Pairwise Comparisons for the GMRT Pretest and Delayed Posttest Scores by Condition for the AVK Group.....	146
Table 4.15 Results of the One-Way ANOVA for the Effects of Condition on the Normalized GMRT Pretest Scores	151
Table 4.16 Results of the One-Way ANOVA for the Effects of Condition on the Normalized ECT Pretest Scores	152
Table 4.17 Results of the One-Way ANOVA for the Effects of Condition on the Normalized GMRT Delayed Posttest Scores	153
Table 4.18 Results of the One-Way ANOVA for the Effects of Condition on the Normalized ECT Delayed Posttest Scores	154

LIST OF FIGURES

	Page
Figure 3.1 Instructional Conditions	93
Figure 4.1 Results of the Means of ESS Pretest Scores for Condition and Group	122
Figure 4.2 Changes in the Students' Retention as a Function of Time and Group.....	137
Figure 4.3 Results of the Means of GMRT Pretest and Delayed Posttest Scores for Group and Condition.....	148
Figure 4.4 Changes in the Students' Results as a Function of Test and Condition	155
Figure A.1 The Mean Percentages of the Experimenter-Constructed Tests for All Students	217
Figure A.2 The Mean Percentages of the Standardized Tests for All Students.....	218

CHAPTER I

INTRODUCTION

The ability to read is critical to a student's future success because it is the foundation for learning. In schools today, many students experience difficulty with reading. Thirty-two percent of fourth-grade students and 24% of eighth-grade students in the United States read below the basic level, leaving these students unable to achieve even minimum academic expectations (National Assessment of Educational Progress [NAEP], 2017). This means that more than eight million students in Grades 4 through 12 are struggling to read at grade level (NAEP, 2017). These struggling readers, if not provided with early and appropriate interventions, will face serious challenges, including an increased risk of dropping out of school, which occurs in the United States at a rate of more than 600,000 students per year (Stillwell, Sable, & Plotts, 2011). According to Hudson, Isakson, Richman, Lane, and Arriaza-Allen (2011), students who experience early reading difficulty often continue to experience decline in later grades and further in life, emphasizing the importance of providing early intervention.

There is an even greater reason for concern when considering the reading scores of students with disabilities. In 2017, 72% of Grade 12 students and nearly 70% of students with disabilities in fourth and eighth grades were reading below their grade levels (National Center for Education Statistics [NCES], 2018b). Lyon and Chhabra (2004) stated that failure to read by age 9 predicts a lifetime of illiteracy for at least 70%

of struggling readers. This is a powerful statement. Statistics such as these are what drive teachers to become better literacy instructors. Although the overall percentage of students reading below the proficient level has decreased over the past decade, this improvement has not been as pronounced among students with disabilities (Lee, Grigg, & Donahue, 2007). In fact, there are still far too many students not reading at an acceptable level of proficiency. Therefore, educators need practical and proven methods to help students with and without disabilities improve their reading skills and make better progress on reading assessments.

Reading researchers have concluded that a way to strengthen reading skills is to strengthen vocabulary (Gunning, 2013; Herman & Dole, 1988; Nash & Snowling, 2006; Reed, Petscher, & Foorman, 2016). Nash and Snowling (2006) emphasized that “vocabulary, the knowledge of words and their meanings, is one of the best predictors of educational achievement” (p. 336). Other researchers have also confirmed the important connection between vocabulary knowledge and reading comprehension by showing that greater vocabulary knowledge makes comprehension easier (Kame’enui & Baumann, 2012; Lesaux & Kieffer, 2010; Oakhill, Cain, & Elbro, 2015; Reed et al., 2016).

Generally, there are two basic types of vocabulary instruction: (a) intentional and (b) incidental. Intentional vocabulary instruction is defined as instruction with the explicit purpose to teach the meaning of a word (Beck & McKeown, 1991; Harris, Deshler, & Schumaker, 2011). An example of intentional instruction is when a teacher directly provides one or more resources, such as a dictionary or a more knowledgeable person, with the strict intention of having the student learn the meaning of single a word.

Researchers such as Beck and McKeown (1991), Fukkink (2002), and Lenhart, Lenhard, Vaahtoranta, and Suggate (2018) describe incidental vocabulary instruction as an experience where students may increase their word knowledge through an initial encounter with a word. This encounter may come through an oral situation, such as conversation and the media, or through written environments, such as letters, magazines, and books.

Context Clues

The most prominent way students learn words incidentally is through the use of context clues (Beck & McKeown, 1991; Beck, McKeown, & Kucan, 2002; Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998; Nash & Snowling, 2006). Context clues are words or hints found around an unfamiliar word that provide clues to reveal the meaning of the unknown word (Beck & McKeown, 1991). The context in which a word is used can often provide clues that can help students pinpoint a word's meaning independent of a dictionary or a teacher. Using context is one strategy students can adopt to help them become independent word learners, and it also helps account for the words students learn outside of intentional instruction.

There are seven common types of context clues, and these context clues often contain signal words that readers can use to help them learn the meaning of new words. Thus, signal words point out the type of context clue being used. Research has shown that teachers need to provide struggling readers, including those with learning disabilities (LD), with direct instruction in how to use these different types of context clues as well

as their respective signal words (Nash & Snowling, 2006). Table 1.1 includes an explanation of each type of context clue, its related common signal words, and examples.

In context clue studies, instruction and practice center on one or more of these context clue types (Fukkink & de Glopper, 1998). Students are taught to recognize and use certain context clues to elicit the meaning of an unfamiliar word. In some studies, clue instruction is incorporated in a generic strategy that emphasizes the recognition of the instructed clues. This type of instruction is closely aligned to studies in which a classification of context clues is proposed and can, therefore, be labeled as text-oriented.

Table 1.1 Seven Common Types of Context Clues and Related Signal Words

Context Clue	Definition/Explanation	Common Signal Words	Example
Definition / Restatement	The unfamiliar word is defined in a dictionary form or casual form.	<i>are, is, or, that is, also known as, which is to say, sometimes called, meaning, is defined as, means</i> Note: Commas, parentheses and dashes that set a word or phrase apart from the rest of the sentence could also be a signal.	A vegetarian <i>is</i> a <u>person that does not eat, or believe in eating, any food derived from animals.</u>
Synonym	A word is similar in meaning to the unfamiliar word.	<i>likewise, like, especially, or, in that, similarly, in other words, that is, and</i> Note: Commas, parentheses and dashes that set a word or phrase apart from the rest of the sentence could also be a signal.	His simple glance was a harbinger of danger. <i>That is</i> , his eyes were <u>signs foretelling</u> rough waters in their relationship.

Antonym	A word is opposite in meaning to the unfamiliar word.	<i>different, unlike, though, opposite, by contrast, but, some...but others, on the other hand, not, despite, although, yet, then again, whereas</i>	Now they were <u>confident</u> , despite their usual diffidence . Diffidence = Lacking confidence
Example / List	Examples of the unfamiliar word are given through what the word or phrase might be like, not what they have in common.	<i>such as, for example, including, for instance, in one case</i>	Some animals are omnivores . Bears, for instance , <u>eat whatever meat or plants they can find</u> .
Cause & Effect	Words describing an action in which a cause or effect suggests the meaning of an unfamiliar word.	<i>because, so, therefore, consequently, since, as a result, if... then, due to</i>	Because the horse was so fatigued , he <u>collapsed before he finished</u> . Fatigued = tired
Compare & Contrast	Comparing determines how two words are alike. Contrasting determines how two words are different.	<i>especially, like, likewise, also, and, as well as, that resembling, identical, similar to, as, in the same way</i> <i>unlike, on the other hand, in contrast to, on the contrary</i>	My brother is enthralled by birds in the same way that I am <u>fascinated</u> by insects.
Description / Inference	The use of background knowledge, logical guess, and context clues to determine meaning.	n/a	<u>When Joe's dog passed away</u> , he was filled with grief .

The Fertility/Futility Debate: A Conceptual Framework for the Study

This study is grounded in the Fertility/Futility Debate about reading vocabulary growth and instruction. There has been an ongoing argument regarding reading-vocabulary growth and instruction. One point of view is that, given the unreliability of context clues (Beck, McKeown, & McCaslin, 1983; Graves, 2006; Schatz & Baldwin, 1986), specific vocabulary can be taught in effective, *fertile* ways through direct

instruction of specific words (Beck, McKeown, & Omanson, 1984). Indeed, numerous studies indicate that students can be effectively taught the meanings of specific new words through a variety of instructional strategies (Elleman, Steacy, Olinghouse, & Compton, 2017; see also reviews by Anderson & Nagy, 1991; Baumann & Kame'enui, 1991; Beck & McKeown, 1991; Blachowicz & Fisher, 2000; Graves, 1986; Jitendra, Edwards, Sacks, & Jacobson, 2004; Miller & Gildea, 1987; Stahl & Fairbanks, 1986).

The counterargument is that it is *futile* to attempt to teach words individually because of the vast number of words students must learn and limited instructional time (Ford-Connors & Paratore, 2015; Graves, 2016; Nagy & Herman, 1987). Instead, students' growth in vocabulary can be best accounted for by independent reading (Harris, Schumaker, & Deshler, 2011; Karbalaei, Amoli, & Tavakoli, 2012; Nagy, Anderson, & Herman, 1987), listening to stories read aloud (Hennebry, Rogers, Macaro, & Murphy, 2017; Lenhart et al., 2018; Sénéchal & Cornell, 1993; Teng, 2016), and exposure to enriched oral language (Dickinson, Cote, & Smith, 1993; Kame'enui & Baumann, 2012; Kamil et al., 2008).

Missing from the fertility/futility debate, however, is the acknowledgment that vocabulary growth can occur through the application of generalizable linguistic knowledge in the form of morphemic and contextual analysis. *Morphemic analysis* involves unlocking a word's meaning by examining its morphemes, or meaningful parts, such as base words, prefixes and suffixes, inflected endings, and Latin or Greek roots (Kieffer & Lesaux, 2012). *Contextual analysis* involves inferring a word's meaning by examining surrounding text, which includes syntactic and semantic linguistic cues

provided by preceding and succeeding words, phrases, and sentences (Helman, Calhoon, & Kern, 2015). Nagy and Scott (2000) acknowledged the prevalence of linguistic cues by noting that “context and morphology (word parts) are the two major sources of information immediately available to a reader who comes across a new word” (p. 275). Although morphemic or contextual analysis may not be as effective for vocabulary learning when compared to direct instruction in the meanings of *specific* words (Baumann & Kame’enui, 1991; Cobb & Blachowicz, 2014; Jenkins, Matlock, & Slocum, 1989), instruction in morphemic and contextual analysis has the potential to equip a learner with the ability to infer the meanings of *numerous* words in an independent manner.

Statement of the Problem

There is growing evidence suggesting that a reader must employ certain cognitive processes (or strategies) in order to most efficiently make use of contextual information (Çakici, 2017; see also Sternberg & Powell, 1983). These processes involve several components, such as planning, monitoring, and decision-making as well as execution of strategic behaviors, such as selecting and integrating information. In other words, when readers encounter new words, they must decide what information will determine word meanings and what information will receive the most attention, and then integrate this with previous contextual information and prior knowledge. At the same time, the learners must also test and make adjustments in hypothesized word meanings based upon subsequent contexts.

Contextual knowledge is one type of simple-level word knowledge. Contextual knowledge is a word meaning derived from context, which can include a sentence, a passage, a discussion, or a picture (Lewis, 2009). According to Rupley and Nichols (2005), “Contextual knowledge often has a stronger connection to the text than definitional knowledge” (p. 242). The research suggests that students whose vocabularies are most in need of support are less likely to be able to get information from context (Beck et al., 1983; Graves, 2006; Schatz & Baldwin, 1986). Additionally, several studies confirm that deriving word meaning from context is a complex process that is susceptible to errors at several points (Baumann, Edwards, Boland, Olejnik, & Kame’enui, 2003; Fukkink, 2002, 2005; Fukkink & de Glopper, 1998). These findings emphasize that teaching students how to use context to derive word meaning is quite important. Contextual analysis is not meant to teach specific words. Instead, it is meant to be a general strategy aimed at helping students to contend with unfamiliar words in a wide variety of texts (Kuhn & Stahl, 1998).

The purpose of teaching strategies to improve the learning of word meanings from context is to help students learn more words incidentally as they are encountered in everyday reading. Doing so should, in turn, lead to a larger vocabulary over time as students read texts containing unknown words (Kuhn & Stahl, 1998). A few such studies have been undertaken, but researchers have not met with overwhelming success (e.g., Fukkink, 2002; Patberg & Stibbe, 1985). The findings of some studies (e.g., Goerss, Beck, & McKeown, 1999) showed that an instructional task based on the process of using context to derive word meaning information is a powerful model for one-on-one

instruction; yet, it is unclear how useful the procedure would be if it were adapted for small groups and classrooms.

Providing explicit vocabulary instructions in various cognitive and metacognitive strategies to help students, especially students with LD, determine word meanings is an important factor in vocabulary development (Gersten, Fuchs, Williams, & Baker, 2001; Swanson, Mink, & Bocian, 1999; Torgesen et al., 2007; Vaughn, Gersten, & Chard, 2000; Wong, Harris, Graham, & Butler, 2003). Although successful readers can increase vocabulary by encountering unknown words during incidental reading (Elleman et al., 2017; Sternberg, 1987), this is problematic for students with LD since the possibility of learning any word during a first encounter is low, especially given challenging texts (Fukkink, 2002; Jitendra et al., 2004; Weiser, 2013). Even if students with LD can identify the printed words, they likely lack the vocabulary necessary to understand grade-level texts (Biemiller, 1999). Additionally, the findings of literature review research on vocabulary instruction showed that students with LD often had minimal to no training in deriving meanings for unfamiliar words using context (Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998).

Since students with reading difficulties, especially students with LD, are not equipped with effective word learning strategies, they often have fragmented knowledge of words, particularly a narrower understanding of word features (Jitendra et al., 2004; Swanson et al., 1999; Swanson & Vaughn, 2010). Hence, the best solution may involve providing explicit vocabulary instruction in various cognitive and metacognitive strategies to assist these students in finding the meaning of unknown words

independently during reading. Explicit instruction can be described as instruction that (a) provides a clear description of the task, (b) encourages students to pay attention, (c) activates prior knowledge, (d) breaks the task into small steps, (e) provides adequate practice throughout each step, and (f) provides teacher feedback (RAND Reading Study Group, 2004). According to Karbalaie et al., (2012), if students are explicitly taught how to use context as a vocabulary learning strategy, their ability to learn words independently may increase.

However, lack of research on such strategies to improve word knowledge appears to be one of the critical obstacles to enhanced vocabulary development for students, particularly students with reading difficulties including those with LD (Jitendra et al., 2004; Stahl & Nagy, 2006). In fact, quite a few studies have been conducted to investigate the utility of the context clue strategies as a teaching device to improve word knowledge, vocabulary acquisition, and comprehension with elementary, typically achieving students (e.g., Martin-Chang, Levy, & O'Neil, 2007; Nash & Snowling, 2006; Yuen, 2009). While most of the investigators obtained results that pointed to the superiority of the context clues as a teaching strategy, few of them could report results that were statistically significant. To date, the question of whether instruction in context clue strategies can improve the ability of students with reading difficulties, including students with LD, to use context to derive the meanings of unfamiliar words has not been explored and remains an open question, one addressed in this study.

Significance of the Study

The context clue refers to a source of information that helps readers understand the unfamiliar word. A context clue strategy, which provides students with specific steps to determine meaning of unknown words, should be explicitly taught by teachers to enable students to use context clues (e.g., synonyms, antonyms, definition, example; Cobb & Blachowics, 2014). Research has shown that average 12th graders know something like 50,000 word families and learn from 3,000 to 4,000 words each year (Anderson & Nagy, 1992; Anglin, 1993; Graves, 1986, 2016; White, Graves, & Slater, 1990). Hence, students need to learn six to eight new words per day (Beck, McKeown, & Kucan, 2013; Cain 2007; Justice, Meier, & Walpole, 2005; Stahl & Shiel, 1999). Since it is impossible to teach every single word that students need to know, learning to use the context clues strategy will help them to independently know a larger number of vocabulary words.

When students do not possess the adequate knowledge of using a word-learning strategy (i.e., context clues), they may face several long-term complications. First, this may impact student achievement in negative ways. The *National Reading Panel* report by the National Institute of Child Health and Human Development (2000) insists that the level of vocabulary impacts reading comprehension; students are unable to understand their reading clearly if they have limited vocabulary, which prevents students from being independent readers. Accordingly, their success at school and readiness for college and career may not be guaranteed.

Second, the inadequacy of knowledge in the use of context clues for deriving meanings of unknown words can lead students to not learn the specific word-learning strategy they need to know in order to help them read and understand the meanings of new words in a text. In fact, using context clues aligns with one of the English Language Arts (ELA) Common Core State Standards (i.e., Common Core State Standards Initiative, n.d.). This standard requires students to “determine or clarify the meaning of unknown and multiple-meaning words and phrases by using context clues, analyzing meaningful word parts, and consulting general and specialized reference materials, as appropriate” (n.d., para. 5). Consequently, lacking the skill may represent an obstacle to students’ capability to achieve the standard.

Additionally, most vocabulary is gained through reading (Duff, Tomblin, & Catts, 2015). Therefore, it is important that students recognize and take advantage of context clues (Sternberg, 1987). If, on the one hand, students keep trying to skip over unfamiliar words, they may come to the end of the passage with a very unclear idea of what they have read. On the other hand, if students stop to look up every unknown word, they may become irritated at the slow rate of their reading and stop altogether. In other words, using the context clues strategy will help students accelerate their reading speed and enhance comprehension. Additionally, this strategy helps students to learn how to reference a dictionary correctly because they will select the right definition from the multiple meanings listed in the entry.

Mastering this strategy serves the long-term development of student achievement. As research has suggested, using the context clues strategy provides positive effects on

students' word learning (e.g., Baumann et al., 2002; Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998; Martin-Chang et al., 2007). For example, students will gradually develop more extensive reading vocabulary as they become more proficient in using context clues in the text. It will also help students read faster and enhance reading comprehension as their knowledge of words grows (Kuhn & Stahl, 1998). As a result, students will be independent readers and be ready for both college and careers.

Research Purpose and Questions

The purpose of this study was to examine the influence of vocabulary instruction that is based on a combination of a strategy and certain types of context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. In other words, the study determined the extent to which an instructional intervention—modeling the process of word-meaning acquisition and guiding students through the process—may overcome problems that fourth-grade students with and without reading difficulties had in deriving word meaning from context. Thus, the results of this study may apply to the school setting, as it helps teachers provide reading instruction that might facilitate understanding the process of deriving word meaning from contexts.

I used a well-recognized standardized reading test (Gates-MacGinitie Reading Tests [GMRT], Fourth Edition) and an experimenter-constructed test designed for this study to answer the following research questions:

1. Does the vocabulary instructional intervention have any effect on the performance of the students on the experimenter-constructed test? If so, for which group of

students (students with poor vocabulary knowledge [PVK] and students with adequate vocabulary knowledge [AVK]) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (1A–1D):

Subquestion 1A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) over time? I hypothesize that students in the treatment condition will perform better than students in the control condition on the experimenter-constructed test.

Subquestion 1B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the experimenter-constructed test.

Subquestion 1C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the experimenter-constructed test.

Subquestion 1D

Is there a difference in retention between short-term learning and long-term learning (Time 2 vs. Time 3) for students in the treatment condition? I hypothesize that the performance of the students will be better in short-term learning (Time 2) in the treatment condition on the experimenter-constructed test.

2. Does the vocabulary instructional intervention have any effect on the performance of the students on the standardized test (Gates-MacGinitie Reading Tests [GMRT])? If so, for which group of students (students with PVK and students with AVK) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (2A–2C):

Subquestion 2A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) at the pretest and delayed posttest? I hypothesize that students in the treatment condition will perform better than students in the control condition on the standardized test.

Subquestion 2B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the standardized test.

Subquestion 2C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the standardized test.

3. Are the students' results on both measures of vocabulary knowledge (the GMRT and the ECT) consistent across the two conditions (treatment condition and control condition)? I hypothesize that students' results on both measures of vocabulary knowledge across the two conditions will be consistent, which will confirm that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

Definitions of Key Terms

In order to avoid misunderstandings or unclear language, in this section, I define key terms relevant to the research study. Chapter two covers most of these terms in-depth.

- **Context.** Context is “the language that surrounds a given word or phrase” (Moats, 2005, p. 66).
- **Explicit.** Explicit is a type of lesson delivery: direct, precise, and unambiguous.
- **Gates-MacGinitie Reading Tests (GMRT).** GMRT are group-administered paper and pencil tests to measure reading achievement (Maria & Hughes, 2008).

- **Individuals with Disabilities Education Act (IDEA).** IDEA is the federal special education law concerning schools and the education of students with disabilities.
- **Instructional Strategies.** Instructional strategies are the direct teaching or planned experiences that facilitate a growth in knowledge and/or understanding (Woolfolk, 2010).
- **Learning Disabilities (LD).** Learning disabilities are disorders that affect the ability of the individual to understand or use spoken or written language. LD can be identified by difficulties with listening, thinking, speaking, reading, writing, spelling, or mathematical calculations (Wright & Wright, 2009).
- **Literacy.** Literacy is a complex set of skills that comprise the interrelated processes of reading and writing required within varied socio-cultural contexts (National Joint Committee on Learning Disabilities [NJCLD], 2008).
- **No Child Left Behind Act (NCLB).** NCLB was originally passed as a bill by the U.S. Congress in 2001 but was enacted by the U.S. federal government in 2002. NCLB mandates the use of research-based strategies in the classroom (NCLB, 2002)
- **Reading Comprehension.** Reading comprehension is the process of gaining an understanding of written text through a process of extracting and constructing meaning (Spencer, Quinn, & Wagner, 2014).
- **Reading Difficulties.** This term can be used with students with serious word reading difficulties but who have not been formally identified as having reading

disabilities (Denton & Al Otaiba, 2011). These students may just slightly lag behind their peers, requiring more time to learn certain things. They may require more specialized reading instruction than has been provided, or the students may have previously received poor reading instruction.

- **Reading Disabilities (RD).** Reading disabilities are specific deficits in reading ability within the broader term of learning disability.
- **Research-Based Strategies.** Research-based strategies are techniques for instruction that have been empirically tested, researched, and proven effective.
- **Struggling Reader.** The struggling reader is one who may or may not possess the ability to decode words quickly and accurately and fails to comprehend meaning from text. For example, students who struggle with fluency are typically ones who can decode words and meaning, but at such a slow rate that comprehension is lost, causing them to become frustrated and lose interest in reading. Without intervention, these students will continue to fall behind their peers.

Summary

The process of acquiring new words to use in daily life is called *vocabulary development*. The aim of vocabulary development is helping students learn the meanings of new words and concepts in various contexts and across all academic content areas. Research shows that there is a variety of ways to learn vocabulary including direct instruction, incidental learning, and context clues. Researchers such as Carlo, August, McLaughlin, and Snow (2004), Herman and Dole (1988), and Martin-Chang et al. (2007) pointed out that even though students can sometimes learn a new word when the

definition is given, there are other times when they need strategies for using context to decipher unfamiliar words. Using context clue strategies helps students to understand the reading and improve their achievement. Also, research has shown positive effects of teaching the use of context clues on students' word learning (e.g., Baumann et al., 2002). When students possess the adequate knowledge of using a word-learning strategy (i.e., context clues), this will help them to become independent readers as well as serve their continued success in their lives after secondary school.

The aim of this quasi-experimental research study was to investigate effects of using the context clues strategy to provide vocabulary instruction to fourth-grade students with adequate and poor vocabulary knowledge. In this chapter, I introduced the issues of a limited vocabulary and the multifaceted problems that face students with limited knowledge in the use of context clues for deriving meanings of unknown words. These concerns established a foundation for this quantitative dissertation. I provided the research questions used to guide this study as well as the significance of the study. I also presented the theoretical framework in this chapter. In chapter two, I provide a literature review of the five essential components of reading instruction, issues related to the topic of vocabulary instruction for students with and without LD, and finally, the effects of using the context clues to improve students' ability to use context to derive the meanings of unfamiliar words. In chapter three, I offer an in-depth description of the research design and methodology employed in this present study, the participants, instrumentation that were used, administration, and data analyses. In chapter four, I present the research findings that emerged from this study, through numerical and narrative description.

Lastly, in chapter five, I discuss the study's findings and conclusions in detail, describe the limitations of the study, and provide recommendations and implications for future research.

CHAPTER II

REVIEW OF RELEVANT LITERATURE

As indicated in chapter one, this study examined the influence of vocabulary instruction based on context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. In this chapter, I provide a review of the literature that is relevant to this study, beginning with a description of the five major components of reading instruction that are essential to both reading success and school achievement. Second, I review literature about students with LD, focusing on the definition of the term learning disability (LD), statistics about the number of students with LD, and their general characteristics—particularly their characteristics in reading. This is followed by an explanation of the importance of vocabulary knowledge and development for students, including students with LD. In the fourth section of this chapter, I explore and investigate issues related to the topic of vocabulary instruction for students with and without LD. Lastly, I review and critique research examining the effectiveness of using context clues to improve students' ability to use context to infer the meanings of unfamiliar words, before closing the chapter with a chapter summary.

Five Components of Reading

The National Institute of Child Health and Human Development (NICHD) formed the National Reading Panel (NRP) to address the issue of teaching children to read (NRP,

2000). This panel sought to identify the most effective evidence-based methods for teaching reading. After soliciting public input and reviewing more than 100,000 research studies conducted on reading instruction, the panel published a report, *Teaching Children to Read: An Evidence-Based Assessment of the Scientific Research Literature on Reading and its Implications for Reading Instruction* in 2000. The University of Oregon's Center on Teaching and Learning (n.d.) published its interpretation of this report as a website resource called, "Five Big Ideas in Reading." The resource includes a list of the five core components of reading instruction that are essential to both reading success and school achievement: (a) phonemic awareness, (b) alphabetic principle (phonics), (c) fluency, (d) vocabulary, and (e) comprehension.

In order to be successful readers, students need to be proficient in these five areas. Phonemic awareness is an overarching term that relates to the awareness and understanding that spoken words are made up of smaller sounds (e.g., "cat" is made up of three separate phonemes: /k/ /a/ /t/). Mastering this skill allows students to recognize that written language is similar to oral language (López & Greenfield, 2004). Phonics, the second component, describes the relationship between graphemes (letter symbols) and their corresponding sounds (Joseph & Schisler, 2009). Fluency is the ability to read a text orally with speed, accuracy, and comprehension (Kim, Wagner, & Foster, 2011). In order for children to be fluent readers, they need to grow their vocabulary. Vocabulary is defined as the knowledge of the meanings of words (Butler et al., 2010). Increasing the knowledge of word meanings leads to increased reading comprehension. Comprehension, the fifth component, is a student's skill at processing meaning from text through

construction, integration, and metacognition (Graves, Juel, Graves, & Dewitz, 2011). Denton and Al Otaiba (2011) stated that comprehension is “the ultimate goal of [the] reading process” (p. 3). Other researchers have also echoed this statement (see Boyle, 2008; Kuhn, 2005; Rupley, 2009; Tolman, 2005; Torgesen, 2002).

Addressing these five components provides students with comprehensive literacy instruction. The purpose of the literacy curriculum is to allow students to build connections (Graves et. al, 2011). As students build connections, they begin to develop a higher order of thinking necessary in society (Mainali, 2013). Rupley, Blair, and Nichols (2009) pointed out that, when it comes to mastering the reading process, these “five instructional tasks or content strands represent the major thrust of reading acquisition” (p. 135). They concluded, “These major instructional tasks are inseparable parts of one total instructional process” (p. 135). Some readers pick up these components easily, while other readers need extra support and practice to master these skills. According to Swanson and Vaughn (2010), these five components of reading are particularly significant for students with LD to understand.

Learning Disabilities

Although the designation of LD as disabilities occurred in U.S. federal legislation in 1968, and researchers have been intensely studying LD throughout this time, there still exists considerable controversy over what a learning disability actually is (Learning Disabilities Association of America [LDA], n.d.). The most commonly used definition for the school-age population is found in the federal special education law, the Individuals with Disabilities Education Act (IDEA, 2004). IDEA uses the term *specific*

learning disability (SLD). Friend (2017) pointed out that the federal definition of LD articulated in P.L.94-142 in 1975 has changed very little since then. According to IDEA:

Specific learning disability means a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, that may manifest itself in imperfect ability to listen, think, speak, read, write, spell, or to do mathematical calculation, including conditions such as perceptual disabilities, brain injury, minimal brain dysfunction, dyslexia and developmental aphasia. Specific learning disability does not include a learning problem that arises primarily as the result of visual, hearing, or motor disabilities; of mental retardation; of emotional disturbance; or of environmental, cultural, or economic disadvantage. (IDEA 20 U.S.C. Stat. 1401 [2004], 20 C.F.R. Stat. 300.8[c][10]; see also Friend, 2014, p. 127)

IDEA's (2004) definition, although still maintaining the core assumption of an underlying, intrinsic, psychological processing disorder, differs from previous definitions (e.g., United States Office of Education's definition of 1977) by reducing the reliance on discrepancy and exclusionary identification methods in favor of a more criterion-based emphasis on a failure to achieve. However, issues with identifying students with LD have been an increasing concern for parents, teachers, assessment specialists, and researchers over the past decades (Alzahrane, 2019; Fletcher, Coulter, Reschly, & Vaughn, 2004; Gargiulo & Bouck, 2017; Higgins, Raskind, Goldberg, & Herman, 2002; May & Stone, 2010; Smith, Osborne, Crim, & Rhu, 1986).

In sum, a learning disability can be defined as a neurological condition that interferes with an individual's ability to store, process, or produce information. LD can affect an individual's attention, memory, coordination, reasoning, emotional maturity, selection and focusing on relevant stimuli, and the perception and processing of visual and/or auditory information (NJCLD, 2008). These processing difficulties are presumed

to be the underlying reason why students with LD experience one or more of the following characteristics: reading problems, deficits in written language, underachievement in math, poor social skills, attention deficits and hyperactivity, and behavioral problems (King-Sears, Swanson, & Mainzer, 2011).

Prevalence

According to Pullen, Lane, Ashworth, and Lovelace (2011), “As a field, the area of learning disabilities is constantly growing and changing, and it is one of the most active areas of special education research” (p. 187). In fact, LD are among the most common disorders in school-age children across countries and cultures worldwide, with approximately 5–15% prevalence rates (American Psychiatric Association [APA], 2013). In the United States, between 1997 and 2013, the proportion of children identified as having a learning disability varied only slightly, staying between 7–8%, and was 8% in 2013 (Child Trends Databank, 2014). The National Center for Education Statistics (NCES, 2018a) reported that the number of children and youth ages 3–21 receiving special education services was 6.7 million, or about 13% of all public-school students, in 2015–2016, with 34% of these children receiving special education services under IDEA for specific learning disabilities more than any other type of disability.

It is also of note that boys are more likely to be labeled with a learning disability than girls. In 2013, 9% of boys and 6% of girls, ages 3–17, had a learning disability (Child Trends Databank, 2014). Researchers have not yet determined a reason for this. These statistics show LD to be the most common type of disability in the field of special education, which is the reason why LD are considered to be a high-incidence disability.

The amount of time students with disabilities spend in general education classes is significant. Previous research shows that the percentage of students with LD who spent most of the school day (i.e., 80% or more) in general education classes increased from 47% in 2000–2001 to 63% in 2015–2016 (NCES, 2018a). More than two-thirds (70%) of students with LD spent most of the school day in general education classrooms.

Characteristics

Students with LD are typical children and have average or above-average intelligence. In fact, research shows that students with LD have the ability to learn, and their ability may be even higher than that of their typical peers, if they are taught in the ways they learn best (Ali & Rafi, 2016; Carr & Bertrando, 2012; Sze, 2009). Also, students with LD are an extremely heterogeneous group, meaning that no two students possess the identical profile of strengths and weaknesses. Students differ, too, in their coping skills. According to Bowe (2005),

Some [students] learn to adjust to LD so well that they ‘pass’ as not having a disability, while others struggle throughout their lives to even do ‘simple’ things. Despite these differences, LD always begins in childhood and always is a life-long condition. (p. 71)

Lerner (2000) identified nine learning and behavioral characteristics of individuals with LD:

1. Disorders of attention
2. Reading difficulties
3. Poor motor abilities
4. Written language difficulties

5. Oral language difficulties
6. Social skills deficits
7. Psychological process deficits
8. Mathematical disorders
9. Information processing problems

Not all students with LD will display these characteristics, and many students who demonstrate these same behaviors are quite successful in the classroom (Gargiulo & Bouck, 2017). As Epler-Brooks (2018) stated, “For students with a learning disability, it is the quantity, intensity, and duration of these characteristics that lead to problems in school and elsewhere” (p. 47). Due to the effect on cognitive processes, students with LD may have difficulty in a variety of academic areas (Malmgren & Trezek, 2009; Melekoglu, 2011; National Center for Learning Disabilities [NCLD], 2013) as well as social and emotional development (Hughes et al., 2011; Kavale & Forness, 1996; Milsom & Glanville, 2010; NJCLD, 2008). While a student with a learning disability may have difficulties in all academic areas, difficulty with reading is by far the most common characteristic of students with LD (Fletcher, Lyon, Fuchs, & Barnes, 2007; Jitendra & Gajria, 2011; Wei, Blackorby, & Schiller, 2011).

The terms learning disabilities (LD) and reading disabilities (RD) are not interchangeable since not all students identified with LD have difficulty with reading (Denton & Al Otaiba, 2011). However, it is estimated that as many as 85–90% of students with LD have reading problems so significant that they cannot read and understand grade-level material (Hasselbring & Bausch, 2006; Kaluger & Kolson 1978;

Melekoglu, 2011), and even the low estimates are approximately 60% (Bender, 2001). In most cases, this difficulty in reading, contributes to the growing achievement gap between students with LD and their typical peers (Malmgren & Trezek, 2009).

In 2011, Melekoglu noted that the reading performance of students with LD is an average of 3.4 grade levels behind their peers without disabilities. This may be due to decreased remedial reading instruction that students received at the upper-elementary and secondary grades (Vaughn et al., 2010). To illustrate, the impact of reading instruction fades starting in fourth grade, when content area learning (e.g., science, history, and mathematics) becomes the main focus of daily instruction, causing a decline in reading achievement for upper-elementary and secondary students with LD. Much of this decline is due to problems in reading comprehension. Therefore, a learning disability in reading affects the student's ability to decode and/or understand the meaning of words and passages (Denton & Vaughn, 2008). This delay in developing foundational skills in reading—phonemic awareness, phonics, fluency, vocabulary, and comprehension—results in delays in other academic areas that require the use of these skills (e.g., writing, spelling, science, math, and social studies; NCLD, 2013).

Vocabulary Knowledge and Development

Vocabulary is one of the five essential components of reading that every child needs in order to grow into a confident and fluent reader. Vocabulary is generally defined as the knowledge of word and word meanings. More precisely, reading vocabulary refers to the kind of words that students should know in order to read and comprehend increasingly demanding text (Butler et al., 2010). Stahl (2005) stated, "Vocabulary

knowledge is knowledge; the knowledge of a word not only implies a definition, but also implies how that word fits into the world” (p. 95). Since vocabulary knowledge is something that expands and deepens over the lifespan, it can never be completely mastered (Honig, Gutlohn, & Diamond, 2012). The process of acquiring new words to use in daily life is called vocabulary development. The aim of vocabulary development is to help students learn the meanings of new words and concepts in various contexts and across all academic content areas. To do so, teachers must provide students with explicit instruction on important words in the text or with strategies for independently learning word meanings (Weiser, 2013). Either way, vocabulary development is a vital skill.

Importance of Vocabulary Knowledge and Development

The NRP (2000), the RAND Reading Study Group (2002), and the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010)—all major study groups in the United States—identify vocabulary as an essential component of reading instruction. In addition, research shows that the beginning of vocabulary development occurs before any formal educational instruction ever takes place. For example, early preschool vocabulary skills were found to be predictive of comprehension in the early stages of reading (Bianco et al., 2012). This means that students will be behind in literacy skills at the start of formal instruction if they are lacking in key experiences or interactions that build vocabulary.

Moreover, teaching vocabulary helps students understand and communicate with others, because without sufficient vocabulary, students may have difficulty understanding others or expressing their own ideas sufficiently. Wilkins (1972) wrote that “...while

without grammar very little can be conveyed, without vocabulary nothing can be conveyed” (pp. 111–112). This point reflects the experience of many English language learners; even without grammar, with some useful words and expressions, they can often manage to communicate (Alqahtani, 2015). Vocabulary knowledge also plays a significant role in both the learning-to-read and reading-to-learn phases (Pullen, Tuckwiller, Konold, Maynard, & Coyne, 2010).

Learning-to-read phase. Numerous researchers found that when children are *learning to read* in the initial primary grades (grades K–3), their early capabilities to decode are dependent on oral vocabulary (Carlson, Jenkins, Li, & Brownell, 2013; Hudson, Isakson, Richman, Lane, & Arriaza-Allen, 2011; Pullen et al., 2010). When applying letter-sounds to printed materials, the reader can gain an understanding of the text only if the resultant oral vocabulary is a known word in the learner’s oral vocabulary. The NRP (2000) explained, “When the word is not in the learner’s oral vocabulary, it will not be understood when it comes in print” (pp. 3–4). However, when children grow up and begin the phase of reading to learn, their reading abilities change.

Reading-to-learn phase. In the later primary grades (fourth grade and above) when children are *reading to learn*, there is evidence that vocabulary knowledge also plays an important role in comprehending texts. Once students enter fourth grade, most of the information they need is given to them in textual format, and the focus changes from learning to read to reading to learn. That means, the tasks required in content-area classrooms expect readers to be proficient in reading to learn from expository texts, a skill that Chall, Jacobs, and Baldwin (1990) asserted is not emphasized during the earlier

years of reading development. As students encounter vocabulary related to content areas other than English (e.g., social studies, mathematics, and science), they are required to utilize their background knowledge and prior experiences about a topic to make sense of the word (Weiser, 2013). Lee, Grigg, and Donahue (2007) indicated that nearly two-thirds of fourth-grade students cannot read with understanding in fourth-grade-level content-area materials. Therefore, students use their concept knowledge and context to make meaning. However, many fourth-grade students may begin to fall behind and experience difficulty reading informational textbooks; this decline in performance has been referred to as the “fourth-grade slump” (Chall et al., 1990).

This deficit often increases over time, affecting the students’ achievement. This problem begins early in childhood and becomes increasingly apparent over time. By the time they graduate from high school, students with limited vocabulary know only one-fourth as many words as their academically successful peers (Lubliner & Smetana, 2005; Stanovich, 1986). Consequently, limited vocabulary may prevent students from being able to comprehend the more academic texts that they encounter in the later primary grades, resulting in a lack of content knowledge in those subject areas as well as poorer reading achievement. The question remains to what extent the vocabulary knowledge is crucial to reading comprehension and in determining how well students will be able to comprehend the texts they read in later grades (late elementary, middle, and high school).

Reading Comprehension and Vocabulary Connections

The strong connection between the levels of vocabulary knowledge and reading comprehension has been widely documented (e.g., Carlson et al., 2013; Chall et al., 1990;

Cunningham & Stanovich, 1997; Freebody & Anderson, 1983; Kame'enui & Baumann, 2012; Lesaux & Kieffer, 2010; Oakhill et al., 2015; Reed et al., 2016; Rupley, Logan, & Nichols, 1998; Wagner & Meros, 2010). In their article, Rupley, Logan, and Nichols (1998) stated, "Vocabulary is the glue that holds stories, ideas, and content together... making comprehension accessible for children" (p. 339). Comprehension is far more than recognizing words and remembering their meanings. Indeed, reading comprehension requires reading quickly (reading fluency), keeping track of what words are in the sentences (working memory), vocabulary meaning and text conventions (semantics/grammar), and understanding what is read (receptive language; Berninger & Richards, 2002; Cutting, Materek, Cole, Levine, & Mahone, 2009). Fresch (2008) found that knowing the meaning of words is the main source of difficulty for students, if they are even able to pronounce them.

It is also true that comprehension is impossible if a student does not know the meanings of a sufficient proportion of the words in the text. In fact, earlier studies have shown that readers cannot understand a text adequately without knowing at least 90–95% of the words in the text (e.g., Hirsch, 2003; Hu & Nation, 2000; Laufer, 1989; Nagy & Scott, 2000), which may explain why several researchers described the relationship between vocabulary and reading comprehension to be *reciprocal* (e.g., Freebody & Anderson, 1983; Oakhill et al., 2015; Stanovich, 1986; Wagner & Meros, 2010). Thus, a reader who knows more words is likely to have better reading comprehension; likewise, a reader who is successful with comprehension will frequently have the opportunity to learn more words. In contrast, a reader who struggles with vocabulary will be less likely

to comprehend text and less likely to learn new vocabulary. Hence, it appears that acceptable levels of comprehension occur when the reader knows what most of the words in the text mean. Put simply, the ultimate outcome of great vocabulary knowledge in students is improved comprehension (Graves, August, & Mancilla-Martinez, 2012; Nagy & Scott, 2000; Stahl & Nagy, 2006).

Vocabulary knowledge is not only fundamental for reading comprehension, but it is also considered an important predictor of how students will be able to comprehend the texts they read in late elementary grades, middle, and high school (Chall et al., 1990; Graves, 2006; Lesaux & Kieffer, 2010). According to Kame'enui and Baumann (2012), one of the major limiting factors for most students in third grade and above is vocabulary, not reading mechanics (i.e., decoding print into words). Other studies have shown that (a) students with limited vocabulary by third grade have low comprehension scores in late elementary school (Grades 4–6; Chall et al., 1990), and (b) vocabulary that is tested orally at the end of first grade is an important predictor of reading comprehension 10 years later (Cunningham & Stanovich, 1997). These studies demonstrate a relationship between vocabulary differences and subsequent comprehension abilities in students. Once students fall behind in word knowledge in earlier grades, the gap widens between them and their academically successful peers without intervention, thus affecting the reading comprehension ability for those students in later grades. This large vocabulary gap will further widen for students with LD who are even more behind in acquiring vocabulary. This necessitates paying greater attention to their vocabulary development in order to address the deficits in their vocabulary knowledge.

Importance of Vocabulary Development for Students with Learning Disabilities

Students with LD often read less per day than their typically developing peers, making it more difficult for them to acquire new vocabulary and comprehend academic texts (Weiser, 2013). Although students with LD may struggle with basic reading skills such as decoding words, comprehension is the greater weakness; this affects their ability to understand the meanings of words and passages (Denton & Vaughn, 2008). As previously discussed, students need to understand the meaning of critical words they will be reading to promote comprehension. Weiser (2013) illustrated that students with LD often struggle with reading comprehension because they do not possess the oral vocabulary that is a prerequisite to their understanding and retention of content-area texts; this weakness in oral language skills directly impacts their vocabulary development. Researchers such as Kamil et al. (2008) and Loftus and Coyne (2013) demonstrated that as those students with LD get older, it is critical for both oral and written vocabulary development to increase so that they can comprehend increasingly more complex grade-level texts.

Limited vocabulary knowledge will not only hinder the ability to be successful in comprehending grade-level texts for students with LD, but it will also impede their participation in classroom discussions and doing well on assignments, quizzes, and tests (Weiser, 2013). These deficiencies often increase over time, and the best solution includes consistent and persistent investment in vocabulary development, which can be implemented through a variety of ways (Ebbers & Denton, 2008). Without intervention, vocabulary deficits occurring in students' early years of education will negatively affect

their performance in reading and other academic areas as they progress into middle and high school.

While many researchers have examined the correlation between vocabulary and comprehension, other researchers have described how a larger vocabulary contributes to other areas of school success. Manzo, Manzo, and Thomas (2006) asserted, “Word learning can improve the *capacity* to learn,” and “a rich vocabulary increases comprehension and, therefore, most all learning” (p. 615). Simply stated, Lubliner and Smetana (2005) declared, “Children with larger vocabularies find reading easier, read more widely, and do better in school” (p. 163). Thus, recognizing and understanding more words increases the likelihood that students will comprehend what they are reading and, therefore, perform better in school.

In fact, in addition to improving students’ reading comprehension and writing quality, providing vocabulary instruction also improves students’ listening vocabulary and their speaking vocabulary (Joshi, 2005; Kame’enui & Baumann, 2012). Vocabulary knowledge, including both oral and written vocabulary, is critically important for a student’s success in school (Kamil et al., 2008). Additionally, students with larger vocabularies usually articulate responses to questions and ask better questions than their peers with limited vocabularies (Blewitt, Rump, Shealy, & Cook, 2009). Not only can vocabulary contribute to achievements in language arts classes, but it improves other areas of study as well.

Factors that Influence the Vocabulary Development

The lack of devoting enough time for vocabulary instruction in classrooms is one of the significant factors negatively affecting vocabulary development for many students with and without disabilities. Students with vocabulary deficits—such as those with LD—need more time to learn strategies to help them acquire words in order to reduce their vocabulary gap (Lockavitch, 2010; Nash & Snowling, 2006; Swanson & Vaughn, 2010). Unfortunately, students typically receive little vocabulary instruction in their classrooms (Blachowicz & Fisher, 2006). In a classic classroom observational study, Durkin (1979) found that only 19 minutes (i.e., 0.4%) of the 4,469 minutes of observed reading instruction were devoted to vocabulary instruction and that almost no vocabulary-development instruction took place during content instruction such as social studies. In another early study on the amount of time vocabulary instruction occurs in classrooms, Roser and Juel (1982) found that teachers spent an average of only 1.67 minutes on vocabulary during each reading lesson. The findings of these and other studies on the effects of the lack of attention to vocabulary instruction led Beck et al. (2002) to draw this research-based conclusion, “All the available evidence indicates that there is little emphasis on the acquisition of vocabulary in school curricula” (p. 15). More recent studies also have revealed that time devoted to vocabulary instruction has not increased (see Berne & Blachowicz, 2008; Cunningham, Zibulsky, Stanovich, & Stanovich, 2009), with one study showing that less than 12 minutes was devoted, on average, to vocabulary instruction throughout the week (Graves et al., 2012). This

explains at least one contributing factor to why students, especially students with LD, continue to have gaps in their vocabulary knowledge.

Additionally, students with LD need considerable repetition in order to acquire vocabulary words (Chall, Conrad, & Harris-Sharples, 1991). Several researchers noted that multiple exposures to vocabulary words being taught increase struggling readers' understanding of word meanings and their use of targeted words (e.g., Biemiller & Boote, 2006; Hudson et al., 2011; Justice et al., 2005; Pullen et al., 2010). Consequently, students with LD are more likely to benefit from repeated exposure to the same unknown word. Justice et al. (2005) mentioned that exposing students to novel words, either within the text of a single book or through repeated readings of the same book, facilitates their learning of those words. Research also indicated that vocabulary is learned gradually (Baumann, Kame'enui, & Ash, 2003). Therefore, students—especially students with deficits in vocabulary—need to be given time to process, hear words, and use them in context to their meaning. Word knowledge grows slowly and gradually, requiring multiple exposures to words. Misulis (1999) stated, “In order for words to be truly learned, that is, to be used and committed to long-term memory, they must be reinforced many times in meaningful ways” (p. 25). It appears that when a student is exposed to a new word multiple times, this will not only reinforce the word and its meaning, but it also moves the word from short-term to long-term memory. Dale and O'Rourke (1981) summarized the four incremental stages of word knowledge as:

1. I never saw it before;
2. I've heard of it, but I don't know what it means;

3. I recognize it in context—it has something to do with...;
4. I know it (p. 10).

Thus, the more exposure a student has to a word, the more information accumulates about that word. By gathering more information about a word, the student will be able to define, comprehend, and remember the word. Providing good vocabulary instruction that builds repetition into the learning process enables students with LD to learn more words more quickly. Furthermore, using and applying numerous examples of a word in various contexts enhances word knowledge (Justice et al., 2005) and provides those students with LD with experiences and clues to the word's meaning that builds over time and will help shape their understanding of the unknown word (Stahl & Nagy, 2006).

Lastly, students learn best when they are actively engaged with and, thus, can deeply process the materials (Richek, 2006). Unfortunately, many times students are not engaged with the delivery of the content and then left with holes in their vocabulary expansion. Teachers must create engaging methods of delivering instruction to further assist in students retaining words. Considering language engagement, Coyne, McCoach, and Kapp (2007) implemented two studies aimed to evaluate the effectiveness of an extended instructional approach for teaching vocabulary to kindergarten students within small-group intervention models during storybook reading in order to determine the amount and quality of word learning that students experience as a result of extended instruction. The researchers described extended vocabulary instruction as a robust approach that (a) provides direct, explicit teaching of word meanings that includes using both contextual and definitional information; (b) increases students' exposure to target

words in various contexts; and (c) encourages students' deep processing of vocabulary word meanings by describing how the target words relate to other words and to their own experiences.

The purpose of the first study was to compare extended instruction of target words to incidental instruction (i.e., reading the story without direct instruction). The second study was conducted to compare extended instruction to embedded instruction (i.e., providing simple definitions within the context of the story). The results of the two studies showed that when students in the extended instruction group were given the opportunity to learn the targeted words through dialogue and interactive experiences that extended beyond only listening to the oral reading of the text, they produced a complete knowledge of the targeted words. Additionally, the students maintained knowledge of the words for eight weeks after instruction. In contrast, incidental exposure to the words resulted in minimal word learning, and embedded instruction resulted in only partial word learning. Therefore, the researchers concluded that, in both studies, extended instruction produced a statistically significant difference that resulted in greater word learning than either embedded or incidental instruction on all three measures used. The researchers also found that in order to facilitate students' deeper understanding of word meanings, vocabulary instruction must involve teacher-student discussion and interactive activities that focus on the new words.

Leung (2008), who investigated preschoolers' learning of scientific vocabulary, reported similar results. The researcher explored the efficacy of retelling and hands-on science activities that were relevant to concepts presented in three informational picture

books. Participants in this study were 37 preschoolers (ages 3–4 years) enrolled at an urban child development center in the southeastern United States. The students were randomly assigned to two groups—retelling or no retelling conditions. All students in the study engaged in small group, interactive, repeated readings of informational books on a science topic: light and color. Then, half of the students immediately retold the books. After the retellings, all students participated in hands-on activities. The researcher used two standardized vocabulary measures, the PPVT-III and the Expressive Vocabulary Test (EVT), to posttest students on 32 targeted words.

The findings of Leung’s study indicated that students who participated in the book retellings condition made the greatest gains in word knowledge. The author concluded that young children can learn vocabulary and the scientific names for complex concepts by using repeated interactive reading (a hands-on activity) plus a retelling approach that related to the targeted words and meanings. These results suggest that incorporating language-engagement activities, such as a hands-on activity and story retelling, while reading a storybook, facilitate young children’s acquisition of unfamiliar words as well as improve their deeper understanding of the targeted word meanings.

So far, the studies reviewed emphasize that vocabulary instruction does lead to gains in comprehension, and students’ achievements can be negatively affected by poor vocabulary knowledge and poor reading comprehension. As Jenkins, Matlock, and Slocum (1989) stated, “With each year of schooling, texts take on a larger role in instruction, and factors that may inhibit comprehension of these texts, such as a lack of vocabulary knowledge, can be expected to have increasingly detrimental effects on

achievement” (p. 217). Therefore, reading researchers emphasize the importance of teachers using effective vocabulary instruction methods—appropriate to the reader’s age and ability—to improve comprehension as early as possible. Although it appears that students can benefit from teachers who include vocabulary instruction in their lessons, there is little research that conclusively identifies the best methods or combinations of methods of vocabulary instruction. The next section reviews the existing knowledge base on vocabulary acquisition and instructional practices.

Vocabulary Instruction

The terms *intervention* and *instruction* are often treated as synonyms. What distinguishes intervention from instruction is that intervention is individualized to the student’s needs, provides supports and scaffolds for student success, and is often conducted in one-on-one sessions or small groups (Nelson & van Meter, 2006). Additionally, the terms are sometimes combined as *instructional intervention*, meaning a specific program or set of steps to help a student improve in an area of need. Instructional interventions are formalized, but they can be flexible too (Lee, 2018). For instance, if a particular program is not helping a student, the teacher or interventionist may change it. This might mean increasing the amount of time a student receives reading support each week. Or it might mean getting more intense support, such as moving from small-group instruction to one-on-one help.

In order to frame my research questions, I have explored and investigated issues related to the topic of vocabulary instruction, collecting related literature. The literature search centered on the following questions:

- What is vocabulary instruction?
- What effect does vocabulary instruction have on student reading achievement?
- What makes an effective vocabulary instructor?
- What are effective methods of implementing vocabulary instruction?

I consider these four questions in the following sections and then explore vocabulary instructional strategies and interventions for students with and without LD.

What is Vocabulary Instruction?

Although vocabulary instruction is typically thought of as a process in learning words, there is no clear agreement about an appropriate definition. In the past, vocabulary instruction most often was unplanned and incidental, primarily driven by student questions and “teachable moments.” When students encountered unfamiliar words, they were directed to look words up in a dictionary, write definitions, and use words in a sentence (Bromley, 2007), or simply, they were given a quick oral definition of the unfamiliar word. Although requiring students to find word meanings independently may save time for the teachers and enables them to progress to content matter, several researchers have asserted that such traditional vocabulary instruction is problematic (Ford-Connors & Paratore, 2015; Irvin, 2001; Nagy, 2005; Richek, 2006) for at least two reasons. First, asking students to look up words in the dictionary and write corresponding sentences is not necessarily the most engaging method for all students (Richek, 2006). Making vocabulary learning fun and creative can foster a sense of competence in students, and thus, students want to learn and use new word meanings.

Second, the traditional vocabulary instruction is of limited value, particularly in improving students' reading comprehension (Ford-Connors & Paratore, 2015; Stahl & Fairbanks, 1986). Students need to know how a word functions in various contexts. Specifically, even though students may learn the meaning of a word related to certain content, they should be able to transfer vocabulary meanings to other contexts. This is not to say that using the definitional approach should be avoided. Rather, learning definitions of words can be effective when students already have an understanding of the underlying concept of the term. Therefore, the focus of effective vocabulary practices should be placed on improving comprehension, not just word knowledge alone (Beck & McKeown, 1991; Stahl & Fairbanks, 1986). As previously mentioned, word knowledge and comprehension are reciprocal. The question remains to what extent the vocabulary instruction influences students' reading achievement.

What Effect Does Vocabulary Instruction have on Student Reading Achievement?

There is a considerable amount of scientific research showing that effective methods of vocabulary instruction improve students' reading comprehension, which in turn, increases their reading achievement. For instance, Rupley et al. (2009) stated, "As children's vocabulary grows their ability to comprehend what they read grows as well; furthermore, as their comprehension skills grow so do their abilities to learn new words in context" (p. 336). This statement emphasizes how crucial it is for teachers to implement effective vocabulary instruction, since it has direct links to improving student reading comprehension. Students with broad vocabulary knowledge are better able to infer meanings of unfamiliar words in the texts that they read (Rupley et al., 1998). As

previously mentioned in the report published by the NRP (2000), researchers noted that reading comprehension is a complex cognitive process that “cannot be understood without examining the critical role and importance” that vocabulary development and vocabulary instruction play in the understanding of what has been read (p. 228). The NRP report concluded that vocabulary should be taught directly and indirectly to help student reading achievement.

Consequently, engaging students in explicit methods of vocabulary instruction is a critical element in language development and also in students maintaining high levels of reading achievement. By implementing effective methods of vocabulary instruction, teachers have the potential ability to increase student reading achievement. Consequently, vocabulary instruction is what the instructor makes it. And for this reason, I find it worthwhile to examine what makes an effective vocabulary instructor.

What Makes an Effective Vocabulary Instructor?

Research shows that effective methods of vocabulary instruction currently remains a topic of interest and concern for classroom teachers, reading researchers, and teacher educators. Berne and Blachowicz (2008) pointed out that teaching vocabulary may be problematic because many teachers are not confident about the best practice in vocabulary teaching and, at times, do not know where to begin informing an instructional emphasis on word learning. Simply stated, Bromley (2007) declared, “Many teachers know they need to do a better job teaching vocabulary to students who find reading difficult” (p. 528). However, selecting the most appropriate method of vocabulary instruction is a difficult task. Gambrell, Morrow, and Pressly (2007) explained the

difficulty for teachers as being “increasingly faced with a diverse group of learners in terms of current word knowledge, linguistic background, learning styles, and learning abilities. It is up to teachers to make word learning enjoyable, meaningful, and effective.” (p. 179).

Helping teachers to be effective vocabulary instructors, who can meet the diverse needs of all learners, reading researchers Blachowicz and Fisher (2004) suggested guidelines that apply to most classrooms in most situations of vocabulary instruction. In fact, those researchers, with the findings of the report compiled by the NRP (2000), conducted a study and found five guidelines that apply when defining an effective vocabulary instructor. An effective vocabulary instructor:

1. Builds a word-rich environment in which students are immersed in words for both incidental and intentional learning and the development of “word awareness” (p. 67).
2. Helps students develop as independent word learners.
3. Uses instructional strategies that not only teach vocabulary but also model good word learning behaviors.
4. Provides explicit instruction for important content and concept vocabulary, drawing on multiple sources of meaning.
5. Uses assessment that matches the goal of instruction.

These evidence-based guidelines are proven to be effective. Thus, classroom teachers or reading researchers can confidently use these guidelines when preparing the teaching-related activities that will be included in their vocabulary instructional

interventions. These guidelines are important because, in order to develop and design an explicit and effective vocabulary instructional intervention, the five guidelines must be in place to ensure that the designed lesson plans in the vocabulary instructional intervention are the most powerful they can be. Once the researcher or classroom teacher has evaluated and concluded that these guidelines are present, they can then continue to examine the effective methods for implementing the vocabulary instructional intervention that they designed.

What are Effective Methods of Implementing Vocabulary Instruction?

Countless vocabulary instructions and interventions have been tried and examined in elementary classrooms over the years, yet there is not a singularly accepted method for implementing vocabulary instruction currently. In a chapter written more than three decades ago, Beck, McKeown, and Omanson (1987) synthesized the findings from a series of vocabulary research experiments in which they engaged, and they concluded, “Research has provided much useful information about vocabulary learning and instruction. What it has not provided is a simple formula for optimal instruction, because no such formula can exist” (p.150). Additionally, from the best-evidence synthesis of vocabulary research, the NRP committee (2000) concluded that the research base is inadequate for determining the best method for teaching vocabulary. However, the NRP committee advised teachers to utilize a variety of strategies (e.g., direct, indirect, multiple exposures, computer use) when teaching vocabulary to students. The committee also explained that the selection of a strategy may depend on the goal of instruction, the

assessment tool to monitor learning, the development level of the student, the objective for vocabulary learning, and how much time is available.

Vocabulary instruction for students with and without learning disabilities.

Stahl and Fairbanks (1986) conducted a pivotal meta-analysis on vocabulary instructional strategies for students with typical development, which continues to be a cornerstone for current work on vocabulary instruction in the area of vocabulary acquisition. The researchers examined the components of effective vocabulary instruction and investigated the effect of vocabulary instruction on comprehension. After analyzing 52 studies, conducted between 1932–1986, the researchers identified three key characteristics of effective vocabulary instruction for mainstream education, which remain the gold standards for effective vocabulary instruction: (a) definitional and contextual information, (b) in-depth teaching of the meanings of words, and (c) multiple repetitions or exposures to new words.

Graves (2006) described the strategies of effective vocabulary instruction slightly differently: (a) review, rehearse, and remind “students about the word in various contexts over time” (p. 70); (b) discuss word meanings to actively involve students; and (c) spend time teaching, discussing, and learning about each word. Moreover, Graves provided guidance about what should *not* be done during vocabulary instruction:

- Do not give words out of context or ask students to look up meanings in a dictionary.
- Do not use speeded trials with individual words.
- Do not use word mazes.

- Do not teach words as an alternative label when they represent new and challenging concepts.
- Do not teach spelling rather than vocabulary.
- Do not assume that contextual clues are enough to yield precise word meanings.

Similar to typically developing students, these guidelines for vocabulary instruction are found to be suitable for students with LD. In their widely cited article, Jitendra et al. (2004) reviewed and summarized the findings of published research on vocabulary instruction involving students with LD. Specifically, Jitendra et al. reviewed 19 vocabulary studies that comprised 27 investigations conducted between 1978 and 2002. This review of 19 studies included 17 group-design studies and 2 single-subject design studies. A total of 901 students with LD participated in the studies reviewed. The researchers found the studies reviewed suggested that vocabulary instruction for students with LD could lead to gains in word knowledge, and the students learned more through direct instruction than incidental learning through context. Furthermore, they found that the guidelines for vocabulary instruction with typically achieving students were also appropriate for students with LD. In particular, the researchers supported the vocabulary instructional strategies that provided a combination of definitional and contextual information and encouraged students to use vocabulary expressively.

Types of vocabulary interventions. Vocabulary interventions conducted in the last few decades fall into two categories: *non-generative* and *generative*.

Non-generative vocabulary interventions. This type of intervention, with the aid of a strategy and/or a device, teaches students the meaning of a single word (Harris et al., 2011). In other words, non-generative vocabulary approaches help students write and memorize one definition of the word. Examples of non-generative strategies include: (a) dictionary-use strategies, (b) keyword mnemonic instruction, (c) semantic feature analysis, and (d) semantic mapping. Although these strategies may help students learn the meaning of the targeted vocabulary word, students do not learn more than that word. Research shows that students are exposed to roughly 88,700 word families over their 13 years of schooling (Nagy & Anderson, 1984), average students learn to read approximately 3,000–4,000 words each year, and average students acquire reading vocabularies for about 50,000 words by the time they graduate from high school (Graves, 1986, 2016; White et al., 1990). Given these figures, it seems apparent that non-generative strategies alone may not be the most practical instruction for accomplishing such a significant task. This is especially true for students who have vocabulary deficits, such as students with LD, who need to learn thousands of words to decrease the gap between their performance and the performance of their typically achieving peers (Harris et al., 2011). This is why non-generative strategies, which teach students only one word at a time, lack the power to build vocabulary at a sufficient rate (Nagy & Anderson, 1984).

Considering dictionary usage, for example, many researchers indicated that reliance on the use of the dictionary alone is poor practice because students are more likely to face difficulty in associating the proper meaning with an unfamiliar word or to

select the first listed meaning of a word. As Marksheffel (1966) stated over 50 years ago, “The particular word may have a number of definitions, and the student may be confused as to which meaning is associated with the puzzling word” (p. 249). In 1987, Miller and Gildea conducted a study that revealed the extent to which students can misunderstand definitions. They studied the ability of fifth and sixth graders to generate appropriate sentences after reading traditional dictionary definitions. The researchers judged over 60% of the sentences that students constructed to be odd, often because students chose only a fragment of the definition on which to base their sentences. For instance, while a definition of the term *eroding* in a dictionary included the phrase “eating out,” one student participant in Miller and Gildea’s study generated the sentence, “Our family erodes a lot” (p. 98). Researchers such as Marksheffel (1966) and Graves (2016) warned that although using the dictionary to define words is possible, students’ dictionary usage does not guarantee that they will associate the appropriate meaning with an unknown word. Graves (2016) asserted that “using the dictionary to define words is possible but difficult for elementary students and frequently not 100% successful even for college students” (p. 32).

All in all, there is good evidence to suggest that the non-generative vocabulary interventions traditionally used in elementary classrooms are not helping students acquire the amount of vocabulary needed to become proficient readers. Consequently, these interventions are not efficient or effective in providing the amount of words or approaches required to close the gap in vocabulary development for students, especially students with poor vocabulary knowledge, including students with LD. Thus, it is

important for teachers to teach students a variety of independent word-learning strategies to be able to acquire vocabulary independently and become independent word learners to learn more words.

Generative vocabulary interventions. In addition to teaching students the meanings of unfamiliar words, generative vocabulary interventions allow students to infer the meaning of related new words. Specifically, these interventions assist students in becoming independent word learners by teaching them word-learning strategies, which allow them to use vocabulary knowledge that can transfer to the learning of new words (Nagy, Berninger, & Abbott, 2006). A large body of research supports teaching students vocabulary strategies in a way that can build students' ability to use the strategies on their own (Blachowicz & Fisher, 2006; Harris et al., 2011; Nash & Snowling, 2006). Unfortunately, although teaching these independent word-learning strategies may promote students' self-monitoring and comprehension to increase, few studies have investigated the role of generative strategies in vocabulary acquisition, and even fewer have focused on elementary students with LD. Baumann, Kame'enui, and Ash (2003) noted the lack of attention to generative strategies to vocabulary acquisition, "In spite of the conventional wisdom that instruction in morphemic analysis is an appropriate transferable and generalizable vocabulary strategy, research on the efficacy of such instruction is fairly limited" (p. 773).

Two effective generative vocabulary strategies found to support vocabulary acquisition for students, especially during incidental reading of content area texts, are: (a) context clues (i.e., teaching students how to use context or identify clues around an

unfamiliar word that will help them define the word), and (b) word parts (i.e., teaching students how to analyze words and word parts such as prefixes, roots, and suffixes).

These independent word-learning strategies are also known as *contextual analysis* and *morphemic analysis* respectively. For the purpose of this research, I focused on describing the literature related to the contextual analysis strategies.

Contextual analysis strategies. Contextual analysis is defined as the use of clues within the context of the text to derive word meanings. Deriving word meanings from written context is a significant source of vocabulary expansion for students at all ages (Fukkink, Blok, & de Glopper, 2001). Until the late 1990s and early 2000s, only a few reviews of the contextual analysis literature investigated the instructional effects of the skill of determining word meaning from written context (Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998; NRP, 2000; Swanborn & de Glopper, 1999). The NRP committee (2000) promoted the importance of contextual analysis but also emphasized that research in the types of contextual analysis strategies and techniques that are most effective are in a “state of infancy” (p. 29). Specifically, the committee noted the importance of contextual analysis since students learn words incidentally through reading. In fact, research has shown that contextual analysis can occur incidentally or can be explicitly taught.

Most of the studies conducted until the late 1990s that examined the role that contextual analysis played in word learning focused mainly on incidental word learning from written context—students read a short passage and have to define the word orally (Fukkink, 2005; Fukkink et al., 2001). Nagy, Anderson, and Herman (1987) examined

incidental word learning during natural reading and concluded that students can, and indeed do, derive and learn new vocabulary incidentally while reading. In a meta-analysis of 20 studies that included a total of 2,130 students assigned to experimental and control groups, Swanborn and de Glopper (1999) investigated students' word learning from context when not prompted to do so. The researchers found that incidental word learning from context during reading takes place, and students learn an average of 15 of the 100 unfamiliar words they encounter during reading when word learning is measured directly after the reading event. The researchers also found that students at higher grade levels and students with higher reading ability are better able to use context and that text containing fewer unknown words better facilitate learning from content.

Although the aforementioned studies showed that students can benefit from the incidental learning of word meanings, research conducted with students who have poor vocabulary knowledge (Nash & Snowling, 2006) or students with reading disabilities (although limited) has demonstrated the need for more explicit strategy instruction, without which they will learn far fewer words incidentally compared to their typically developing peers (Baumann, Edwards et al., 2003; Nagy & Scott, 2000). Research also has shown that students, particularly those with reading difficulties, have better vocabulary outcomes with explicit instruction (Carlo et al., 2004; Ebbers & Denton, 2008; Kieffer & Lesaux, 2008; Scott & Nagy, 2000). *Explicit instruction* involves instructing students on different types of context clues (e.g., synonym, antonym, definition, cause and effect, inference) that surround the unknown words or instructing

students on a general strategy aimed at improving their ability to use context for learning unfamiliar word meanings.

A study done by Buikema and Graves (1993) suggested that teaching students to use context clues can be effective only if the instruction is explicit, scaffolded, and provides practice and feedback. Many other researchers have argued that if context clue strategies are explicitly taught to students using appropriate contexts, they were able to generalize what they learned (Stahl & Fairbanks, 1986; Stahl & Nagy, 2006). Students, including students with reading difficulties or with poor vocabulary knowledge, need explicit and systematic instruction in context clue strategies to improve their vocabulary acquisition. The next question for researchers is: To what extent can instruction on context clue strategies affect the abilities of students, particularly students with LD or who have poor vocabulary knowledge, to derive the meanings of unknown words from context and to improve their vocabulary acquisition?

Review of the context clues literature. I conducted an in-depth review of the research literature over the past 60 years to determine the existing knowledge base regarding descriptive and experimental vocabulary studies that included context clue strategies. To identify related studies, I searched multiple web-based databases: Education Resources Information Center (ERIC), PsycINFO, Education Full Text (EBSCO), Journal Storage (JSTOR), and Council for Exceptional Children (CEC). I used numerous key words in the search, including: *vocabulary instruction, vocabulary acquisition, vocabulary learning, contextual analysis, context, context clues, contextual clues, contextual cues, learning from context, incidental word learning, word learning*

strategies, word learning skills, generative vocabulary strategies, elementary students, students with learning disabilities, students with reading disabilities, students with reading difficulties, and students with poor vocabulary knowledge. Additionally, I hand searched and accumulated papers from conferences and other unpublished sources. I used the following criteria for selecting studies to review. The study (a) included elementary students who were in kindergarten through fifth grade, (b) examined the effects of contextual analysis interventions on students' word learning and/or vocabulary performance, (c) used an experimental design or quasi-experimental design with experimental control or alternative treatment condition, and (d) was published in a peer-reviewed journal.

The computer search produced a large number of studies examining contextual analysis interventions, however, not all the identified studies met my inclusion criteria. A number of studies involved students of English as a second language (e.g., Elgort, Brysbaert, Stevens, & Van Assche, 2018; Fischer, 1994; Hu, 2013; Hu & Nassaji, 2014; Huckin & Zhendong, 1986; Mondria & Wit-De Boer 1991; Montelongo, Hernández, Herter, & Cuello, 2011; Walters, 2004). Several other studies involved listening instead of reading (e.g., Eller, Pappas, & Brown, 1988; Elley, 1989; Hennebry et al., 2017; Lenhart et al., 2018; Sénéchal & Cornell, 1993; Teng, 2016). Because these studies were beyond the scope of my research, I excluded them from review. Although my focus was on elementary students, some studies included a range of participants that encompassed both middle and high school students. I accepted these studies due to the dearth of studies investigating the effect of contextual analysis strategies on elementary students.

However, I explicitly noted participants' ages or grade levels in the presentation of each study.

When applying the aforementioned restrictions, a search of the databases using the identified keywords yielded 15 studies. Among the included studies, 13 were quantitative using experimental or quasi-experimental designs, and two were meta-analyses. Additionally, most of these 13 experimental or quasi-experimental studies were not included in either of the two meta-analyses reviewed. In sum, this review included only those studies that taught students a generalized strategy or list of context clues with the intention of improving their efficiency in using context to learn unfamiliar word meanings.

Meta-analysis studies on teaching context clues. Two teams of researchers (Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998) have conducted meta-analyses of studies examining instruction in the use of context to infer word meanings. In one of the most well-known meta-analyses, Kuhn and Stahl (1998) investigated instructional strategies aimed at teaching students to be more efficient at learning words from context. In particular, they examined 14 studies to improve words students were learning from context, through instruction on using context clues. The researchers grouped the findings of each study by type of measure and examined commonalities among studies. Of the 14 studies reviewed, the researchers found 10 studies that included a control group and measured students' ability to derive word meanings from context. Based on the findings across those 10 studies, Kuhn and Stahl determined that students benefit from explicit instruction on how to use context clues to derive word meanings of unfamiliar words

from context. This conclusion is explained by their finding that in the studies that included treatment and control groups, the students in both conditions did not differ significantly on the outcome measure, suggesting that practice in deriving words, rather than the strategies, may make a difference in vocabulary development. Kuhn and Stahl asserted that context clue is a general strategy aimed at helping students contend with unfamiliar words in a wide variety of texts.

A similar result was found in a meta-analysis conducted by Fukkink and de Glopper (1998). Based on their meta-analysis of 21 studies of instruction in context clues, the researchers concluded that instruction aimed at enhancing the skill of deriving word meaning from context during reading does have a positive effect. In addition, Fukkink and de Glopper noted that the vocabulary instruction that is based on context clues for deriving word meanings appears to be more effective than other instruction types or just practice.

Both of these meta-analyses—which overlap to some extent, with six studies simultaneously represented in both—found a positive effect for instruction in the use of context, although Kuhn and Stahl (1998) cautioned that in studies that employed a practice-only condition, no difference was seen between treatment and practice groups, suggesting that students benefit as much from repeated practice opportunities as they do from specific instruction. Kuhn and Stahl found that context clue instruction was just as successful as either cloze exercises (i.e., students were provided texts that contained blank spaces with specific words omitted, drawing students' attention to the context) or general strategy instruction, while Fukkink and de Glopper found that context clue

instruction was superior to other forms of instruction. These findings contrast with Sternberg's (1987) finding that training in specific types of context clues was least effective.

However, the two reviews have four limitations. First, some relevant studies were not included in the reviews, which are discussed in a following section. Second, most of the published studies in both reviews included students in middle and high schools. Third, several of the studies examined, especially in Fukkink and de Gloppe's meta-analysis, have used the researcher-developed tests as the singular measurement instrument to measure the effect of the instructional interventions provided in these studies. And finally, the statistical power of the reviewed studies is small, and statistical significance of the results is, therefore, not the best method to evaluate findings.

Consequently, findings from both meta-analyses should be interpreted cautiously due to the absence of robust research evidence that used explicit instruction to teach context clues. This also confirms the claim that research assessing the effects of explicitly teaching students to use context is still in its infancy (Kieffer & Lesaux, 2012; NRP, 2000). Caution is also necessary because there is no evidence that instruction in one kind of context clue transfers to other kinds of clues. Nevertheless, these meta-analyses make an important contribution to what is known about training students to use context. They also enable us to conclude, with some degree of confidence, that training in this skill appears to be worthwhile, at least for students who have vocabulary deficits, while at the same time highlighting areas for future research.

Experimental studies on teaching context clues. Over the past six decades, 13 experimental studies have examined the effectiveness of using context clues to improve students' ability to use context to infer the meanings of unfamiliar words. These studies can be roughly divided into three broad categories: (a) those that examine the effects of teaching specific types of context clues, (b) those that look at the effects of teaching a general strategy for dealing with unknown words, and (c) those that examine the effects of a combination of strategy and selected types of context clues.

Context clue instruction. Early studies examining the effectiveness of making students aware of specific types of context clues include one by Hafner (1965), who carried out a study with elementary school students, to determine what effect an experimental program consisting of lessons in the use of selected context clues would have on tests measuring reading comprehension, vocabulary-in-context, and context comprehension. The experimental program specifically included lessons sought to improve the ability of students to use the following selected types of context clues to infer the meanings of unknown words: contrast, explanatory words and phrases, meaning expressed in a single sentence and gained through interpretation of the sentence, and indirect explanations (inference). A total of 75 fifth-grade students from three classes in two elementary schools in the same city participated in this study. Students in one of these classes were used as the experimental group while students in the other two classes were used as the control groups. Results of Hafner's study indicated that the experimental group showed significant gains on the test of vocabulary, but gains on the comprehension test were not significant, and there was no gain on the context comprehension (measured

by cloze) test. However, in comparison to control groups, Hafner noted that the experimental group showed the highest percentage gains in comprehension, vocabulary, and context comprehension. Furthermore, when controlling for intelligence and background, the researcher found that more students with training made gains than did students without training.

A similar early experimental study conducted by Askov and Kamm (1976) investigated whether instruction in certain types of context clues enhanced students' ability to use those context clues. The researchers specifically sought to answer these questions: (a) can instruction in given categories of context clues improve students' abilities to use those context clues, and (b) if so, do they retain their use of context clues over a period of time? All third-, fourth-, and fifth-grade students ($n = 133$) enrolled in two public elementary schools in Minnesota participated in the study. The classroom teachers taught students in the experimental groups two types of context clues: cause-effect and direct description. The results of the two-week instructional intervention showed that the experimental groups demonstrated significantly more positive change in the ability to infer meaning from these two types of context clues than did the control groups. The researchers concluded that "teaching a classification of context clues, such as cause-effect and direct description, will promote greater use of such clues and enhance the student's ability to determine the meaning of an unknown word in a sentence" (p. 343). Additionally, the results of the study led the researchers to suggest that "teaching a classification system seems to be effective at all grade levels, even at the beginning of grade 3" (p. 341).

In another study of this type, Patberg, Graves, and Stibbe (1984) compared the effects of three instructional conditions: (a) active instruction in contextual analysis, (b) presentation of guidelines and practice without active instruction, and (c) control group. All fifth-grade students who were not in the control group were exposed to three 30-minute instructional sessions to teach them how to use synonym clues and contrast clues. In this study, the researchers used three measures: (a) multiple-choice test of words taught, (b) multiple-choice test of words determined from synonym and contrast clues, and (c) multiple-choice test of words determined from a single 394-word passage. Findings showed that both instructed groups outperformed the uninstructed control group on words taught to them on the first measure, both instructed groups outperformed the control group (on process taught) on the second measure, and there were no significant differences between the three groups on the measure of connected text. It appears that teaching students how to use different types of context clues (e.g., synonym clues and contrast clues) helped students in this study to determine the meanings of unfamiliar words presented in short texts containing the sorts of clues taught. However, when Patberg and Stibbe (1985) replicated the same study, they found no effects of instruction in using context clues on students' ability to infer the meanings of unfamiliar words.

Two studies conducted by Baumann and his colleagues (Baumann et al., 2002,; Baumann, Edwards et al., 2003) are considered the most ambitious to date (Graves, 2016). Both studies investigated two types of inferential strategies: contextual analysis and morphemic analysis. The eight types of context clues presented in the contextual analysis condition in both studies were: word definitions; synonyms; appositives;

antonyms; examples; summary; figurative language; and mood, tone, or setting. In the 2002 study, using a quasi-experimental design, 88 fifth-grade students in five classrooms were assigned to one of four instructional groups: morpheme-only, context-only, combined morpheme and context, or an uninstructed control group. Except for students in the control group, the experimenters provided the instructional program to all students in the other three groups. The instructional program in Baumann et al.'s (2002) study consisted of twelve 50-minute lessons that followed an explicit instruction model (verbal explanation, modeling, guided practice, and independent practice); gradual release of responsibility; and provided students with declarative, procedural, and conditional knowledge about the strategy they were learning. Results indicated that students in both the contextual group and the morphemic group were better at inferring the meanings of transfer words—that were not taught—on an immediate posttest but not on a delayed posttest that was administered five weeks after the immediate posttest.

In the 2003 study, Baumann, Edwards et al. employed a quasi-experimental design with 157 fifth-grade students in eight social studies classes to examine the students' ability to derive word meaning following instruction on a combination of morphemic and contextual analysis on social studies textbook vocabulary instruction. Specifically, the experimenters taught students 25 lessons from the fifth-grade social studies curriculum that lasted approximately 45 minutes each. Then, the experimenters compared students' learning to that of students who were taught the vocabulary of the social studies texts in a traditional fashion. Results indicated that students receiving the experimental treatment were more successful at inferring the meanings of

morphologically and contextually analyzed words on a delayed posttest three weeks after learning the words but not on an immediate posttest. There was one major limitation to this study: Some groups of students received instruction in two different sets of words; thus, it is difficult to determine whether the instructional program or characteristics of the word sets produced the differences.

On the whole, it appears from most of these six studies that context clue instruction has a positive effect on students' ability to infer the meanings of unfamiliar words from context, when the instructed clues are present. However, there are several areas of concern. First, the results from the study by Hafner (1965) are inconclusive, shedding a positive light on context clue instruction only after some statistical manipulation. Second, the lasting effect of such training was measured in only three of these studies, with conflicting results. Askov and Kamm (1976) and Baumann, Edwards et al. (2003) found that the gains achieved in ability to infer word meanings from context were maintained after three or six weeks, but Baumann et al. (2002) found no such lasting effect on a delayed posttest, administered five weeks later.

Third, these studies did not provide evidence that training in specific types of context clues transferred to other kinds of context clues, thus leaving open the question of the efficiency of this method. The amount of time spent on each clue in the studies ranged from one hour to two hours, with the study demonstrating the strongest effect—significant gains lasting over six weeks (Askov & Kamm, 1976)—invested the most time per clue. If there is no carry-over to other types of clues, context clue instruction could consume a significant amount of time and effort in the classroom. Consequently, this

might discourage classroom teachers from implementing the context clue instruction at all, or they may implement it incorrectly. Several researchers pointed out that low demand on an already busy teacher is an important consideration when implementing instructional interventions or behavior modification strategies into the classroom (Alzahrane, 2016; Amato-Zech, Hoff, & Doepke, 2006). Therefore, interventions that require less of the teacher's time are more likely to be both implemented and implemented correctly in comparison to interventions that are time intensive and take away from classroom instruction.

Finally, it is assumed that the time and effort spent instructing students in the context clues to improve their ability in inferring unfamiliar words from context stems from a desire for both better reading comprehension and faster vocabulary development. Again, only three of the previous studies included a measure of reading comprehension (Hafner, 1976; Baumann et al., 2002; Baumann, Edwards et al., 2003), and none of the three studies showed significant gains, although both of Baumann et al.'s studies concede possible measurement issues with their true/false comprehension questions. None of the six studies measured vocabulary development. Given these concerns, it would be inappropriate to recommend context clue instruction without further investigation.

General strategy instruction. The second category of studies—those studies that investigated the effects of teaching a general strategy for coping with unknown words while reading—includes a study by Carnine, Kame'enui, and Coyle (1984), which is the study that first applied instruction in the derivation of word meaning. In this study, the researchers examined the differential effects of three intervention strategies (rule-plus-

systematic-practice, systematic practice only, and no intervention), designed to facilitate the use of context clues in learning the meanings of unfamiliar words, on 37 fourth-, fifth-, and sixth-grade students from three multi-graded classrooms in three elementary schools located in the same city. Their strategy involved a rule like, “When there’s a hard word in a sentence, look for other words in the story that tell you more about that word” (p. 197). Students were also told that the unfamiliar word gave information about a character in the story, or what and how something is done. Classroom teachers modeled two of the three intervention strategies. The results of the five-day intervention period indicated that rule-plus-systematic-practice and systematic-practice-only conditions produced higher transfer scores than a no intervention condition.

In 1989, Jenkins et al. compared the effects of two different vocabulary instructional methods with 135 fifth-grade students in six classrooms from three different schools in the same school district. One instructional method emphasized direct teaching of the individual meanings for a set of unfamiliar words. Rather than teaching specific meanings, the second instructional method emphasized teaching students how to use a strategy to derive meanings of unfamiliar words from sentence context. After randomly assigning three classrooms to each of the vocabulary instructional methods, and to allow comparison across levels of implementation as well as between treatments, each of the six classrooms was randomly assigned to low (nine sessions), medium (11 sessions), or high (20 sessions) amounts of practice for their respective instruction method. The classroom teachers in both instructional methods used scripted lesson plans that the researchers created to ensure proper implementation of the experimental techniques.

Additionally, the researchers adopted a four-step strategy from Kranzer (1988), encapsulated by the acronym SCAR: Substitute, Check the fit, Acept the substitution, or Rethink, if necessary. The Jenkins et al.'s study (1989) used a modified strategy: SCANR. The modified procedure involved five steps: Substitute a word or expression for the unknown word, Check the context for clues that support your idea, Ask if substitution fits all context clues, Need a new idea?, and Revise your idea to fit the context. Kranzer combined these two last steps into one step, *rethink, if necessary*. The students used the acronym to help them easily remember the steps. The classroom teachers modeled the SCANR steps first in Jenkins et al.'s study. The researchers administered two pre and posttests to all students to assess their mastery of the specific word meanings taught and their ability to derive meanings from unknown words. More specifically, one of these pre and posttests required students to write synonyms or definitions for words in isolation, and the other required students to write synonyms or definitions for words given in context.

With a maximum score of 20, the results of the pretests revealed that students scored only an average of 0.73 in the isolation assessment, and an average of 1.12 for the words in context assessment. On the posttests, students from all three levels of practice (low, medium, and high) for individual meanings instruction outperformed students from all three levels of deriving meaning instruction. In particular, the students from the high practice individual meanings group outscored (10.05 – words in isolation test, 13.24 – words in context test) all other students on the posttests. Based on their data, the researchers concluded that instruction in deriving meaning was effective with medium

and high amounts of practice but not with less practice. This led Jenkins et al. to suggest that providing higher levels of instruction for individual word meanings leads to a deeper understanding of those words, and providing higher levels of training in deriving meaning improves students' ability to derive word meanings from context. After comparing the two instructional methods to "previous studies of economical ('lean') and more time-consuming ('rich') forms of vocabulary instruction," the researchers suggested the potential of these two instructional methods for improving students' vocabulary learning in the classroom (p. 215).

Nash and Snowling (2006) also investigated the effects of using two different vocabulary instructional methods—the definition method and context clues method—to improve vocabulary knowledge and reading comprehension. A total of 24 British students (seven- and eight-year-olds) with poor vocabulary knowledge participated in the study. Those participants were equally assigned in closely matched pairs to one of the two designed teaching programs. Thus, 12 students were taught new vocabulary words using definitions; the other 12 students were taught a strategy for deriving word meanings from written context. In particular, students in the definition program received the definitions of words and then wrote these words on paper, while the students in the context program received a few sentences for each word in context and then discussed context clues.

Each program involved two 30-minute sessions a week for six weeks, resulting in six teaching hours in total. The same experimenter (Nash) taught students in both programs two words per lesson, one noun and one verb, for a total of 24 words. Nash

reversed the noun and verb order in every other lesson. As a result, the researchers were able to ensure that only the method of instruction varied between groups because both groups in this study had the same instructor, amount of time per lesson, days of instruction, vocabulary words, and number of students. Researchers assessed students from each intervention group using two vocabulary posttests, one immediately after teaching (posttest 1) and then three months later (posttest 2).

Results indicated that, when tested immediately after teaching, both groups improved equivalently in vocabulary knowledge for the taught words. However, three months following the intervention, the context group demonstrated significantly better expressive vocabulary knowledge, particularly with nouns. On the transfer test, the context group also showed significantly better comprehension of text containing a number of the taught words and demonstrated that they could independently use the newly acquired strategy to derive meanings from written context. The researchers concluded that the context method was a more effective intervention than the definition method in increasing vocabulary knowledge and improving reading comprehension in students with poor vocabulary knowledge. Consequently, the researchers recommended using the context clues method with students who require extra help developing vocabulary and comprehension skills. One potential limitation was that, since the experimenter taught both programs, there could potentially be experimenter bias. In addition, there was an absence of an uninstructed control group, and a small sample size limited statistical power.

In another experimental study comparing the effects of using the definition method and context clues method, Martin-Chang et al. (2007) examined how well students retain and apply these vocabulary strategies in future reading tasks. A total of 28 second-grade students between the ages of 6 and 8 years, with average reading ability, from different schools participated in this study. The researchers did not include English as a Second Language (ESL) students or those with developmental delays (e.g., Down Syndrome, Autism) in the sample.

For each training program, sessions lasted for 15–20 minutes over two weeks. The training program schedule was as follows: day one was a pretest, days two through four were training days, days five through 11 were a retention period, and on day 12 a retention task and transfer task were given. This program was implemented twice—once for context training and once for isolated word training. Therefore, each student was an active participant in the study for approximately one month. The researchers created individualized instructional materials for each study participant. These materials included only those words that average readers in second grade were unable to read or name in context during the pretest.

Students in the isolated word training program read words from personalized sets of cue cards, and they were given 1.5 seconds to respond. If the student was correct, no feedback was given, and the experimenter showed the next card. If the student was incorrect or did not respond in the allotted time, the experimenter said the word correctly and then moved on to the next word. A total of 12 repetitions (2-word repetitions per list x 2 trials x 3 days = 12 word repetitions) of each word occurred over the three-day

training period. In contrast, students in the context training program read individualized reading passages with the experimenter. Each story contained target vocabulary words, which the students read, while the experimenter read the rest of the story aloud to the student. Following the same procedure as the isolated word training, students were given 1.5 seconds to respond. Correct responses meant the experimenter continued reading. Responses taking longer than 1.5 seconds or incorrect responses resulted in the experimenter saying the word and then continuing to read the rest of the passage.

Eight days after training, all student participants were exposed to a retention task and then a transfer task. Retention tasks utilized the same materials used in training, while the transfer task used the original screening passage, which students had only seen one time 12 days earlier in the pretest. The results of the retention task, when students were asked to read the same words eight days later, showed that students from the context group scored 97% accuracy, while the isolated word group scored 94.5% accuracy. For the transfer task, where students read a novel passage, students from the context group scored 85% accuracy, and the isolated word group scored 70% accuracy. These data show a positive relationship between learning words in context and the ability for students to retain and transfer these words. In addition, the data gathered from this experiment suggest that context training promotes word acquisition and retention beyond that experienced from reading words in isolation.

Yuen (2009) showed similar findings in a more recent study. This research explored the use of context clues to gain knowledge of new vocabulary words during reading. A total of 20 third-grade students in a self-contained classroom at a public

school participated in this study. During the three-week intervention, the researcher provided direct instruction to students on how to use three context clues strategies: locating appositives, searching for explicit definitions, and using prior knowledge. Each intervention week was devoted to teaching one of those three strategies. To measure the dependent variable in this study, the researcher implemented pre and posttest assessments, pre and postintervention interviews, and classroom sweeps. The researcher found that teaching students how to use context clues while reading improves their understanding of new vocabulary words. Moreover, results from classroom sweeps demonstrated students became more attentive to their reading throughout intervention, which implies that they were implementing context clue strategies to assist their reading. Lastly, postintervention interviews suggested that students instilled context clues in their awareness to help decipher new words in reading.

All in all, the positive results demonstrated by the experimental groups in these five studies confirm that teaching students how to use a general strategy for coping with unknown words while reading helps them to derive meanings of unfamiliar words from written context. Additionally, of these five studies, three studies compared the effects of two vocabulary instructional methods (the definition method and context clues method) on students' ability to infer meanings of novel words during reading. The results revealed that the context method was a more effective intervention than the definition method in increasing students' vocabulary knowledge and improving their reading comprehension. Moreover, the findings of the study by Martin-Chang et al. (2007) showed that context training also promoted word acquisition and retention beyond that experienced from

reading words in isolation. This is an important finding because “successful reading instruction entails not only acquiring new words but also remembering them after training has finished and accessing their word-specific representations when they are encountered in new text” (p. 37). Thus, providing instruction in context clues as a strategy has a sound and persuasive rationale because even a small improvement in ability to infer word meaning of unfamiliar words would result in a sizeable number of words learnt over time.

However, findings from these five studies should be interpreted cautiously due to several limitations. First, none of these studies were conducted with students with special needs, including those with LD. In fact, there was only one study (Nash & Snowling, 2006), where the researchers conducted the study with students with poor vocabulary knowledge. Second, the majority of these studies did not include a delayed posttest. It is important to include a delayed posttest after the immediate posttest to examine the students’ retention of the learned context clues strategy and to measure the lasting effect of such a strategy on students’ growth in reading ability, specifically in vocabulary. Knowing this will help us to answer this important research question: Is there a difference in retention between short-term learning (measured by the immediate posttest) and long-term learning (measured by the delayed posttest) for students receiving the context clues strategy in the treatment condition?

Third, although it is significant to include both treatment and control groups in a study to compare the influence of the vocabulary instructional intervention on the performance of students in the treatment groups with their peers in the control groups who did not receive the intervention, only one study (Carnine et al., 1984) included both

groups. Fourth, the amount of time spent on each intervention in the aforementioned studies was somewhat different, ranging from a three-day training period to intervention that lasted for 20 sessions. These problems, along with the problem of vast disparity among the interventions and the amount of time spent on each intervention presented, raise questions: How do we present the context clues strategy in the classroom, and how effective is it in improving reading comprehension and vocabulary development for all students in the classroom, including students with LD?

Combination of strategy and context clues. Two studies (Buikema & Graves, 1993; Fukkink, 2002) incorporated a mixture of specific types of context clues and general strategy training. In one notable study, Buikema and Graves (1993) investigated the effectiveness of an instructional unit designed to improve seventh- and eighth-grade students' ($n = 38$) ability to use context to infer the meanings of unfamiliar words. The instructional unit guided the students through a strategy of dealing with unfamiliar words using descriptive context clues, beginning with *boxing in* the word—listing the words and phrases in the context that gave information regarding sense, action, or purpose aspects of the word—and then thinking about what the unfamiliar word might mean. Students were encouraged at this point to bring in their experience of the world as well as their knowledge of parts of speech. Finally, students guessed at the meaning of the word. Although the instructional unit focused on only one type of context clue (descriptive clues), the researchers' emphasis was on using the strategy rather than on the clue type. When comparing students in the experimental group with uninstructed control-group students on tests of word knowledge and inferring from context, the results showed

significant positive differences between the groups, with the experimental group better able to infer word meanings from context.

Fukkink (2002) also combined specific types of context clues with a general strategy. In a randomized experiment, the researcher assessed the effects of a 12-lesson instructional program on deriving word meaning from written context and incidental word learning with 145 Dutch fourth-graders of below-average reading ability. The program consisted of direct instruction in both a general strategy and selected types of context clues (synonyms, antonyms, and direct explanation clues). However, the majority of the program's time was spent on instruction in and practice of the strategy. The strategy involved four steps: (1)

1. Searching the context for a possible answer (by searching for the three types of clues taught)
2. Thinking of an answer (i.e. guessing)
3. Checking the guess for appropriateness
4. Formulating a definition

Posttests measuring incidental word learning and the skill of deriving word meaning from context showed no significant improvement in the experimental group as a result of instruction. This finding is in direct contrast to those of the majority of studies that are included in this literature review of the effects of strategy or clue instruction (e.g., Baumann et al., 2002, Baumann, Edwards et al., 2003; Buikema & Graves, 1993; Martin-Chang et al., 2007; Nash & Snowling, 2006; Yuen, 2009).

Fukkink (2002) offered two possible explanations for his surprising results. First, Fukkink noted that the instructional programs in most previous studies he reviewed have been provided by the researchers or specially trained teachers (Fukkink & de Gloppe, 1998; see also Baumann et al., 2002; Helman et al., 2015; Tomesen & Aarnoutse, 1998). Fukkink suggested that the results of such studies might not be generalizable to the regular classroom. Thus, to avoid this obstacle, the instructional program in Fukkink's (2002) study was implemented by the students' regular classroom teachers. However, this might represent threats to the internal validity of Fukkink's study since the regular classroom teachers might not have implemented the program as closely as intended, particularly the direct instruction and strategies because of their rather complex nature. Second, Fukkink speculated that perhaps the short time period of the program or the absence of sufficient practice opportunities during the program prevented students from gaining the full benefits of the instructional program. Unfortunately, Fukkink's study did not include a delayed posttest, which might have shown a long-term effect for the training, in spite of a lack of evidence for immediate effect.

In brief, the two studies described in this section examined the effects of a combination of a strategy and certain types of context clues in different ways and with diametrically opposite results. The studies differ in several respects: Buikema and Graves (1993) worked with U.S. seventh and eighth graders, presumably of mixed reading ability, while the student participants in Fukkink's (2002) study were Dutch fourth graders, specifically chosen for their low reading ability. Therefore, it is possible that the differences in age, reading ability, and/or language background might account for the

disparity in results between the two studies. Other differences include the strategies employed and the types of context clues presented. Given these differences, it is difficult to determine the source of the disparity in results. Furthermore, none of these studies included a delayed posttest, so it is not possible to pinpoint the effects of the learned context clues strategy in the long term on students' ability to glean the meanings of unfamiliar words.

Summary

To determine whether teaching students contextual analysis strategies provides effective vocabulary instruction, studies examined investigated such an instructional approach. Not all studies from the past 60 years have produced positive results, however, some have been successful. Baumann, Edwards et al. (2003) also confirmed that not all instruction in using context clues has been successful. In fact, teaching students to use context clues is a challenging task. As Fukkink (2005) noted, "Deriving the meaning of an unknown word from the written context is a complex and demanding task" (p. 24).

Additionally, research shows that students with LD are tremendously behind in the number of vocabulary words they need to know in order to succeed academically (Denton & Al Otaiba, 2011; Denton & Vaughn, 2008; Loftus & Coyne, 2013; Swanson & Vaughn, 2010; Weiser, 2013). When comparing generative and non-generative strategies, although non-generative strategies are effective in teaching students isolated words, these types of strategies are not as effective in helping students learn related relevant words (Harris et al., 2011). Effective instruction for students with reading difficulties, especially students with LD, must increase their vocabulary acquisition in a

small duration of time. Therefore, students with reading difficulties, particularly students with LD, need explicit and systematic instruction in context clue strategies to improve their vocabulary acquisition (Ebbers & Denton, 2008; Kieffer & Lesaux, 2012; Scott & Nagy, 2006). A generative vocabulary strategy, such as contextual analysis, can help identify context clues surrounding unknown vocabulary to help students figure out word meaning. Thus, teaching students contextual analysis skills helps them derive word meanings from text, and generative strategies, such as context clue strategies, are recommended (Kieffer & Lesaux, 2012).

Based on recent vocabulary studies, context clue strategies have led to improved vocabulary acquisition and comprehension with elementary students (Martin-Chang et al., 2007; Nash & Snowling, 2006; Yuen, 2009). However, whether instruction in context clue strategies can improve the ability of students with reading difficulties, including students with LD, to use context to derive the meanings of unfamiliar words has not been explored, thus remaining an open question, which was researched in this study. In particular, the purpose of the study was to examine the influence of vocabulary instruction that is based on a combination of strategy and certain types of context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. In addition, students' abilities to maintain and retain their skills following the intervention was investigated.

CHAPTER III

METHODOLOGY

This study was designed to find out more about how using the context clues strategy combined with learning certain types of context clues would help fourth-grade students with adequate and poor vocabulary knowledge figure out the meaning of those unfamiliar words during reading. In this chapter, I present the methods that were employed in the study. First, I provide a list of the research questions and related hypothesis that were tested. This is followed by a description of both the research design and data analysis procedures. In the next section, I describe the research setting, the participants in the study, and how they were recruited. Then, I present the vocabulary instructional intervention and all the instruments that were used to measure the variables examined in the study. Lastly, I provide details about the data collection and study procedures employed in the study, before closing the chapter with a summary.

Research Questions and Hypotheses

In chapter two, I identified three major research questions:

1. Does the vocabulary instructional intervention have any effect on the performance of the students on the experimenter-constructed test? If so, for which group of students (students with poor vocabulary knowledge [PVK] and students with adequate vocabulary knowledge [AVK]) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (1A–1D):

Subquestion 1A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) over time? I hypothesize that students in the treatment condition will perform better than students in the control condition on the experimenter-constructed test.

Subquestion 1B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the experimenter-constructed test.

Subquestion 1C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the experimenter-constructed test.

Subquestion 1D

Is there a difference in retention between short-term learning and long-term learning (Time 2 vs. Time 3) for students in the treatment condition? I hypothesize that the performance of the students will be better in short-term learning (Time 2) in the treatment condition on the experimenter-constructed test.

2. Does the vocabulary instructional intervention have any effect on the performance of the students on the standardized test (Gates-MacGinitie Reading Tests [GMRT])? If so, for which group of students (students with PVK and students with AVK) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (2A–2C):

Subquestion 2A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) at the pretest and delayed posttest? I hypothesize that students in the treatment condition will perform better than students in the control condition on the standardized test.

Subquestion 2B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the standardized test.

Subquestion 2C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the standardized test.

3. Are the students' results on both measures of vocabulary knowledge (the GMRT and the ECT) consistent across the two conditions (treatment condition and control condition)? I hypothesize that students' results on both measures of vocabulary knowledge across the two conditions will be consistent, which will confirm that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

Research Designs and Data Analysis

A quasi-experimental design was applied to this study in order to examine the influence of vocabulary instruction that is based on context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. Specifically, this study involved a comparison of two approaches: (1) *business as usual* instruction was used as a control condition and (2) a nine-day vocabulary instructional intervention was used as a treatment condition. Subjects were fourth-grade students with and without learning disabilities (LD), and measures were taken at pretest, immediate posttest, and delayed posttest (three weeks follow-up). The dependent variable in the study was a measure of the effects of a vocabulary instructional intervention on students' vocabulary acquisition. A 2 X 2 X 3 repeated-measures design with Conditions and Groups as between-subject factors and Time as a within-subject factor was employed in this study. The two between-subjects variables and one within-subjects variable is also called a Three-Way Mixed ANOVA

design (Howell, 2013). The design may be diagrammed as follows in Table 3.1, where G_i represents the i th group of participants.

Table 3.1 The Three-Way Mixed ANOVA Design Table

	Control Condition			Treatment Condition		
	Pretest	Immediate Posttest	Delayed Posttest	Pretest	Immediate Posttest	Delayed Posttest
PVK	G_1	G_1	G_1	G_2	G_2	G_2
AVK	G_3	G_3	G_3	G_4	G_4	G_4

Note: PVK = Students with poor vocabulary knowledge; AVK = Students with adequate vocabulary knowledge

A three-way mixed ANOVA design was used to analyze the results of an experimenter-constructed test (ECT) that was created for this study. In fact, this analysis tool is used primarily to understand if there is an interaction between the within-subjects factor and between-subjects factors on the dependent variable (Howell, 2013). Hence, the three-way mixed ANOVA design was used to determine whether any change in vocabulary acquisition (i.e., the dependent variable) was the result of the interaction between the between-subjects factors (i.e., conditions and groups) and Time (i.e., the within-subjects factor, consisting of three time points; vocabulary acquisition is measured “at the beginning of the study” [time point #1], “immediately following the intervention” [time point #2] and “three weeks after the immediate posttest” [time point #3]). The three-way mixed ANOVA also allowed for the examination of two-way interactions and

the main effects of the other factors to determine whether any change in vocabulary acquisition was due to one or two of the factors (i.e., conditions, groups, or time).

Additionally, I used a standardized norm-referenced reading test, known as the Gates-MacGinitie Reading Tests (GMRT), Fourth Edition, in the pretest as a general benchmark of students' vocabulary knowledge prior to the study and in the delayed posttest, given three weeks after the immediate posttest, to see if there was any growth in reading ability, specifically in vocabulary. I employed the Pretest-Posttest, Nonequivalent Group Design as another quasi-experimental design in this study. The term *nonequivalent* means that assignment to a group was not random. It does not mean that it was impossible to make a case for the similarity of the groups on relevant variables or characteristics (Wiersma & Jurs, 2009). The design may be diagrammed as follows in Table 3.2, where G_i represents the i th group of participants.

Table 3.2 The Pretest-Posttest, Nonequivalent Group Design Table

	Control Condition		Treatment Condition	
	Pretest	Delayed Posttest	Pretest	Delayed Posttest
PVK	G_1	G_1	G_2	G_2
AVK	G_3	G_3	G_4	G_4

Note: PVK = Students with poor vocabulary knowledge; AVK = Students with adequate vocabulary knowledge

This design aids in checking the extent of group similarity, and the pretest scores were used for statistical control and for generating gain scores (Wiersma & Jurs, 2009). Using the GMRT, all participants in both the control and treatment groups were pretested

and then posttested three weeks after the immediate posttest was completed.

Consequently, I was able to calculate the *gain score* by subtracting the pretest score from the delayed posttest score. Then, I compared this gain score in the GMRT with the results of the ECT, which was analyzed using the three-way mixed ANOVA analysis method, to see if the students' results were consistent. The consistency between those two measures confirms that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

Participants and Setting

A pilot study using a pre-post design on one group of participants was conducted in order to obtain an estimate of the potential effect size of the intervention, information that was then used to determine the sample size for the study. Results of the pilot study indicated a Cohen's *d*, effect size of 0.79 with a corresponding eta-squared effect size of 0.138. This is considered a moderate effect size for the proposed intervention on the performance of students without disabilities (Cohen, 1998). Subsequently, a power analysis using this effect size applied to a three way-mixed ANOVA with two between-subjects factors and a one within-subject factors with three time points indicated that a minimum sample size of 40 participants will result in power of 0.80, at the nominal alpha level of .05. An overview of the pilot study results, its effect size, and required sample size are available in Appendix A.

Since I was able to recruit more participants than the minimal sample size for the present study, I chose to increase the sample size to get strong and generalizable results. Research has shown that significant results issued from larger studies usually receive

more credit than those from smaller studies because of the risk of reporting exaggerating treatment effects with studies of smaller samples or of lower quality, and small trials are believed to be more biased than others (Biau, Kernéis, & Porcher, 2008). However, there is no statistical reason a significant result in a trial including 2000 participants should be given more weight or preference than a trial including 20 participants since the significance level chosen is the same in both trials. Small but well-conducted trials may yield a reliable estimation of treatment effect. Nonetheless, a large sample size does have some advantages, including that it allows a more precise estimate of the treatment effect, and usually, it is easier to assess the representativeness of the sample and to generalize the results (Biau et al., 2008).

Five fourth-grade classrooms, with a total of 60 students, participated in the present study. One of the student participants moved to a different school during the study. Thus, the total of students participating in the study was 59 students. This sample size was sufficient to achieve the aim of the study (Noordzij et al., 2010). These five classrooms came from four separate elementary school campuses, two public schools in the same school district and two private schools that exclusively serve students with diagnosed LD or learning differences such as Attention Deficit/Hyperactive Disorder (ADHD). These schools are located in three different urban cities in the southern United States.

Thus, for the purpose of this study, two groups of students participated: (a) students with diagnosed learning disabilities (LD; $n = 25$), and (b) students without disabilities (students with typical development [TD]; $n = 34$). Table 4.1 includes

information about the number of students in both groups (PVK and AVK) across both instructional conditions (treatment and control conditions) as well as the number of students with LD and TD in each group. In addition to these inclusion criteria, all student participants were elementary schoolers aged between 9 and 11 years old (fourth-grade students). I secured permission to conduct research from the Human Subjects Committee at the University of North Carolina at Greensboro (UNCG), the participating school district's research review board, the principals of the selected public and private schools, classroom teachers, parents of all students, and the students. A detailed description of the recruitment procedures that were implemented for this study is provided in the next section.

Data Collection and Study Procedures

In this section, I describe the data collection and study procedures that were employed for (a) recruiting participants, (b), conducting the vocabulary instructional intervention, and (c) choosing and creating the vocabulary instructional intervention and all the instruments that were used to measure the dependent variable examined in the study.

Procedures for Recruiting Participants

In order to recruit participants for this study, I contacted four elementary school principals, via letters, to request their permission to conduct the study at their schools. After approval was granted, I met all fourth-grade teachers in these schools to present the purpose, procedures, and participant inclusion criteria of the study. I clarified to teachers that only two fourth-grade teachers from each school were needed for participation in this

study. Additionally, I gave the teachers recruitment letters (information sheet templates) that were to accompany the adult consent form. Then, the classroom teachers who chose to participate gave their consent, and then they sent home consent forms to all students' parents in their classrooms. There were two copies of the parental permission form in each envelope along with a cover letter, so the parents were aware of the purpose of the consent form. Both the cover letter and parental permission form were also translated to Spanish for Spanish-speaking parents.

The parental permission form explained to parents the aim of the study and its procedure and asked for their consent to include their children in the study. Families were instructed to keep one copy of the parental permission form for their records. They were also instructed to return one signed form indicating consent or refusal sealed in the provided envelope. All students were given stickers for returning their family's sealed envelopes. Then, I collected the sealed envelopes from the teachers. On a different list, I recorded the names of parents who gave their consent to let their children participate to determine children who were included in the study. Finally, upon parental approval, I obtained students' signatures on the assent forms at the beginning of the study process.

Costs and payments to the participants. There were no costs to student participants or payments made for participating in this study. However, student participants received a small incentive (e.g., candy, pencil, eraser, toy) at the end of every completed lesson and test to increase their desire to continue participation in the study. Regarding teacher participants, there were no costs to them for participating in this study. In fact, to encourage teacher participation, all teacher participants in this study were paid

for their participation. Teacher participants, who provided the vocabulary instructional intervention for the treatment groups, were paid a total of \$400. Teacher participants in the control groups, who followed their usual instructional techniques in teaching unknown words, were paid a total of \$200. All payments were made two weeks after the completion of the study.

Potential risks to participants. The Institutional Review Board (IRB) at UNCG has determined that participation in this study poses minimal risk to participants. There were no known or foreseeable risks involved with this study. The classroom teachers provided the vocabulary instructional intervention. Thus, there was minimal emotional distress and embarrassment risks involved for participation in this study because this was just like any typical school day; students were learning, and if some students struggled, it was no more than what they experience during their normal classwork.

Procedures for maintaining confidentiality of the data collected. The research team has done everything possible to make sure that all information obtained in this study is kept strictly confidential unless disclosure is required by law. In any sort of report we might publish or present, we will not include any information that will make it possible to identify a participant. To do so, first, participants in this study were listed in a separate sheet using their first names and middle initials, and then were linked to special coding identifiers (e.g., A1, A2, A3, and A24). Using the codes ensured accuracy while entering student participants' testing scores and allowed for confidentiality at the same time. The printed coding sheet that had the special coding identifiers was sorted in print at my office. The electronic version that linked participants' full names to the coding document

was stored securely in the UNCG Box cloud storage, associated with the research team's access information, and only the research team had access to the research records. In this way, the list linking the participants' names to their codes was stored separate from the testing data, which ensured confidentiality of the testing data and the participants' identity.

The Coding Identifier Sheet will be stored electronically for five years before deleting it permanently since the UNCG Access to and Retention of Data policy requires that data be kept for five years after study completion. However, the printed version of the Coding Identifier Sheet, which had participants' first names and middle initials, was stored in my personal office as long as the study was conducted. As soon as the study was completed, all the printed materials related to the study were either stored electronically (when needed) on UNCG Box or destroyed; the printed version of the Coding Identifier Sheet, which had participants' first names and middle initials, was among the destroyed documents.

Videotapes were locked in a safe place and were destroyed after completion of the study. Only the research team viewed the videotapes for measuring the fidelity of the classroom teachers' implementation of the vocabulary instructional intervention.

Videotapes were not used in any public presentation.

Procedures for protecting the privacy of potential subjects during recruitment. The classroom teachers sent and received the consent forms from parents in sealed envelopes. However, I was the only one who opened these sealed envelopes. This ensured that no one, not even the classroom teachers, knew which students had elected to

participate in this study except me. Additionally, at the school site, only the coding identifiers were used to refer to participants when recording and documenting performance. The sheet with students' names attached to their coding identifiers was stored in my personal office and never appeared at the school sites. Finally, before beginning the instructional intervention, I obtained students' signatures on the assent forms.

Teacher intervention training. I provided training for the study procedures to the fourth-grade classroom teachers involved in the treatment groups in this study. Since this study was conducted in a whole group instruction format (sometimes called whole class instruction), the teacher participants in the treatment groups videotaped themselves as they were conducting the vocabulary instructional intervention using the Swivl and iPad devices that I provided to the teachers in order to allow the research team to measure the fidelity of their implementation of the instructional intervention. The primary focus of the videotapes was on the teachers' instruction and *not* on the students in the classroom.

The teacher participants in the treatment groups were trained twice (individually) for a total of one hour, about 30 minutes in each session, on how to conduct the vocabulary instructional intervention and how to use the Swivl and iPad devices. These training sessions were reduced from four to two based on Helman et al.'s (2015) recommendation, since teacher participants quickly acquired the target skills. These training sessions were conducted after school in the teachers' classrooms. Except for conducting the vocabulary instructional intervention, the teacher participants in both the control and treatment groups were not required to conduct any pre or posttest or collect

any data about the student participants in this study. I was responsible for conducting all the pre and posttests required in this study. Thus, there was no data collected by or from the teachers participating in this study.

Testing time. Before conducting any test, I discussed with the classroom teachers the best time and location for conducting the pre and posttests to avoid test strain and conflicts around the timing of the project, considering that these tests should be about the same time for all classrooms participating in the study. After consulting with the classroom teachers and upon their decision, I scheduled the training sessions for teacher participants, obtained students' signatures on the assent forms, began the instructional intervention, and conducted the testing sessions to administer the data collection tools.

Study Procedures

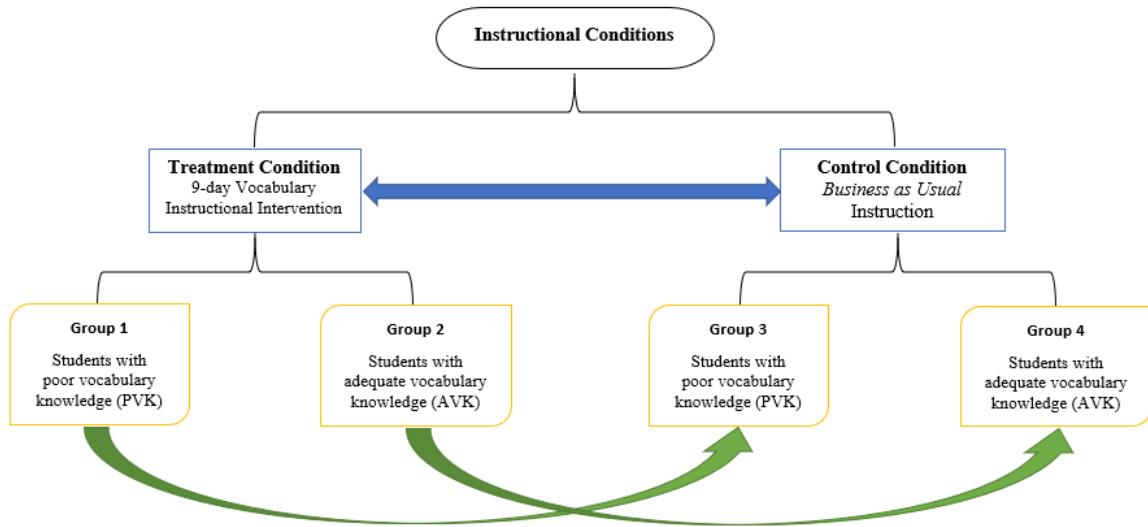
The study procedures are summarized in the following four steps:

Step 1: Vocabulary pretests. The study began with conducting two pretests on those students—in both the control and treatment groups—who had parental permission to participate in this study. However, before conducting any test, I discussed with the classroom teachers the best time and location for conducting the pre and posttests to avoid test strain and conflicts around the timing of the project. Upon the teachers' decisions, all student participants were exposed to two pretests. I conducted these pretests to evaluate student participants for possible pre-experimental differences in vocabulary ability and to classify them into two groups (those with adequate and poor vocabulary knowledge). These pretests also served as covariates in posttest analyses. These two pretests were completed on two separate days.

- Pretest 1: The GMRT, Fourth Edition, Form S was administered as a general benchmark of students' vocabulary knowledge prior to the study. Students were given 55 minutes to take this test, per the test's instructions (for more details about the GMRT and how and why it was used, see the section titled *The standardized test*).
- Pretest 2: An experimenter-constructed test (ECT, Form A), comprised of two segments (multiple-choice recognition and a meaning recall), was administered to evaluate student participants for possible pre-experimental differences in vocabulary ability and to use their scores on this test to serve as a covariate in posttest analyses. Forty-five question items were included in this test, and the total time required to complete it was 45 minutes (for more details about the ECT and how and why it was used, see the section titled *The experimenter-constructed test*).

Step 2: Instructional conditions. Figure 3.1 shows how students with adequate or poor vocabulary knowledge were divided into groups and instructional conditions.

Figure 3.1 Instructional Conditions



The control groups were used to control for the Hawthorne effect and using classrooms from different schools for the control and treatment conditions also controlled the Hawthorne effect (Oswald, Sherratt, & Smith, 2014). The Hawthorne effect refers to a phenomenon in which human subjects change their behavior simply because they are part of an experiment or study (McCambridge, Witton, & Elbourne, 2014). The students in the control groups followed a *business as usual* condition, meaning the students did not have explicit instruction on vocabulary strategies. Classroom teachers addressed vocabulary through informal discussions of words from the trade books as unknown words naturally occurred in the context of social studies lessons. However, all students in the control groups received the vocabulary instructional intervention once the study was completed and the intervention was found to be effective.

On the other hand, all students in the treatment groups in the other fourth-grade classrooms participated in the social studies lessons with context clues, but data were

only collected from those who had parental permission. Specifically, all students in these classrooms were exposed to the vocabulary instructional intervention for 20 to 30 minutes, which was part of the 50 minutes of social studies lessons, for nine consecutive days. The classroom teachers provided instruction on consecutive days in order to avoid the discontinuity that would occur with daily experimenter rotations among treatments that were provided (Baumann et al., 2002).

Step 3: Immediate vocabulary posttests. After completing the vocabulary instructional intervention, all student participants—in both the control and treatment groups—were administered an immediate posttest (Form B) to compare the influence of the vocabulary instructional intervention on the performance of students in the treatment groups with their peers in the control groups who did not receive the intervention. The total time required for completing the immediate posttest was 45 minutes, which was completed in one day.

Step 4: Delayed posttests. All student participants in the control groups and treatment groups completed two unannounced delayed posttests three weeks after the immediate posttest.

- Delayed Posttest 1: Delayed Context Recognition and Context Production Posttests. This was a readministration of the ECT but with a different test form (Form C), three weeks after the immediate posttest, to examine the students' retention of the learned strategy.
- Delayed Posttest 2: The standardized test (GMRT) was administered again but with a different test form (Form T) as a posttest, three weeks after the

immediate posttest, to see if there was any growth in reading ability, specifically in vocabulary.

These two delayed posttests were completed in two separate days. Table 3.3 includes information about the test forms, the number of questions involved in each test form, and the total testing time in minutes.

Instrumentation and Other Documents

This section lists and illustrates all instruments, materials, and documents that were used in this study.

Instructional intervention. I conducted a pilot study of the proposed vocabulary instructional intervention at one school. One Grade 4 classroom of 21 students participated in the pilot study. Due to difficulties in securing permission, there was not a control group involved in the pilot phase of the study. Therefore, the Grade 4 classroom that participated was involved as the treatment condition to test the proposed instructional intervention, the adequacy of research instruments, and assessment tools. I observed the lessons and adjusted the processes for the present study. An example of the lesson plans used in the vocabulary instructional intervention is available in Appendix B.

In the present study, students in the treatment groups received a nine-day lesson (the vocabulary instructional intervention), but data were only collected from those who had parental permission. These nine lessons included the direct instruction model for teaching and learning implicit skills (Archer & Hughes, 2011; Coyne, McCoach, Loftus, Zipoli, & Kapp, 2009; Shippen, Houchins, Steventon, & Sartor, 2005; Stahl & Hayes, 2013). Direct instruction aims at applying the strategy by providing declarative,

procedural, and conditional knowledge (Almasi & Fullerton, 2012; Iwai, 2016; López, Torrance, Rijlaarsdam, & Fidalgo, 2017; Zepeda, Richey, Ronevich, & Nokes-Malach, 2015). I included explicit think-aloud strategies to represent the mental processing involved in the strategy (Caldwell & Leslie, 2010; Kucan & Beck, 1997; Laing & Kamhi, 2002; McClintock, Pesco, & Martin-Chang, 2014). This nine-day lesson was designed to improve students' understanding of the use of context clues via signal words as a strategy to determine the meaning of unfamiliar words. To illustrate, context clues often contain signal words that readers can use to help them learn the meaning of new words. Thus, signal words point out the type of context clue being used.

Accordingly, the goal of the vocabulary instructional intervention was to improve the students' ability to derive word meanings from context by learning a strategy of dealing with both simple and complex words in textual contexts with varying degrees of contextual support. For example, students learned to distinguish between:

- Definition clue – “When the sun hit its *zenith*, which means right overhead, I could tell it was noon by the tremendous heat.”
- Antonym clue – “The soldier was very *intrepid* in battle, unlike the person next to him who was quite *cowardly*.”
- Synonym clue – “Captain Jackson’s uniform was *impeccable*. In fact, it was so *perfect* that she always got the highest score during inspections.”
- Comparison and contrast clue – “My brother is *enthralled* by birds similar to the way that I am *fascinated* by insects.”

- Example/list clue – “All animals share the same *vital* needs, such as food, water, and shelter.”
- Cause and effect clue – “If the snow continues, then it may *impede* our progress.”
- Inference/general sense clue – “*Patriotism* was a very strong force in the South. People loved their part of the country and were very proud to be a Southerner.”

To make the instructional intervention more practical and relevant, I took nonfiction texts from the fourth-grade social studies curriculum to use as part of the instructional intervention. I chose these texts after consulting with the participating classroom teachers.

Consequently, students in the treatment condition learned about the different types of context clues as well as their respective signal words. The classroom teachers also used the gradual release model strategy—I Do, We Do, You Do—for the instruction and practice of these 20–30 minute lessons. This model proposed a plan of instruction that included demonstration, prompts, and practice. This scaffolded instruction is broadly recognized as a successful approach for moving classroom instruction from teacher-centered whole group delivery to student-centered collaboration and independent practice (Fisher & Frey, 2008).

Additionally, the classroom teachers used a context clues graphic organizer as a tool to determine the meaning of unfamiliar words in a passage, text, or nonfiction excerpt. Thus, in addition to learning the different types of context clues, students in the

treatment condition learned how to use the context clues graphic organizer. The graphic organizer involved five steps:

1. Determining the unfamiliar word. The first step involves looking for the unfamiliar word in the sentence and underlining it.
2. Searching for an answer. After determining the unfamiliar word, the second step involves looking for signal words and a synonym, antonym, or other word clues that help students to infer the meaning. Starting with explicit clues in supportive contexts proved effective in helping younger readers (Fukkink & de Glopper, 1998). If there are no clues in the same sentence, students read a few sentences before and after the one containing the unknown word.
3. Thinking of the answer. The third step involves generating a tentative answer, making a guess about what the unknown word means.
4. Checking the answer. The fourth step involves substituting a word in the original sentence and evaluating its appropriateness in the specific sentence and context. Substituting an answer in the original sentence has been applied with some success by elementary students (Daalen-Kapteijns, Schouten-van Parreren, & de Glopper, 1997). Specifically, students look for evidence in the sentences around the word to confirm or deny guesses about what the word means. If it is right, the hypothesis is confirmed. If not, they try again.
5. Formulating a definition. The final step is defining the unfamiliar word as precisely and clearly as possible. This definition style, which may be associated with or even encouraged by the substitution strategy, often leads to

partially correct answers with conceptually complex words. Sometimes students need more words to generate a clear definition. For example, in his study, Fukkink (2002) observed some students defining a target word like “dilemma” with relatively rich, tentative answers, such as “problem in choosing” or even “problem in choosing between two things” (p. 41). In the end, however, they selected only the word “problem” as a final answer.

Dependent measures. I used two measurement instruments to measure the dependent variable in this study. Table 3.3 includes information about the measurement instruments that were used in each phase of the intervention and the time needed to complete them.

The standardized test. The *Gates-MacGinitie Reading Tests* (GMRT), Fourth Edition, standardized norm-referenced tests, are designed to provide a general assessment of reading achievement. Arthur Gates published the Gates Silent Reading Test and the Gates Primary Reading Tests in 1926. They are the most widely used tests and, over the years, have been revised to reflect new concepts in reading and to establish a new national norm (Maria & Hughes, 2008).

Table 3.3 Test Forms and Times

Phase	# of Test	Test	Number of Questions	Testing Time in Minutes	Total of Testing Time in Minutes	Note
Before the intervention	2 pretests	Standardized test (GMRT)	93	55	100	These two tests were completed in two separate days.
		+	+	+		
		Experimenter-constructed test (ECT)	45	45		
Immediately following the intervention	1 immediate posttest	ECT	45	45	45	This test was completed in one day.
Long-term follow up tests (three weeks after the immediate posttest)	2 delayed posttests	GMRT	93	55	100	These two tests were completed in two separate days.
		+	+	+		
		ECT	45	45		
Total of Testing Time in Hours					4 hours and 5 minutes	

The GMRT have been used at the national level in the United States by school districts, classroom teachers, doctoral students, researchers, reading specialists, and in national studies sponsored by the U.S. Department of Education (Cook, Gerber, & Semmel, 1997; Drummond et al., 2011; Fisher, 2001; Gilbert, 2009; Johnson & McCabe, 2003; Nelson & Stage, 2007; Rowe, Ozuru, O’Reilly, & McNamara, 2008; Tilstra, McMaster, Broek, Kendeou, & Rapp, 2009). The current Fourth Edition contains the following grade levels: PR (Pre-Reading), BR (Beginning Reading), Levels 1 through 6, Level 7/9, Level 10/12, and AR (Adult Reading). Hence, the GMRT can assess mean levels ranging from Pre-Reading to Adult Reading. Levels 2 through AR have two

forms—Form S and Form T—for test and retest. Levels 3 through AR include two subtests each for Vocabulary and Comprehension.

The Vocabulary subtest assesses a student’s reading vocabulary. This subtest contains 45 questions, each consisting of a test word in a brief context followed by five other words or phrases. The student’s task is to choose one word or phrase that conveys the nearest meaning to the given word or phrase. The level of difficulty increases from beginning to end. The vocabulary test words are of general usefulness and not obscure or specialized words. Many vocabulary questions include one or more wrong answers. These wrong answers are of three different types: visual similarity, miscue, and association (MacGinitie, MacGinitie, Maria, Dreyer, & Hughes, 2008). The test is a measure of word knowledge, not being able to derive meaning from context. The brief context is not intended to provide clues to the meaning of the test word. The Vocabulary subtest is a timed 20-minute test in which students mark their answers on a test answer sheet.

The Comprehension subtest measures a student’s ability to read and understand different types of prose. This subtest consists of 11 passages of various lengths about diverse subjects, such as literature, social studies, natural sciences, and humanities. The passage type includes both narratives and expository texts (Lipson & Wixson, 1991; MacGinitie et al., 2008). The passages are selected from various authors and are not on very familiar topics or from popular books likely to have been read by many students (Lipson & Wixson, 1991; MacGinitie et al., 2008; Maria & Hughes, 2008). The student’s task is to read the passage and answer a number of multiple-choice questions. Each

comprehension question is presented with four choices. A total of 48 questions probed the student's understanding of the passages. The time given to complete the subtest is 35 minutes. Females and males of various ethnic groups are equally represented in the test content.

Both the Vocabulary and Comprehension subtests can be machine or hand scored. After scoring, each section of the test and a total score will receive a Normal Curve Equivalent (NCE), National Percentile Rank (PR), National Stanine (Stanine), Grade Equivalent (GE), and Extended Scale Score (ESS). The reading facilitator uses the total grade equivalent score to group students according to their reading levels (MacGinitie et al., 2008).

Validity and reliability of the Gates-MacGinitie Reading Tests. In establishing the national norm for the tests, about 65,000 students, studying in both public and private schools from all parts of the United States, were tested in the fall of 1998 and spring of 1999 for the Fourth Edition (MacGinitie et al., 2008). Johnson and McCabe (2003) reviewed the Fourth Edition of the GMRT. They pointed out that the GMRT showed strong total test and subtest internal consistency levels, ranging from 0.88 to 0.90. The significant statistic figures are listed as follows:

- Coefficient values were at or above 0.90 for all test materials.
- Alternate form correlations for the total tests were at or above 0.90.
- Alternate form correlations for the subtests ranged from 0.74 to 0.92.
- Total test coefficient values were at or above 0.88.
- Test-re-test reliability had been reported as above 0.88.

In the review of the Fourth Edition, Johnson and McCabe (2003) affirmed solid evidence for test validity. The researchers pointed out that the content validity of the GMRT is reinforced through a widespread test development process and that scores are reported to correlate well with the scores of comparable measures such as the Standard Achievement Test. The GMRT was also significantly correlated with the verbal or English sections in the Preliminary Scholastic Assessment Test (PSAT), Scholastic Assessment Tests (SAT), American College Testing Program (ACT), and grade point averages (GPAs; Lipson & Lang, 1991). The researchers also found significant correlations between the Third and Fourth Editions, ranging from 0.91 to 0.93, and the design of the two editions was found to be very similar (Johnson & McCabe, 2003).

In 2008, Rowe et al. also examined the GMRT. They determined the difficulty of various standardized reading tests currently used in the United States. They cross-examined the Level 7/9 and Level 10/12 of the GMRT and concluded that the GMRT contains a variety of passages with varying ranges of difficulty, differing in a number of dimensions. Furthermore, the researchers determined that the tests contain questions of several different types, most of which cannot be answered by merely eliminating distractors. The test extensively measures numerous different subcomponents implicit in the reading comprehension of the text in the context of various reading conditions.

How and why the Gates-MacGinitie Reading Test was used for this study? I used Level 4 of the Fourth Edition of the GMRT in the pretest as a general benchmark of students' vocabulary knowledge prior to the study and in the delayed posttest, given three weeks after the immediate posttest, to see if there was any growth in reading ability,

specifically in vocabulary. Consequently, the GMRT provided baseline data in vocabulary knowledge at the beginning of the study and established a baseline proficiency level for each student. The baseline level is the highest level at which a student can demonstrate proficiency without teacher help. This is the level at which a teacher can expect the student to perform at proficiency on high stakes testing. Therefore, the GMRT helped to show students where they were before the intervention, where they should be, and what skills and behaviors are in between.

Additionally, I used scores on the Gates–MacGinitie Reading Vocabulary (GMRV) subtest (MacGinitie et al., 2008) to group students according to their vocabulary knowledge. More specifically, students were classified as students with adequate or poor vocabulary knowledge for purposes of analysis. Poor vocabulary knowledge was defined as those students whose scores fell at or below the 25th percentile, whereas those who scored above the 25th percentile were designated as students with adequate vocabulary knowledge (Denton et al., 2015). A total of 55 minutes was allowed to complete both subtests of the GMRT. I used the following test materials to administer this standardized test in the research project:

- Two GMRT Reusable Test Booklets (Form S) Level 4
- Two GMRT Reusable Test Booklets (Form T) Level 4
- Two GMRT Machine-Scorable Answer Sheets (Forms S/T) Level 4
- One GMRT Directions for Administration (Form S and T) Levels 4–6

- One GMRT Manuals for Scoring and Interpretation (Forms S and T) Levels 4–6
- One GMRT Bundled Technical Report Supplement only

The experimenter-constructed test. I created a researcher-made test that is comprised of two segments (multiple-choice recognition and a meaning-recall) with three alternate or parallel forms (Form A, B, and C) for this study. The purpose of using different forms of the experimenter-constructed test (ECT) for the pretest, posttest, and delayed probes was to avoid what is called the *testing effect* and to obtain meaningful results (Endres & Renkl, 2015). In simple terms, these forms were designed to have similar measurement characteristics, but they contained different items. Therefore, items differed on each form, but each form was measuring the same thing (for more details, see the section titled *Validity and reliability of the experimenter-constructed test*). All student participants in the control groups and treatment groups completed the ECT three times in this study.

1. Before the intervention: I used Form A of the ECT in the pretest to evaluate student participants for possible pre-experimental differences in vocabulary ability and to use their scores on this test to serve as a covariate in posttest analyses. Forty-five question items were included, and the total time required to complete this test was 45 minutes.
2. Immediately following the intervention: After completing the vocabulary instructional intervention, all student participants were exposed to an immediate posttest, using Form B of the ECT, to compare the influence of the

vocabulary instructional intervention on the performance of students in the treatment groups with their peers in the control groups who did not receive the intervention. Forty-five question items were included, and the total time required to complete this test was 45 minutes.

3. Long-term follow up tests: This was an administration of Form C of the ECT three weeks after the immediate posttest to examine the students' retention of the learned strategy. Forty-five question items were included, and the total time required to complete this test was 45 minutes.

Validity and reliability of the experimenter-constructed test. Research shows that it is possible for researchers and teachers to create reliable and valid instruments to measure student vocabulary knowledge (Stahl & Bravo, 2010). In order to address the limitations of forced-choice measurement instruments found in the standardized tests, I created two instruments for this study. I pretested both experimenter-constructed instruments (multiple-choice recognition and meaning-recall) in a pilot study. Specifically, 21 participants in the pilot study completed both instruments prior to the start of the pilot study, 24 hours after completing the instructional intervention, and three weeks after the immediate posttest. The following subsections provide a detailed description of the steps followed to ensure the validity and reliability of each instrument included in the ECT.

Multiple-choice recognition instrument. I constructed a multiple-choice instrument with 30 question items to measure students' ability to use their knowledge to identify correct definitions for unfamiliar words presented in sentences or short

paragraphs. One possible example of the question items included in the multiple-choice recognition instrument was: “I think you should be skeptical about those claims. That is, you should doubt or question them first.” Thus, the stem for each item simply included the unfamiliar word (e.g., *skeptical*), signal words (e.g., *That is*), and the clues in context (e.g., *doubt or question*). Each question possessed four choices, one correct choice and three plausible distractors. I selected the answer choices based on length (number of words), relevance to the correct answer (as distractors), and language density (ease of reading).

To obtain meaningful results about students’ performance, I created different forms of the ECT for the pretest, posttest, and delayed probes. Research shows that using an identical test for the pre and posttests can cause what is sometimes called the *testing effect*, causing the assessment to measure familiarity instead of learning (Endres & Renkl, 2015). To illustrate, imagine that a student sits down to take the post-assessment, recognizes it from a week ago, and puts down the same answers as before. “I remember this,” he says to himself, “The answers made that funny word GUZZLEPOP!” The student shows no growth if he does that. Now imagine another student who remembered the questions from the pre-assessment and looked them all up when she arrived home and memorized the right answers in order. She shows an impressive amount of growth, but that does not mean she learned it all. Therefore, the pre and posttests had to be designed for this study in such a way that any change in students’ scores can be reasonably attributed to student learning and not to memory of the pre-assessment.

Different forms of a test are known as parallel forms or alternate forms. Alternate or parallel form reliability indicates how consistent test scores are likely to be if a person takes two or more forms of a test. In order to accomplish a high parallel form reliability, when designing the three alternate forms (Forms A, B, and C) for this study, I ensured that these forms measured the same learning outcomes at the same level of difficulty using similar assessment items and/or methods. This way scores can be compared to one another to show students' growth (Henchy, 2013). Following the completion of the pilot study, I performed an item analysis for each individual question item included in the three alternate forms for both experimenter-constructed instruments. I found a high parallel form reliability coefficient (at or above 0.88), suggesting that the three alternate forms are comparable.

The construction of the multiple-choice recognition instrument reflects best practice for multiple-choice item construction as detailed by Haladyna, Downing, and Rodriguez (2002). In my study, three experts in reading (one professor, one doctoral candidate, and one fourth-grade teacher with more than 19 years of teaching experience) reviewed each of the multiple-choice items for difficulty, clarity, and errors in content or grammar and provided comments for revision. A language specialist also reviewed each item for language consistency and appropriateness among the distractor items. I held individual conversations with the reviewers regarding their critique and ideas for improvement. I revised the multiple-choice instrument until the reviewers confirmed that their concerns had been addressed through a review of the final version of the instrument. I followed the same steps for the meaning-recall instrument.

I constructed a rubric of correct and acceptable answers for both experimenter-constructed instruments during discussion with the reviewers. I graded students' answers to the question items on both experimenter-constructed instruments according to the following system:

- On the multiple-choice recognition instrument, correct answers were given one point each.
- On the meaning-recall instrument, correct answers were given one point, and a word with a similar meaning was given a half point. For example, if the test word's correct answer was *novel* but the student wrote *book*, a half point was awarded. This way students' partial understanding of unfamiliar word meanings was measured.

At the beginning of conducting the pilot study, students were invited to participate in a research study to help evaluate a new method for finding the meaning of unfamiliar words during reading. I explained to student participants,

We would like to find out more about how using words surrounding unfamiliar words will help fourth-grade students with and without reading difficulties figure out the meaning of those unfamiliar words during reading. To accomplish this goal, it is necessary to take multiple tests to figure out which words students already know. Each test contains 30 multiple-choice items and 30 open-ended questions where you will be asked to provide a definition for each new word. You are not expected to know all of these terms. However, please do the very best that you can to figure out the meaning of those unknown words based on their context in sentences or short paragraphs. There is no penalty for incorrect answers.

I gave the directions for both instruments (the multiple-choice and meaning-recall) to students separately, as described here. Following the instructions, I gave the multiple-

choice instrument to students during a class period of their social studies course. In addition to the test form, I gave each student an answer sheet for recording responses.

I scored all multiple-choice items, and a research assistant (RA) entered the scores into a spreadsheet. The RA re-scored 10% of my work, using a copy of the answer key, to ensure fidelity. If the RA found mistakes, I re-scored all the tests and resubmitted them for review again by the RA. I reviewed 10% of the RA's work to ensure accuracy. If I found any mistakes, I reviewed 100% of the RA's work and fixed any errors.

To determine the internal consistency, I calculated Cronbach's alpha following the pretest, immediate posttest, and delayed posttest for the multiple-choice instrument. The alpha levels were .89, .90, and .89 respectively. An alpha level of .70 or higher is typically considered acceptable in social science research (Cronbach, 1951). Consequently, these alpha levels provide strong evidence of the reliability of the experimenter-constructed multiple-choice instrument used in the pilot study.

As noted by Bravo and Cervetti (2008), regardless of the reliability of the measures, the use of an experimenter-constructed multiple-choice instrument alone would not provide sufficient evidence to show that students moved from no knowledge to passive knowledge of vocabulary terms/concepts. The researchers illustrated that passive knowledge requires a demonstration of knowledge that goes beyond identification of a simple definition. As a result, a second assessment instrument was created in order to measure and confirm student learning.

Meaning-recall instrument. The second instrument was meaning-recall (i.e., open-ended). The main purpose of creating this instrument was to assess students' ability

to produce a definition for new, unfamiliar words in writing, as producing written responses is a typical requirement in the coursework of upper-elementary grades. Additionally, this instrument was intended to evaluate students' deeper knowledge of words (e.g., synonyms, antonyms) and contextual understanding based on contextual knowledge. Contextual knowledge is a word meaning derived from context, which can include a sentence, a passage, a discussion, or a picture (Lewis, 2009). In particular, the meaning-recall instrument asked students to write a word's meaning for each of the 30 unfamiliar words. Hence, this instrument required much more than simple matching, a form of vocabulary assessment that has been widely criticized (Stahl & Bravo, 2010).

Again, when implementing this instrument during the pilot study, students were encouraged to do the best that they could, despite the likelihood of not knowing the definition of all 30 terms. In fact, when taking the ECT, students completed the multiple-choice recognition instrument first, took a five-minute break while I collected student multiple-choice answer sheets, and then completed the meaning-recall instrument. This step was built in to prevent students' cognitive fatigue. Sievertsen, Gino, and Piovesan (2016) defined cognitive fatigue as "an increasingly common human condition that results from sustained cognitive engagement that taxes people's mental resources" (p. 2621). Considering students' cognitive fatigue also led me—during the pilot study—to conduct the pretest, immediate posttest, and delayed posttest early in the school day. Research finds that "as the day wears on, students become increasingly fatigued and consequently more likely to underperform on a test" (Sievertsen et al., 2016, p. 2621).

The RA and I independently scored student responses to the meaning-recall assessment, using a rubric of acceptable responses, and then compared scores. When scores did not match, the RA and I achieved 100% agreement for each item through conversation. Preliminary interscorer reliability was 94% with the RA; however, final interscorer reliability with the RA for all items was 100%.

I calculated Cronbach's alpha for the meaning-recall items following the pretest, immediate posttest, and delayed posttest to determine the quality of internal consistency for this measure. The alpha levels were .91, .92, and .91 respectively. An alpha level of .70 or higher is typically considered acceptable in social science research (Cronbach, 1951). Therefore, these alpha levels provided strong evidence of the reliability of the experimenter-constructed meaning-recall instrument used in this study.

Finally, based on feedback from the pilot study regarding students' results on both of the experimenter-constructed instruments (multiple-choice recognition and meaning recall), and considering the limited time that might be given for the formal study when conducting the study measurements in the participating schools, I decided to reduce the number of question items in both instruments using the item-by item analysis method. This analysis method is a widely used and broadly applied statistical technique in the social and behavioral sciences (Izquierdo, Olea, & Abad, 2014).

Using this analysis method, I identified and removed the questions that were found to be "too difficult," meaning items that were either skipped or not answered correctly by many of the students. More specifically, I removed all question items with high percentages of students (35% or more) missing them. Consequently, items included

in the three alternate forms for both experimenter-constructed instruments followed a normal distribution of difficulty surrounding the desired mean. This was calculated by trial and error using a spreadsheet program, in Microsoft Excel. The final version of the three parallel forms for both experimenter-constructed instruments contains a 20-item production test (i.e., write a word's meaning) and a 25-item, 4-option, multiple-choice test. As a result, 45 question items were included in the experimenter-constructed test and were used in this study, and the total time required to complete it was 45 minutes. Examples of the question items included in both experimenter-constructed instruments are available in Appendix C.

Fidelity of implementation. The application of an intervention as it is designed is a critical factor in order to maximize intervention benefits. The term used to describe this concept is fidelity of implementation (FOI), which is “the delivery of instruction in the way in which it was designed to be delivered” (North Dakota Department of Public Instruction, 2010, p. 10). The FOI concept has received increased attention in recent years because the findings of numerous research showed that schools with “high levels of implementation [and] ... uniformity of high implementation across program components” did experience improvements in achievement, especially in the areas of math and reading (Aladjem & Borman, 2006, p. 3).

Therefore, I created a fidelity checklist of critical elements of the intervention for this study to assess implementation of the vocabulary instructional intervention during the instructional intervention lessons. The teacher fidelity checklist contained 20 items, representing specified teacher behaviors, modeling, and cueing throughout the

instructional lessons. Items on the checklist also included the format of the lesson to make sure the teacher began with teacher-led practice (I Do), guided practice (We Do), and ended with independent practice (You Do). The teacher fidelity checklist included items such as, “Teacher reads the practice sentences or paragraphs aloud and then demonstrates examples of the thinking process that good and skilled readers use to understand how using context clues can improve comprehension.” The teacher fidelity checklist is available in Appendix D.

Several researchers recommended that treatment fidelity be obtained for a minimum of 25% of all intervention sessions (Kratochwill et al., 2013; O’Neill, McDonnell, Billingsley, & Jenson, 2011). As a result, in this study, two trained graduate students—who were not involved with the study—assessed treatment fidelity using videotapes during 39% of the instructional sessions with inter-observer agreement (IOA) data that were calculated for all 39% of these sessions. Regarding the standards for IOA, O’Neill et al. (2011) asserted that researchers should collect IOA data during 20% - 30% of the observations or intervention sessions. The two graduate students calculated IOA on an item-by-item basis by dividing each step with agreement by the total number of steps with agreements plus disagreements and multiplying by 100 (Cooper, Heron, & Heward, 2007). According to O’Neill et al. (2011), the conventional minimum standard for acceptable IOA is a mean of 80% to 85% across all observations.

I trained two graduate students (master’s level) on how to assess treatment fidelity using both the videotapes and the teacher fidelity checklist. Specifically, I trained those graduate students by reviewing the fidelity checklist then having them watch a

videotaped lesson demonstration and complete the fidelity checklist simultaneously with me. If any one of the graduate students did not meet the 100% criterion, additional lesson demonstrations were planned with me until s/he met 100% criterion. However, both graduate students met 100% criterion and no additional lesson demonstrations were needed.

Inter-scorer reliability. The two graduate students also conducted the inter-scorer reliability of all assessments that were used in this study. Independently, these individuals scored 25% of all tests that were distributed evenly across conditions and participants. The two scorers used a scoring sheet I created for this study. Scored tests were compared item-by-item to determine the number of agreements and disagreements. The number of agreements were divided by the number of agreements plus disagreements and multiplied by 100% to calculate the percentage of agreement. The Council for Exceptional Children (CEC, 2014) recommended 80% agreement as the minimum acceptable inter-scorer agreement for evidence-based practices in special education. The same two graduate students, who received training for assessing treatment fidelity, also received training on how to score each measurement instrument (GMRT and ECT) used in this study. First, I provided the graduate students with examples of completed assessments and reviewed how to score each test. For the standardized test (GMRT), I provided directions on how the graduate students should use the GMRT examiner manual to score tests. Graduate students scored raw score data for each assessment. If accuracy was less than 100% for any of the assessments scored, I conducted retraining by having

the graduate student score additional assessments until s/he met 100% criterion. I provided a booster session, twice, to a graduate student who scored below 100% for assessment administration.

CHAPTER IV

RESULTS

The purpose of this study was to examine the influence of vocabulary instruction that is based on a combination of a strategy and certain types of context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. A total of 59 urban fourth-grade students participated in the study. To evaluate students for possible pre-experimental differences in vocabulary ability and to classify them into two groups (those with adequate and poor vocabulary knowledge), all student participants were exposed to the standardized reading test (GMRT, Fourth Edition) as a pretest prior to the intervention. More specifically, the GMRV subtest scores were used to classify students into two groups for purpose of analysis: those with adequate or poor vocabulary knowledge.

Using the converted National Percentile Ranks (NPRs) scores of the GMRV pretest scores, poor vocabulary knowledge (PVK) was defined as scores that fell at or below the 25th percentile ($n = 25$), while scores above the 25th percentile were designated as adequate vocabulary knowledge (AVK; $n = 34$). The students with adequate (AVK) and poor vocabulary knowledge (PVK) were further classified into two groups based on instructional conditions (treatment and control conditions; see Figure 3.1). The students were not randomly assigned to the instructional conditions. Student participants were assigned in closely matched pairs to one of the two designed

instructional conditions (treatment and control conditions). A two-way ANOVA method was conducted to verify the equivalency between AVK in the control and treatment conditions and between PVK in the control and treatment conditions (i.e., ensuring that they produce the same results), based on the converted Extended Scale Scores (ESSs) of the GMRV pretest scores. Table 4.1 includes information about the number of students in both groups (PVK and AVK) across both instructional conditions (treatment and control conditions) as well as the number of students with LD and TD in each group.

Table 4.1 Participants Table

Groups	Treatment Condition		
	# of Students with LD	# of Students with TD	Total Students
Group 1 Students with PVK	6	4	10
Group 2 Students with AVK	8	9	17
	Control Condition		
Group 3 Student with PVK	9	6	15
Group 4 Student with AVK	2	15	17
Total Students			59
Note: LD = Students with learning disabilities; TD = Students with typical development; PVK = Students with poor vocabulary knowledge; AVK = Students with adequate vocabulary knowledge			

To determine if the effect of an instructional condition on ESS pretest scores depend on whether a student has PVK or AVK, first it must be determined whether there is a statistically significant interaction effect (between the two independent variables: Condition and Group) in a two-way ANOVA. Prior to running the ANOVA, preliminary analyses were performed to test for outliers as well as the assumptions of normality and homogeneity of variances of the two-way ANOVA. Outliers were assessed visually by an inspection of boxplots. Normality of the distributions was assessed using Shapiro-Wilk's normality test for each combination of instructional condition by vocabulary knowledge. Test of the assumption of homogeneity of variances was assessed by Levene's test. The tests' results showed that there were no outliers, the residuals were normally distributed ($p > .05$), and the assumption of homogeneity of variances ($p = .605 > .05$) was not violated.

Results of the two-way ANOVA showed that the interaction effect between Condition (treatment and control conditions) and Group (PVK and AVK) for ESS pretest scores was not statistically significant, $F(1, 55) = .28, p = .601 > .05$, partial $\eta^2 = .005$. These results confirmed that the differences between AVK and PVK were not dependent on the instructional conditions prior to the intervention. Furthermore, the main effect of instructional conditions (treatment and control conditions) on ESS pretest scores was not statistically significant, $F(1, 55) = 0.09, p = .772 > .05$, partial $\eta^2 = .002$. As expected, there was a statistically significant main effect of Group (PVK and AVK), $F(1, 55) = 83.34, p < .001$, partial $\eta^2 = .602$. Table 4.2 presents the results of the two-way ANOVA.

Table 4.2 Results of the Two-Way ANOVA for the Effects of Condition and Group at the ESS Pretest Scores

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2
Condition	1	46.884	.085	.772	.002
Group	1	45971.462	83.341	.000	.602
Condition x Group	1	152.484	.276	.601	.005
Error	55	551.609			
Total	59				

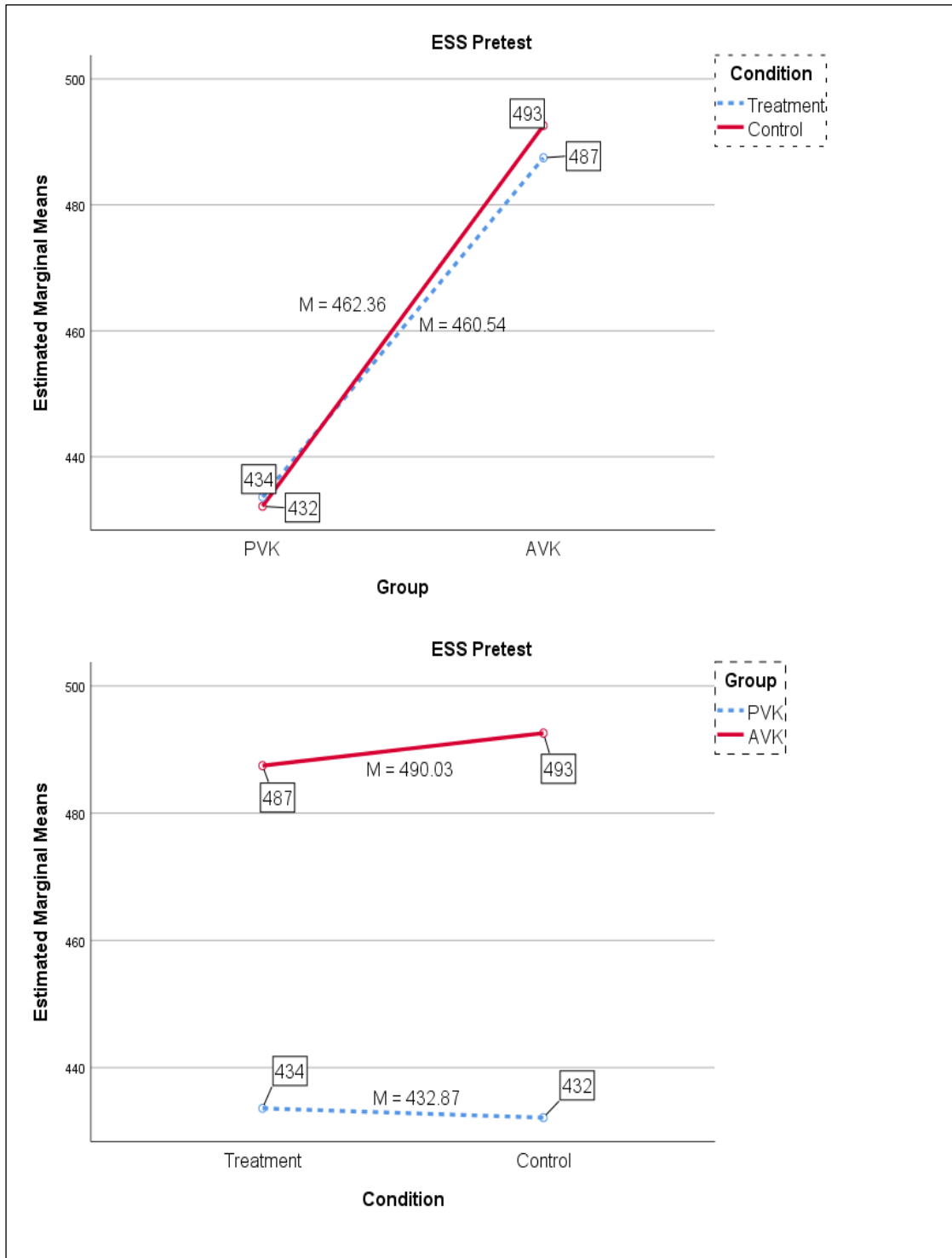
The means of ESS pretest scores for students with PVK and AVK were 432.87 ($SD = 4.79$) and 490.03 ($SD = 4.03$), respectively (see Table 4.3). These results indicate that students with AVK scored an average of 57.16 points higher than students with PVK, a statistically significant difference, $p < .001$. In addition, the means of ESS pretest scores for treatment and control conditions were 460.54 ($SD = 4.68$) and 462.36 ($SD = 4.16$), respectively (see Table 4.3). These results indicated that students in the control condition scored an average of 1.82 points higher than students in the treatment condition. This difference ($p = .772 > .05$) was not statistically significant. The results of the means of ESS pretest scores for Group (PVK and AVK) and Condition (treatment and control conditions) are well visualized in a graph (see Figure 4.1).

Table 4.3 The Unweighted Marginal Means of the ESS Pretest Scores for Condition and Group

		Condition		
Group		Treatment	Control	Marginal Means
	PVK	433.60 (<i>n</i> = 10)	432.13 (<i>n</i> = 15)	432.87 (<i>n</i> = 25)
	AVK	487.47 (<i>n</i> = 17)	492.59 (<i>n</i> = 17)	490.03 (<i>n</i> = 34)
	Marginal Means	460.54 (<i>n</i> = 27)	462.36 (<i>n</i> = 32)	461.45 (<i>n</i> = 59)

Since two measurement instruments (GMRT and ECT) were used to measure the dependent variable (students' vocabulary acquisition) in this study, and because those measures are related to different research questions, the results of the present study are presented separately with respect to the major research questions and subquestions. In addressing major research question 1, the emphasis is placed on the impact of the vocabulary instructional intervention on the performance of the students on the ECT measure. Similarly, the second major research question focuses on the effect of the vocabulary instructional intervention on the performance of the students on the GMRT measure. The third major research question examines whether the students' results on both measures of vocabulary knowledge (GMRT and ECT) were consistent across the two instructional conditions (treatment and control conditions). Interobserver agreement

Figure 4.1. Results of the Means of ESS Pretest Scores for Condition and Group



is presented to verify the treatment fidelity based on the teacher fidelity checklist (see Appendix D). Finally, inter-scorer reliability is presented to show the degree to which all assessments were scored accurately and consistently across conditions and participants.

Research Question 1

The first research question of this study is: Does the vocabulary instructional intervention have any effect on the performance of the students on the experimenter-constructed test? If so, for which group of students (students with PVK and AVK) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (1A–1D).

Subquestion 1A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) overtime? I hypothesize that students in the treatment condition will perform better than students in the control condition on the experimenter-constructed test.

A three-way mixed ANOVA method was conducted to answer this research question and related subquestions. To run a three-way mixed ANOVA, seven assumptions were examined. The first three assumptions relate to the chosen study design: (a) there should be one dependent variable, which was the ECT scores; (b) there should be two between-subjects factors, which were Condition (treatment and control conditions) and Group (students with PVK and AVK); and (c) there should be one within-subjects factor, which was measurement occasions represented in this study by the

variable Time. The ECT scores were collected on three occasions, or time points, namely on the vocabulary pretest, immediate posttest, and delayed posttest. To ensure the internal and statistical validity of the analyses, four assumptions reflecting the nature of the data needed to be met. These were: (a) there should be no significant outliers in any cell of the design; (b) the dependent variable should be approximately normally distributed in every cell of the design; (c) the data does not violate the assumption of homogeneity of variances; and (d) the data meets the assumption of sphericity.

Through examination of these assumptions, it was found that the ECT scores were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there were no outliers in the data, as assessed by an inspection of boxplots for values greater than 1.5 box-lengths from the edge of the box. There was no violation of the assumption of homogeneity of variances for both ECT pretest scores ($p = .246 > .05$) and ECT delayed posttest scores ($p = .215 > .05$). However, this was not true for the ECT immediate posttest scores ($p = .003 < .05$), as assessed by Levene's test for equality of variances. Mauchly's test of sphericity revealed that the assumption of sphericity had not been violated, $\chi^2(2) = 2.91, p = .234$. Accepting the assumption of sphericity indicated that the three-way mixed ANOVA was not biased, and no adjustment to the test was needed.

Results of the three-way mixed ANOVA showed that the three-way interaction between Time, Condition, and Group was not statistically significant, $F(2, 110) = .11, p = .898 > .05$, partial $\eta^2 = .002$. This means that the two-way interactions of Condition x Group were not different at the different points of Time (i.e., not different at the three time points). Partial eta-squared (η^2), which is a measure of effect size (Cohen, 1988),

describes the “proportion of total variation attributable to the factor, partialling out (excluding) other factors from the total non-error variation” (Pierce, Block, & Aguinis, 2004, p. 918). Barely two-tenths of one percent of the variance in this model can be attributed to the interaction between Time, Condition, and Group after partialling out the effects of the other factors (Condition and Group) from the explained variance. A partial η^2 measurement of .002 indicated a negligible effect size for the three-way interaction.

However, the results showed that there was a statistically significant two-way interaction between Time and Condition, $F(2, 110) = 25.27, p < .001$. All other two-way interactions (Time x Group and Condition x Group) were not statistically significant ($p > .05$). A follow-up test was only carried out for the statistically significant two-way interaction between Time and Condition. A test for the simple main effects of Condition was performed at each of the three time points. Statistical significance of a main effect was accepted at a Bonferroni-adjusted alpha level of .0167. Results revealed that there was a statistically significant simple main effect of Condition at the immediate posttest level, $F(1, 55) = 35.17, p < .001$, and at the delayed posttest level, $F(1, 55) = 22.42, p < .001$, as expected, however, not at the pretest level, $F(1, 55) = 1.23, p = .273 > .05$.

Then, all pairwise comparisons were performed for statistically significant simple main effects. Bonferroni corrections were made for comparisons within each simple main effect considered a family of comparisons. Adjusted p -values were reported. Results showed that the mean ECT score was significantly higher in the immediate posttest for students in the treatment condition than students in the control condition, a mean difference of 8.25 points, 95% CI [5.46, 11.04], $p < .001$, with a very large effect size (d

= 1.30). Additionally, the mean ECT score was significantly higher in the delayed posttest for students in the treatment condition than students in the control condition, a mean difference of 7.75 points, 95% CI [4.47, 11.03], $p < .001$, with a large effect size ($d = 1.11$). Table 4.4 shows the results of the pairwise comparisons at each of the three points of Time (ECT pretest, ECT immediate posttest, ECT delayed posttest). Because there were statistically significant mean differences in the ECT scores at the immediate and delayed posttests across both conditions (treatment and control conditions), it can be concluded that students in the treatment condition performed better than students in the control condition on the ECT measure, which positively answers subquestion 1A and supports the hypothesis that students in the treatment condition significantly outperformed students in the control condition on the ECT.

Table 4.4 Pairwise Comparisons for the ECT Scores by Condition x Time

Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
ECT Pretest	Treatment	Control	1.823	1.645	.273	-1.474	5.119	.239
ECT Immediate Posttest	Treatment	Control	8.251	1.391	.000	5.463	11.039	1.296
ECT Delayed Posttest	Treatment	Control	7.747	1.636	.000	4.468	11.025	1.107

Subquestion 1B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the experimenter-constructed test.

For this subquestion, the purpose is to investigate the effect of Condition on the ECT scores for the PVK group only. To answer this subquestion, first it must be determined if there are simple main effects. In an ordinary two-way ANOVA, simple main effects are the effects of one of the between-subjects factors at all levels of the other between-subjects factor (Laerd Statistics, 2015). For example, the effect of Condition at each level of Group (i.e., the effect of Condition for students classified as PVK and the effect of Condition for students classified as AVK).

Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025. The result was a statistically significant simple main effect of Condition for students with PVK on their ECT immediate posttest scores, $F(1, 55) = 27.97, p < .001$. In addition, there was a statistically significant main effect of Condition for students with PVK on their ECT delayed posttest scores, $F(1, 55) = 16.96, p < .001$. These results provided evidence to support the hypothesis that the instructional conditions (treatment and control conditions) *do* lead to different mean ECT immediate and delayed posttest scores for students with PVK.

As with simple two-way interactions, statistically significant simple main effects need to be followed up with simple comparisons. In an ordinary two-way ANOVA,

simple comparisons are tests of the differences between individual group means within a simple main effect (Laerd Statistics, 2015). For example, the difference between the means of the treatment and control conditions for students with PVK at the immediate posttest scores. From the aforementioned findings, the simple main effect of Condition was statistically significant for students with PVK on their ECT immediate and delayed posttest scores. Therefore, the analyses were focused first on analyzing the simple comparisons of the ECT immediate posttest scores between students with PVK in the treatment condition and those students with PVK in the control condition. The results of this comparison are shown in Table 4.5.

Table 4.5 Pairwise Comparisons for the ECT Immediate and Delayed Posttest Scores by Condition for the PVK Group

Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
ECT Immediate Posttest	Treatment	Control	11.267	2.130	.000	6.997	15.536	1.770
ECT Delayed Posttest	Treatment	Control	10.317	2.505	.000	5.296	15.337	1.474

In Table 4.5, Bonferroni corrections were made with all pairwise comparisons within each simple main effect considered a family of comparisons. Adjusted *p*-values were reported. Table 4.5 shows that the mean ECT immediate posttest score (i.e., the “Mean Difference (I-J)” column) was 11.27 points higher for students with PVK in the

treatment condition compared to those students with PVK in the control condition. As shown in Table 4.6, the mean ECT immediate posttest score was significantly higher for students with PVK in the treatment condition ($M = 26.90$, $SD = 7.92$) than those students with PVK in the control condition ($M = 15.63$, $SD = 5.16$), a mean difference of 11.27 points, 95% CI [6.99, 15.54], $p < .001$, with a very large effect size ($d = 1.77$).

When considering the differences in the ECT delayed posttest scores between students with PVK in the treatment condition and those in the control condition, the results of this comparison indicated that the mean ECT delayed posttest score was significantly higher for students with PVK in the treatment condition ($M = 24.65$, $SD = 8.83$) than those students with PVK in the control condition ($M = 14.33$, $SD = 5.28$), a mean difference of 10.32 points, 95% CI [5.30, 15.34], $p < .001$, with a very large effect size ($d = 1.47$; see Tables 4.5 and 4.6). These findings answer subquestion 1B and confirm the hypothesis that students with PVK in the treatment condition significantly outperformed students with PVK in the control condition on the experimenter-constructed test.

Table 4.6 Descriptive Statistics for the ECT Scores by Condition x Group at the Immediate and Delayed Posttests

Time	Condition	Group	<i>N</i>	<i>M</i>	<i>SD</i>
ECT Immediate Posttest	Treatment	PVK	10	26.9000	7.91903
		AVK	17	30.6176	5.20975
		Total	27	29.2407	6.46198
	Control	PVK	15	15.6333	5.15983
		AVK	17	25.3824	2.80919
		Total	32	20.8125	6.36618
ECT Delayed Posttest	Treatment	PVK	10	24.6500	8.82877
		AVK	17	29.2353	6.14694
		Total	27	27.5370	7.43811
	Control	PVK	15	14.3333	5.28024
		AVK	17	24.0588	4.83762
		Total	32	19.5000	6.99885

Subquestion 1C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the experimenter-constructed test.

For this research question, the interest is to investigate the effect of Condition for the AVK group only. Therefore, the same steps that were followed to analyze the simple comparisons of both ECT immediate and delayed posttest scores between students with

PVK in the treatment condition and students with PVK in the control condition were also applied to analyze the simple comparisons for students with AVK.

Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025. The result was a statistically significant simple main effect of Condition for students with AVK on their ECT immediate posttest scores, $F(1, 55) = 8.56, p = .005 < .05$. In addition, there was a statistically significant main effect of Condition for students with AVK on their ECT delayed posttest scores, $F(1, 55) = 6.05, p = .017 < .05$. These results provided evidence to support the hypothesis that the instructional conditions (treatment and control conditions) *do* lead to different mean ECT immediate and delayed posttest scores for students with AVK.

As with simple two-way interactions, statistically significant simple main effects need to be followed with simple comparisons. From the aforementioned findings, the simple main effect of Condition was statistically significant for students with AVK on their ECT immediate and delayed posttest scores. Therefore, the analyses were focused first on analyzing the simple comparisons of the ECT immediate posttest scores between students with AVK in the treatment condition and those students with AVK in the control condition. The results of this comparison are shown in Table 4.7.

Table 4.7 Pairwise Comparisons for the ECT Immediate and Delayed Posttest Scores by Condition for the AVK Group

Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
ECT Immediate Posttest	Treatment	Control	5.235	1.790	.005	1.648	8.822	.822
ECT Delayed Posttest	Treatment	Control	5.176	2.105	.017	.959	9.394	.740

In Table 4.7, Bonferroni corrections were made with comparisons within each simple main effect considered a family of comparisons. Adjusted *p*-values were reported. Results indicated that the mean ECT immediate posttest score was significantly higher for students with AVK in the treatment condition ($M = 30.62$, $SD = 5.21$) than those students with AVK in the control condition ($M = 25.38$, $SD = 2.81$), a mean difference of 5.24 points, 95% CI [1.65, 8.82], $p = .005 < .05$, with a large effect size ($d = .82$; see Tables 4.6 and 4.7). Furthermore, the mean ECT delayed posttest score was significantly higher for students with AVK in the treatment condition ($M = 29.24$, $SD = 6.15$) than those students with AVK in the control condition ($M = 24.06$, $SD = 4.84$), a mean difference of 5.18 points, 95% CI [.96, 9.39], $p = .017 < .05$, with a medium effect size ($d = .74$; see Tables 4.6 and 4.7). These results confirmed the hypothesis that, when receiving the vocabulary instructional intervention, students with AVK in the treatment

condition will perform better than students with AVK in the control condition on the experimenter-constructed test.

Subquestion 1D

Is there a difference in retention between short-term learning and long-term learning (Time 2 vs. Time 3) for students in the treatment condition? I hypothesize that the performance of the students will be better in short-term learning (Time 2) in the treatment condition on the experimenter-constructed test.

This subquestion focuses on the extent to which the students with PVK and AVK, in the treatment condition, retain the learned context clue strategy three weeks post-intervention. In order to answer this subquestion, a two-way mixed ANOVA method was conducted. The primary purpose for running the two-way mixed ANOVA is to test whether there is an interaction between the between-subjects factor (Group) that is categorical with two categories (PVK and AVK) and within-subjects factor (Time) that is categorical with two points (immediate posttest and delayed posttest) on the dependent variable (the ECT scores). To run a two-way mixed ANOVA, there were eight assumptions that were examined. Seven of these assumptions are similar to the assumptions of the three-way mixed ANOVA design that were examined in research question 1, and the additional assumption is that the homogeneity of covariances has not been violated.

Upon examination of these assumptions, it was found that the ECT scores were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there were no outliers in the data, as assessed by an inspection of boxplots. There was no violation of

the assumption of homogeneity of variances for the ECT immediate and delayed posttests scores ($p = .093$ and $p = .244$, respectively), as assessed by Levene's test of homogeneity of variance ($p > .05$). However, there was homogeneity of covariances ($p > .001$), as assessed by Box's test of equality of covariance matrices ($p = .341$). Mauchly's test of sphericity showed that the assumption of sphericity was not violated. No major violation of the assumptions was revealed in this analysis. The assumption of sphericity indicated that the two-way mixed ANOVA was not biased, and no adjustment to the test was needed.

The results of the two-way mixed ANOVA revealed that there was no statistically significant interaction between Group (PVK and AVK) and Time on the ECT scores, $F(1, 25) = .50, p = .485 > .05$, partial $\eta^2 = .020$. This means that the differences in ECT scores for PVK and AVK did not change significantly from Time 2 to Time 3 (i.e., whatever difference existed between the two groups at Time 2, that difference did not change at Time 3). Only 2% of the variance in this model can be attributed to the interaction between Group and Time. A partial η^2 measurement of .020 indicates a small interaction effect of group differences over the two time points. Because the two-way interaction was not statistically significant and the main effects of Time and Group were significant, follow-up tests were performed to probe further the main effect of Time and the main effect of Group.

Results of the main effect of Time showed a statistically significant difference in the mean ECT score at the two time points for students in the treatment condition, $F(1, 25) = 8.80, p = .007 < .05$, partial $\eta^2 = .260$. Post hoc analysis with a Bonferroni

adjustment were performed for statistically significant main effect of Time. Results revealed that the mean ECT score decreased from immediate posttest ($M = 29.24$, $SD = 6.46$) to delayed posttest ($M = 27.54$, $SD = 7.44$), a statistically significant decrease of 1.82 points on average, 95% CI [.56, 3.08], $p = .007 < .05$, with a small effect size ($d = .24$). Tables 4.8 and 4.9 contain information relating to the results of the pairwise comparisons and descriptive statistics for the ECT scores resulting from the two-way mixed ANOVA. These findings showed that there was a small difference in retention between short-term learning and long-term learning for students in the treatment condition, as students performed better in short-term learning (the immediate posttest) compared to their performance in the long-term learning (the delayed posttest).

Table 4.8 Pairwise Comparisons for the ECT Scores from Immediate Posttest to Delayed Posttest

(I) Time Time	(J)	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
					Lower Bound	Upper Bound	
ECT Immediate Posttest	ECT Delayed Posttest	1.816	.612	.007	.556	3.077	.244

Table 4.9 Descriptive Statistics for the ECT Scores by Group for the Immediate and Delayed Posttests

Time	Group	<i>N</i>	<i>M</i>	<i>SD</i>
ECT Immediate Posttest	PVK	10	26.9000	7.91903
	AVK	17	30.6176	5.20975
	Total	27	29.2407	6.46198
ECT Delayed Posttest	PVK	10	24.6500	8.82877
	AVK	17	29.2353	6.14694
	Total	27	27.5370	7.43811

Results of the main effect of Group showed that there was no statistically significant difference in the mean ECT score between intervention groups (PVK and AVK), $F(1, 25) = 2.48, p = .128, \text{partial } \eta^2 = .090$. When comparing the mean ECT score of the delayed posttest for students with PVK in the treatment condition with their mean ECT score on the immediate posttest, the results showed that their mean ECT score was somewhat higher on the immediate posttest ($M = 26.90, SD = 7.92$) compared to their mean ECT score on the delayed posttest ($M = 24.65, SD = 8.83$). Similar results were found when comparing the mean ECT score of the delayed posttest for students with AVK in the treatment condition with their mean ECT score on the immediate posttest. Results showed that students with AVK performed better on the immediate posttest ($M = 30.62, SD = 5.21$) compared to their performance on the delayed posttest ($M = 29.24, SD = 6.15$). These changes in retention, between short-term learning (i.e., the immediate posttest, Time 2) and long-term learning (i.e., the delayed posttest, Time 3) for both

groups of students (PVK and AVK) in the treatment condition, are well visualized in a graph (see Figure 4.2).

Figure 4.2. Changes in the Students' Retention as a Function of Time and Group

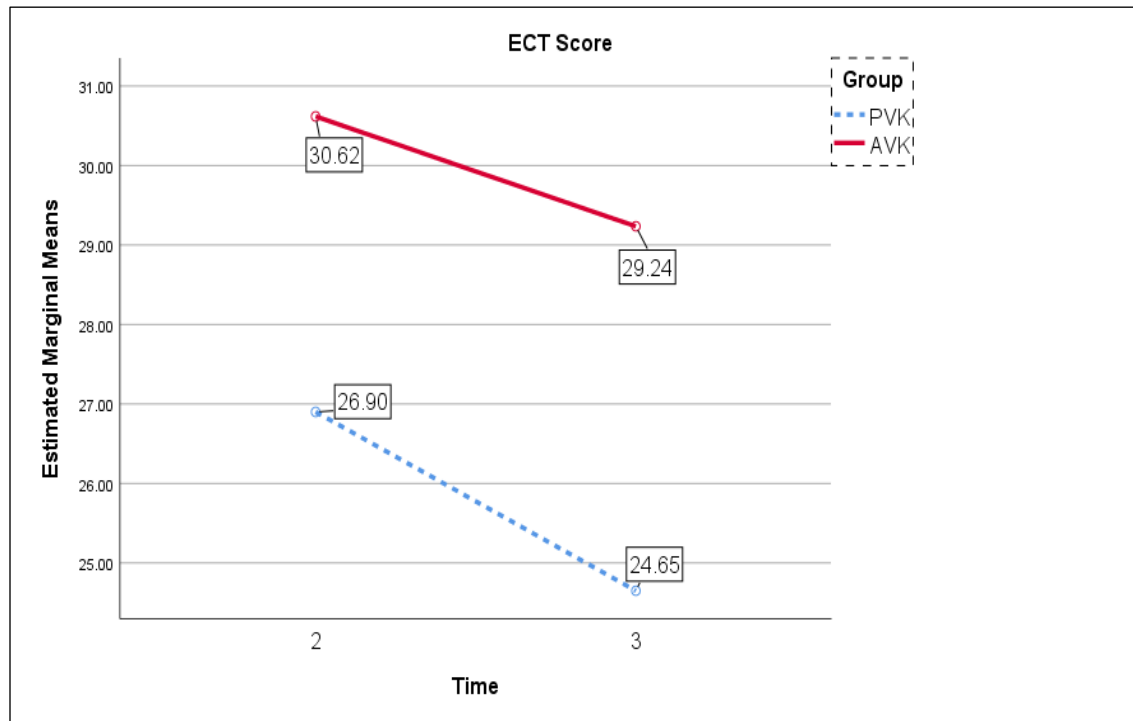


Figure 4.2 shows that there was a small difference in retention between short-term learning (Time 2) and long-term learning (Time 3) for both groups of students (PVK and AVK) in the treatment condition, as both groups of students performed better in short-term learning compared to their performance in the long-term learning. These findings positively answered subquestion 1D and confirmed the hypothesis that the performance of the students in the treatment condition on Time 2 significantly outperformed their performance on Time 3 using the experimenter-constructed test. Additionally, the results

revealed that students with AVK in the treatment condition outperformed their peers with PVK both times, Times 2 and 3. The students with AVK also performed better in short-term learning compared to their performance in the long-term learning.

Research Question 2

The second major research question of this study was: Does the vocabulary instructional intervention have any effect on the performance of the students on the standardized test (Gates-MacGinitie Reading Tests [GMRT])? If so, for which group of students (students with PVK and students with AVK) does the intervention have a significant effect?

This major question was addressed by answering a few subquestions (2A–2C).

Subquestion 2A

Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) at the pretest and delayed posttest? I hypothesize that students in the treatment condition will perform better than students in the control condition on the standardized test.

A three-way mixed ANOVA method was conducted to answer this major research question and associated subquestions. Upon examination of the seven assumptions of the three-way mixed ANOVA, it was found that there were no outliers in the data, as assessed by an inspection of boxplots. The GMRT scores were normally distributed, except for one value ($p = .034$), as assessed by Shapiro-Wilk's test ($p > .05$). This value was accepted because ANOVAs are considered to be fairly *robust* to deviations from normality (Laerd Statistics, 2015), which means that including the outlier ($p = .034$) in

the data analysis would not substantially affect the results. There was homogeneity of variances for the GMRT pretest scores ($p = .102$) but not for the GMRT delayed posttest scores ($p = .026$), as assessed by Levene's test for equality of variances. Because there were only two levels of the within-subjects factor (GMRT pretest and GMRT delayed posttest) and, therefore, only one paired difference, the assumption of sphericity was automatically met. Accepting the assumption of sphericity indicated that the three-way mixed ANOVA was not biased, and no adjustment to the test was needed.

Results of the three-way mixed ANOVA showed that the three-way interaction between Time, Condition, and Group was not statistically significant, $F(1, 55) = 1.40, p = .242 > .05$, partial $\eta^2 = .025$. This means that the simple two-way interactions of Condition x Group were not different at the different points of Time (i.e., not different at the two time points). Only 2.5% of the variance in this model can be attributed to the interaction between Time, Condition, and Group. As partial η^2 can be used in ANOVA as an estimate of the effect size, a partial η^2 measurement of .025 indicates a small effect size for the three-way interaction.

However, results showed that there was a statistically significant two-way interaction between Time and Condition, $F(1, 55) = 14.18, p < .001$, and between Time and Group, $F(1, 55) = 4.27, p = .044 < .05$. There was no statistically significant two-way interaction between Condition and Group, $F(1, 55) = 1.20, p = .278 > .05$. A follow-up test was only conducted for the statistically significant two-way interaction between Time and Condition. A test for the main effects of Condition was performed at the two time points (GMRT pretest and GMRT delayed posttest). Statistical significance of a main

effect was accepted at a Bonferroni-adjusted alpha level of .025. Results revealed that there was a statistically significant main effect of Condition at the delayed posttest level, $F(1, 55) = 10.07, p = .002 < .05$, but, as expected, not at the pretest level, $F(1, 55) = .08, p = .779 > .05$. The main effect of Condition was not statistically significant at the pretest level, which means that the mean GMRT score was not different in the pretest for students in the treatment and control conditions. Table 4.10 presents the results of the univariate tests for the GMRT scores by Condition at the pretest and delayed posttest.

Table 4.10 Results of the Univariate Tests for the GMRT Scores by Condition at the Pretest and Delayed Posttest

Time		<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2
GMRT Pretest	Contrast	2.110	1	2.110	.079	.779	.001
	Error	1465.833	55	26.652			
GMRT Delayed Posttest	Contrast	485.842	1	485.842	10.071	.002	.155
	Error	2653.412	55	48.244			

Note: Each F tests the simple effects of Condition within each level combination of the other effects shown. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

Then, all pairwise comparisons were performed for statistically significant main effects. Bonferroni corrections were made for comparisons within each main effect considered a family of comparisons. Adjusted p -values were reported. Results showed that the mean GMRT score was significantly higher in the delayed posttest for students in the treatment condition than students in the control condition, a mean difference of 5.88 points, 95% CI [2.17, 9.59], $p = .002 < .05$, with a medium effect size ($d = .64$). As a result, it can be concluded that students in the treatment condition performed better than students in the control condition on the standardized test (GMRT) after receiving the

vocabulary instructional intervention, thus positively answering subquestion 2A and confirming the hypothesis that students in the treatment condition significantly outperformed students in the control condition on the standardized test. Table 4.11 shows the results of these pairwise comparisons.

Table 4.11 Pairwise Comparisons for the GMRT Scores by Condition at the Pretests and Delayed Posttests

Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
GMRT Pretest	Treatment	Control	-.387	1.376	.779	-3.146	2.371	.042
GMRT Delayed Posttest	Treatment	Control	5.876	1.852	.002	2.165	9.588	.641

Subquestion 2B

Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the standardized test.

This subquestion focuses on investigating the effect of Condition on the GMRT scores for the PVK group only. In order to answer this subquestion, first it must be determined if there are simple main effects. Hence, the effect of Condition was

investigated at each level of Group (i.e., the effect of Condition for students with PVK and the effect of Condition for students with AVK).

Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025. Results indicated that there were no statistically significant simple main effect of Condition for students with PVK on their GMRT pretest scores, $F(1, 55) = .01, p = .937 > .05$. In addition, as expected, there was a statistically significant main effect of Condition for students with PVK on their GMRT delayed posttest scores, $F(1, 55) = 8.78, p = .005 < .05$. These results provided evidence to support the hypothesis that the instructional conditions (treatment and control conditions) *do not* lead to different mean GMRT pretest scores for students with PVK. This means that students with PVK in the treatment and control conditions did not differ on the GMRT pretest. Furthermore, these results showed that the instructional conditions *do* lead to different mean GMRT delayed posttest scores for students with PVK, meaning that students with PVK were different at the second point of Time (GMRT delayed posttest).

The simple main effects were followed up with simple comparisons to examine the differences between the means of the treatment and control conditions for students with PVK at the pretest and delayed posttest levels. Therefore, the analyses were focused first on analyzing the simple comparisons of the GMRT pretest scores between students with PVK in the treatment condition and those students with PVK in the control condition. The results of this comparison are shown in Table 4.12.

Table 4.12 Pairwise Comparisons for the GMRT Pretest and Delayed Posttest Scores by Condition for the PVK Group

Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
GMRT Pretest	Treatment	Control	.167	2.108	.937	-4.057	4.390	.018
GMRT Delayed Posttest	Treatment	Control	8.400	2.836	.005	2.717	14.083	.917

In Table 4.12, Bonferroni corrections were made with all pairwise comparisons within each simple main effect considered a family of comparisons. Adjusted *p*-values were reported. Table 4.12 shows that the mean GMRT pretest score is .17 points higher for students with PVK in the treatment condition compared to those students with PVK in the control condition. This difference is not statistically significant (i.e., $p = .937 > .05$). The 95% confidence intervals for the mean GMRT pretest score for the students with PVK ranged from -4.06 to 4.39, confirming a non-statistically significant result.

As shown in Table 4.13, the mean GMRT pretest score was higher for students with PVK in the treatment condition ($M = 13.50$, $SD = 3.63$) than those students with PVK in the control condition ($M = 13.33$, $SD = 3.89$), a mean difference of .17 points, 95% CI [-4.06, 4.39], $p = .937 > .05$, with a small effect size ($d = .02$). As expected, this small mean difference ($M = .17$) on the mean GMRT pretest scores between students with PVK in the treatment condition and their peers with PVK in the control condition was not statistically significant ($p = .937 > .05$). Therefore, it can be concluded that there

was no statistically significant difference in the mean standardized (GMRT) pretest scores between students with PVK in the treatment condition and control condition.

Table 4.13 Descriptive Statistics for the GMRT Scores by Condition x Group at the Pretest and Delayed Posttests

Time	Condition	Group	<i>N</i>	<i>M</i>	<i>SD</i>
GMRT Pretest	Treatment	PVK	10	13.50	3.629
		AVK	17	27.41	5.948
		Total	27	22.26	8.556
	Control	PVK	15	13.33	3.885
		AVK	17	28.35	5.968
		Total	32	21.31	9.121
GMRT Delayed Posttest	Treatment	PVK	10	22.20	9.647
		AVK	17	30.71	6.507
		Total	27	27.56	8.706
	Control	PVK	15	13.80	4.693
		AVK	17	27.35	7.202
		Total	32	21.00	9.162

When considering the differences in the GMRT delayed posttest scores between students with PVK in the treatment condition and those in the control condition, the results of this comparison indicated that the mean GMRT delayed posttest score was significantly higher for students with PVK in the treatment condition ($M = 22.20$, $SD = 9.65$) than those students with PVK in the control condition ($M = 13.80$, $SD = 4.69$), a mean difference of 8.4 points, 95% CI [2.72, 14.08], $p = .005 < .05$, with a large effect size ($d = .92$; see Tables 4.12 and 4.13). These findings answer subquestion 2B and confirm the hypothesis that, after receiving the vocabulary instructional intervention, students with PVK in the treatment condition outperformed students with PVK in the control condition on the standardized test (GMRT).

Subquestion 2C

Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the standardized test.

For this research question, the interest is to investigate the effect of Condition for the AVK group only. Therefore, the same steps followed to analyze the simple comparisons of both GMRT pretest and delayed posttest scores between students with PVK in the treatment condition and students with PVK in the control condition were also applied to analyze the simple comparisons for students with AVK.

Statistical significance of a simple main effect was accepted at a Bonferroni-adjusted alpha level of .025. Results indicated that there was no statistically significant simple main effect of Condition for students with AVK on their GMRT pretest scores, $F(1, 55) = .28, p = .597 > .05$. Unexpectedly, there was no statistically significant main effect of Condition for students with AVK on their GMRT delayed posttest scores, $F(1, 55) = 1.98, p = .165 > .05$. These results provided evidence to support the hypothesis that the instructional conditions (treatment and control conditions) *do not* lead to different mean GMRT pretest and delayed posttest scores for students with AVK. This means that students with AVK were not different at the two points of Time (GMRT pretest and GMRT delayed posttest).

The simple main effects were followed with simple comparisons to examine the differences between the means of the treatment and control conditions for students with

AVK at the pretest and delayed posttest levels. Therefore, the analyses were focused first on analyzing the simple comparisons of the GMRT pretest scores between students with AVK in the treatment condition and those students with PVK in the control condition.

The results of this comparison are shown in Table 4.14.

Table 4.14 Pairwise Comparisons for the GMRT Pretest and Delayed Posttest Scores by Condition for the AVK Group

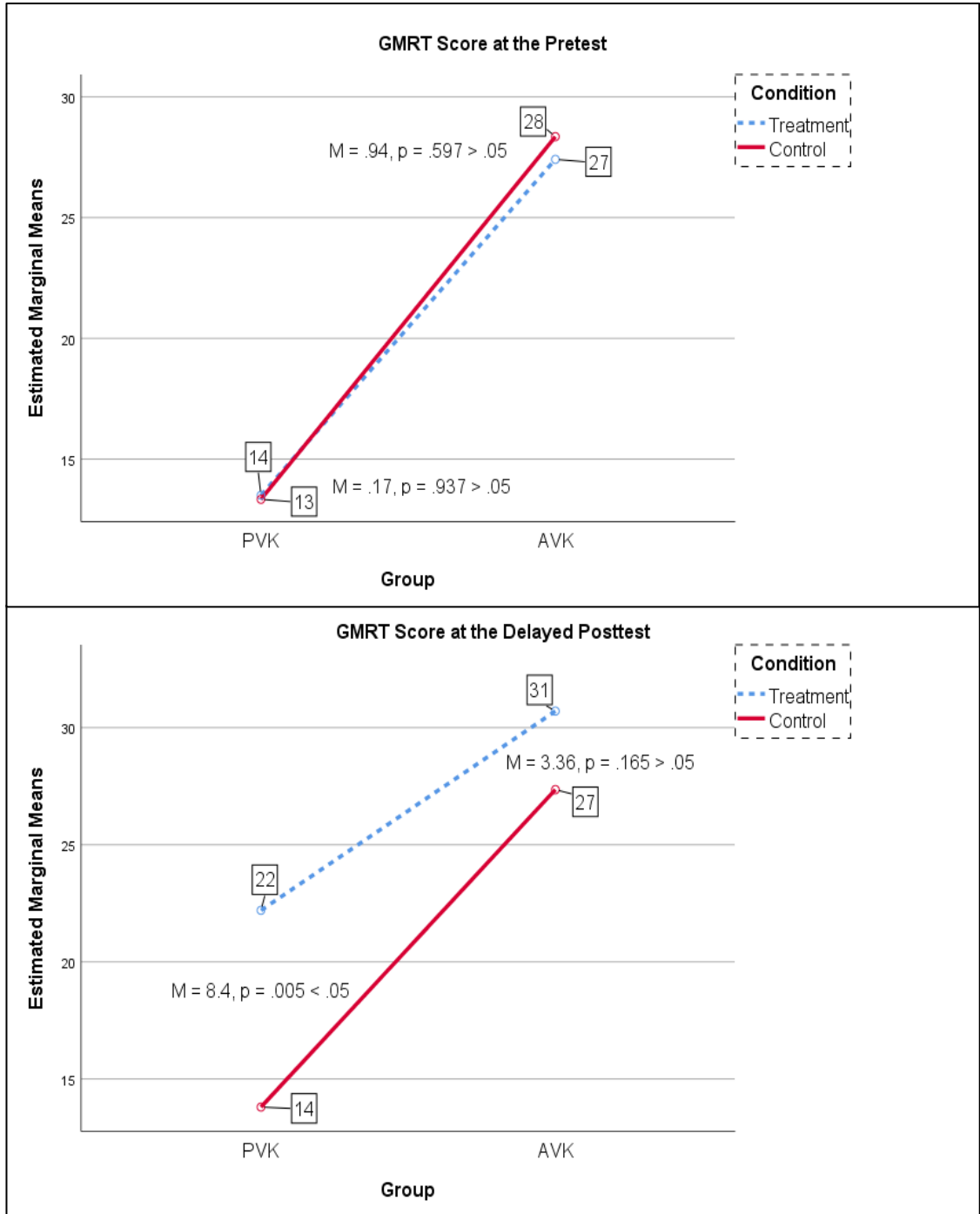
Time	(I) Condition	(J) Condition	Mean Difference (I-J)	Std. Error	<i>p</i>	95% Confidence Interval for Difference		Cohen's <i>d</i>
						Lower Bound	Upper Bound	
GMRT Pretest	Treatment	Control	-.941	1.771	.597	-4.490	2.607	.103
GMRT Delayed Posttest	Treatment	Control	3.353	2.382	.165	-1.421	8.127	.366

In Table 4.14, Bonferroni corrections were made with comparisons within each simple main effect considered a family of comparisons. Adjusted *p*-values were reported. Results indicated that the mean GMRT pretest score was slightly higher for students with AVK in the control condition ($M = 28.35$, $SD = 5.97$) than those students with AVK in the treatment condition ($M = 27.41$, $SD = 5.95$), a mean difference of .94 points, 95% CI [-4.49, 2.61], $p = .597 > .05$, with a small effect size ($d = .10$; see Tables 4.13 and 4.14). Furthermore, the mean GMRT delayed posttest score was higher for students with AVK in the treatment condition ($M = 30.71$, $SD = 6.51$) than those students with AVK in the control condition ($M = 27.35$, $SD = 7.20$), a mean difference of 3.35 points, 95% CI [-

1.42, 8.13], $p = .165 > .05$, with a medium effect size ($d = .37$; see Tables 4.13 and 4.14). However, both of these findings showed that there were no statistically significant differences in the scores of the standardized (GMRT) pretests or delayed posttests between students with AVK in the treatment condition and control condition. Therefore, these results fail to reject the null hypothesis that there is no difference in the standardized test scores between students with AVK in the treatment condition and control condition after receiving the vocabulary instructional intervention. Figure 4.3 illustrates the results of the means of GMRT pretest and delayed posttest scores for both groups of students (students with PVK and AVK) across both instructional conditions (the treatment and control conditions).

To conclude, as shown in Figure 4.3, when examining the performance of both groups of students (students with PVK and AVK), after receiving the vocabulary instructional intervention, only students with PVK in the treatment condition outperformed students with PVK in the control condition on the GMRT delayed posttest. Unexpectedly, as revealed by their results on the GMRT delayed posttest, students with AVK did not show a significant difference in their vocabulary acquisition ability as a result of exposure to the vocabulary instructional intervention.

Figure 4.3 Results of the Means of GMRT Pretest and Delayed Posttest Scores for Group and Condition



Research Question 3

The third, and final, research question of this study was: Are the students' results on both measures of vocabulary knowledge (the GMRT and the experimenter-constructed test) consistent across the two conditions (treatment condition and control condition)? I hypothesize that students' results on both measures of vocabulary knowledge across the two conditions will be consistent, which will confirm that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

A one-way ANOVA method was carried out to answer this major research question. The primary purpose for running the one-way ANOVA was to determine if there is a difference in the means of both measures of vocabulary knowledge (GMRT and ECT) between the two instructional conditions (i.e., treatment and control conditions). The raw scores of both measures of vocabulary knowledge were normalized to ensure that they were on the same standard scales prior to making any comparisons between the two measures. The normalization was achieved by converting the raw scores on either measures to z scores based on their percentile ranks. Henceforth, the z scores rather than the raw scores were used in the analyses.

To run a one-way ANOVA, there were six assumptions examined. Excluding the assumptions of sphericity and homogeneity of covariances, these six assumptions are similar to the assumptions that were examined in research questions 1 and 2 for the three-way mixed ANOVA design and the two-way mixed ANOVA design, respectively. Upon examination of these assumptions, it was found that the first three assumptions related to

the chosen study design (one-way ANOVA) were met. In addition, the scores of both measures of vocabulary knowledge were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$), and there were no outliers in the data, as assessed by an inspection of boxplots. There was no violation of the assumption of homogeneity of variances for all values tested (GMRT pretest scores, ECT pretest scores, GMRT delayed posttest scores, ECT delayed posttest scores), as assessed by Levene's test for equality of variances ($p = .577, p = .951, p = .867, p = .260$, respectively).

The analyses were focused first on analyzing the mean of the normalized GMRT pretest scores between the two instructional conditions (treatment and control conditions). Results showed that there was a very small difference in the mean of the normalized GMRT pretest scores between the two instructional conditions, the control condition ($M = .03, SD = 1.01$) and the treatment condition ($M = .18, SD = 1.02$), and, as expected, the difference was not statistically significant, $F(1, 57) = .33, p = .570 > .05$, with a negligible effect size (partial $\eta^2 = .006$) and almost negligible power of .09. Therefore, this finding fails to reject (i.e., supports) the null hypothesis that there is no difference in the normalized GMRT pretest scores between students in the treatment condition and those in the control condition before receiving the vocabulary instructional intervention. Table 4.15 presents the results of the one-way ANOVA for the effects of Condition at the normalized GMRT pretest scores.

Table 4.15 Results of the One-Way ANOVA for the Effects of Condition on the Normalized GMRT Pretest Scores

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2	Observed Power*
Condition	1	.336	.326	.570	.006	.087
Error	57	1.032				
Total	59					

* Computed using alpha = .05

When comparing the mean of the normalized ECT pretest scores between the two instructional conditions (treatment and control conditions), the results revealed that the mean of the normalized ECT pretest score was somewhat higher in the treatment condition ($M = .21, SD = .99$) compared to the control condition ($M = -.04, SD = .99$). However, this difference between the instructional conditions was not statistically significant, $F(1, 57) = .87, p = .354 > .05$, with a small effect size (partial $\eta^2 = .015$) and a very small power of .15. Therefore, this finding fails to reject (i.e., supports) the null hypothesis that there was no difference in the normalized ECT pretest scores between students in the treatment and control conditions prior to receiving the vocabulary instructional intervention. Table 4.16 presents the results of the one-way ANOVA for the effects of Condition at the normalized ECT pretest scores.

Table 4.16 Results of the One-Way ANOVA for the Effects of Condition on the Normalized ECT Pretest Scores

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2	Observed Power*
Condition	1	.863	.873	.354	.015	.151
Error	57	.989				
Total	59					

* Computed using alpha = .05

When considering the differences in the mean of the normalized GMRT delayed posttest scores between the two instructional conditions, the results of this comparison indicated that the mean of the normalized GMRT delayed posttest score was higher in the treatment condition ($M = .49, SD = .97$) than the control condition ($M = -.18, SD = 1.04$). This difference, as expected, was statistically significant, $F(1, 57) = 6.49, p = .014 < .05$, with a medium effect size (partial $\eta^2 = .102$ and Cohen's $d = .65$). This effect size means the vocabulary instructional intervention resulted in about a two-thirds increase in the average scores relative to the control group. Therefore, this finding rejects the null hypothesis and accepted the alternative hypothesis that there is a statistically significant difference in the normalized GMRT delayed posttest scores between students in the treatment and control conditions, as students in the treatment condition outperformed students in the control condition after receiving the vocabulary instructional intervention. Not only is the result significant, but the probability of correctly rejecting the null hypothesis of no difference between the treatment and control groups and finding a significant difference between the two conditions is close to .71, which is a more than

acceptable statistical power for this type of study. Table 4.17 presents the results of the one-way ANOVA for the effects of Condition at the normalized GMRT delayed posttest scores.

Table 4.17 Results of the One-Way ANOVA for the Effects of Condition on the Normalized GMRT Delayed Posttest Scores

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2	Observed Power*
Condition	1	6.616	6.488	.014	.102	.707
Error	57	1.020				
Total	59					

* Computed using alpha = .05

Similar results were found when comparing the mean of the normalized ECT delayed posttest scores between the two instructional conditions (treatment and control conditions). Results showed that the mean of the normalized ECT delayed posttest score was higher in the treatment condition ($M = .65$, $SD = 1.02$) compared to the control condition ($M = -.39$, $SD = .78$), and this difference, as expected, was statistically significant, $F(1, 57) = 19.62$, $p < .001$, with a large effect size (partial $\eta^2 = .256$ and Cohen's $d = 1.00$). Therefore, this finding rejects the null hypothesis and accepts the alternative hypothesis that there was a statistically significant difference in the normalized ECT delayed posttest scores between students in the treatment condition and those in the control condition, as students in the treatment condition outperformed students in the control condition after receiving the vocabulary instructional intervention. Not only is the result significant, but the probability of correctly rejecting the null

hypothesis of no difference between the treatment and control groups and finding a significant difference between the two conditions is .99, which is a very high statistical power for this type of study. Table 4.18 presents the results of the one-way ANOVA for the effects of Condition at the normalized ECT delayed posttest scores.

Table 4.18 Results of the One-Way ANOVA for the Effects of Condition on the Normalized ECT Delayed Posttest Scores

Source	<i>df</i>	<i>MS</i>	<i>F</i>	<i>p</i>	Partial η^2	Observed Power*
Condition	1	15.829	19.624	.000	.256	.992
Error	57	.807				
Total	59					

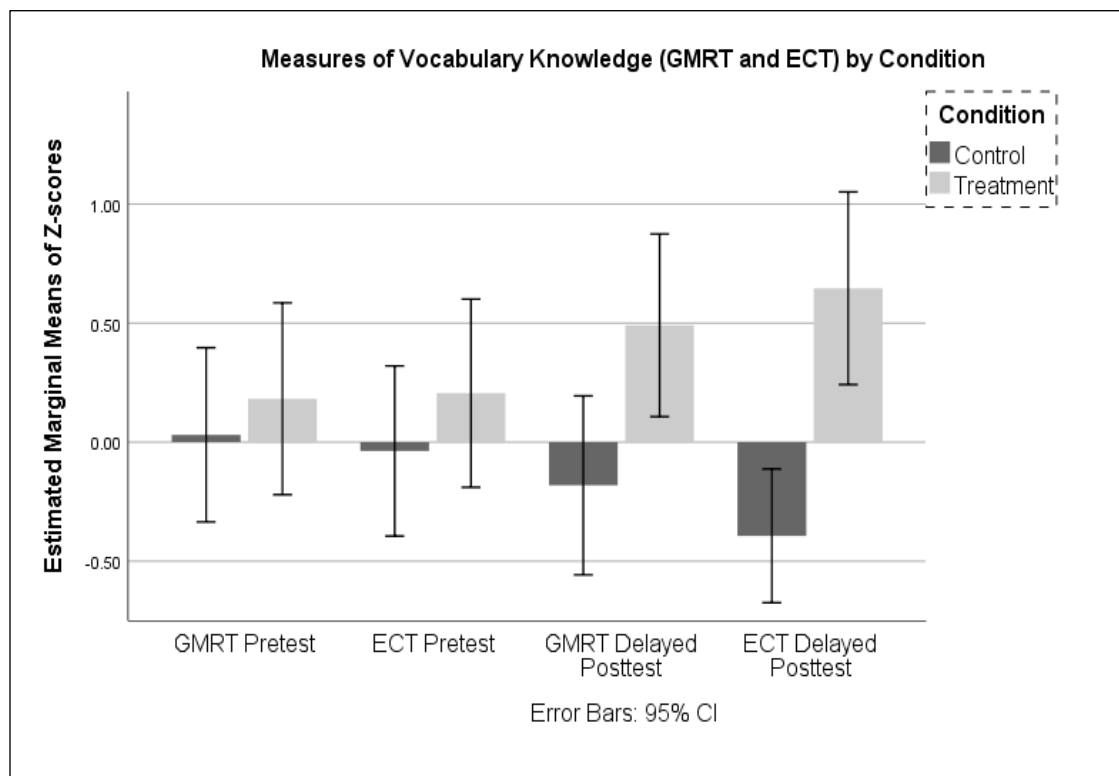
* Computed using alpha = .05

These results provided evidence supporting the hypothesis that the instructional conditions (treatment and control conditions) *do not* lead to different mean pretest scores on either of the two measures of vocabulary knowledge. The mean scores on the two instruments (GMRT and ECT) were not different at pretest for students in the treatment and control conditions. In addition, the difference in instructional conditions *did* lead to different mean delayed posttest scores for both measures of vocabulary knowledge. In summary, the results between the treatment and control groups on the ECT were consistent with the results on the GMRT both at pretest and the delayed posttest.

The results for the two instructional conditions for both measures are shown in Figure 4.4. The cell means plotted in Figure 4.4 show the changes in the students' results on both measures of vocabulary knowledge (GMRT and ECT) from the pretest (Time 1)

to the delayed posttest (Time 3) across the two instructional conditions (treatment and control conditions) separately. As expected, this figure shows that students in the treatment condition performed better than students in the control condition on both measures of vocabulary knowledge after receiving the vocabulary instructional intervention.

Figure 4.4 Changes in the Students' Results as a Function of Test and Condition



In brief, when examining the changes in the students' results on both measures of vocabulary knowledge (GMRT and ECT) from the pretest (Time 1) to the delayed posttest (Time 3) across the two instructional conditions (treatment and control conditions), the findings revealed that students in the treatment condition outperformed

students in the control condition on both measures of vocabulary knowledge three weeks after the completion of the vocabulary instructional intervention (on the delayed posttest). Additionally, these changes in students' results on both measures of vocabulary knowledge across the two instructional conditions were consistent, which confirmed the hypothesis that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

Interobserver Agreement

The classroom teachers in the treatment groups used scripted lesson plans that I created for this study to ensure proper implementation of the experimental techniques. I also created a teacher fidelity checklist of critical elements of the intervention for this study to assess implementation of the vocabulary instructional intervention during the instructional intervention lessons (see Appendix D). A total of 18 videos were recorded during the instructional intervention lessons. Two trained graduate students—who were not involved with the study—assessed the treatment fidelity using videotapes during 39% ($n = 7$) of the instructional sessions with interobserver agreement (IOA) data that were calculated for all these 39% sessions. One point was assigned for each correctly implemented step. To calculate a percentage score for IOA, the total number of agreements on steps implemented correctly was divided by the total number of agreements and disagreements, which was then multiplied by 100 (Cooper et al., 2007). Results revealed that the overall mean IOA for teachers' implementation of the instructional intervention lessons was 93% (range = 90%–100%), which met the criteria

of acceptable IOA across all observations (i.e., 80%–85%) in the field of special education (Cooper et al., 2007; O’Neill et al., 2011).

Inter-Scorer Reliability

Inter-scorer reliability was determined by having the same two graduate students who assessed IOA for the study independently score 25% of all tests, with the selection of tests distributed evenly across conditions and participants. The two scorers used a scoring sheet that I created for this study. Scored tests were compared item by item to determine the number of agreements and disagreements. To calculate percentage agreement, the total number of agreements were divided by the total number of agreements and disagreements, which was then multiplied by 100. Results showed that the mean inter-scorer agreement on the GMRT assessments was 100%, and the mean ECT inter-scorer agreement was 97% (range = 94%–100%), which met the minimum acceptable criteria of inter-scorer agreement (i.e., 80%) for evidence-based practices in the field of special education (CEC, 2014).

CHAPTER V

DISCUSSION

In the last 40 years, numerous researchers have called for promoting independent vocabulary learning strategies that will allow students to access the sheer volume of vocabulary they are likely to encounter in school texts (e.g., Baumann et al., 2002, 2003; Ebbers & Denton, 2008; Fukkink, 2002; Fukkink & de Glopper, 1998; Fukkink et al., 2001; Nagy, 1988; Nagy & Anderson, 1984). In 1988, Nagy argued that there is value in teaching morphemic and contextual analysis:

Two widely used methods of helping students learn to deal with unfamiliar words on their own are context and structural [morphemic] analysis. There is no doubt that skilled word learners use context and their knowledge of prefixes, roots, and suffixes to deal effectively with new words. (p. 38)

Based on their meta-analysis of 21 studies of instruction in generative vocabulary interventions (i.e., contextual analysis strategies), Fukkink and de Glopper (1998) concluded,

The research of strategy instruction lacks a process model that describes the process of deriving word meaning from context of good and poor readers. Such a cognitive process model would be of help in conjunction with other tools in the design of strategy instruction. Although the process of deriving word meaning from context has been investigated in some think-aloud studies (Daalen-Kaptejns & Elshout-Mohr, 1981; Daalen-Kaptejns, Schouten-van Parreren, & de Glopper, 1997; Werner & Kaplan, 1953), we do not know yet if any expert strategy exists that can be taught to novices with some success. (p. 462)

Ebbers and Denton (2008) also asserted,

There is reliable evidence that new vocabulary is primarily acquired through wide independent reading. However, struggling readers tend to avoid reading, resulting in limited word encounters and inadequate vocabulary growth, and they often have difficulties inferring the meanings of new words from context. (p. 90)

One common element to these scholars' ideas is that there is value in teaching students to employ contextual analysis as a word-learning strategy. Furthermore, researchers such as Fukkink (2002) and Matsuda (1987) pointed out that research, teacher surveys, and reading methodology textbooks since 1900 consistently attest to the value of instruction related to deriving word meaning from written context. It is generally accepted that deriving word meaning from written context is an important source of vocabulary expansion for primary-school students, particularly students with reading difficulties (Fukkink et al., 2001). The word in a written context provides more information (e.g., semantic, syntactic, pragmatic) to create a well-specified semantic representation (Nash & Snowling, 2006).

However, Graves (2000) noted that whereas teaching specific words is a common activity in classrooms, teaching students strategies for deriving or inferring word meanings is relatively rare. "I know of no detailed archival description of a well-planned, serious, powerful, and long-term attempt to teach students word-learning strategies. This is very bad news" (Graves, 2000, p. 123). The NRP committee (2000) noted a similar conclusion, "[It] knows a great deal about the ways in which vocabulary increases under highly controlled conditions, but the Panel knows much less about ways in which such

growth can be fostered in instructional contexts” (p. 4-27). Ogle and Blachowicz (2002) explicitly called for research on word-learning strategies in subject-matter classrooms:

We have been struck by the paucity of research studies on vocabulary in informational reading. We hope to see studies that observe the ways in which teachers incorporate vocabulary instruction into their content classrooms using strategies that develop students’ abilities to be independent word learners. (p. 270)

Additionally, as discussed in chapters one and two, a thorough investigation of the available literature revealed a need for further studies into the area of teaching generative vocabulary interventions through a combination of a cognitive independent word-learning strategy and certain types of context clues. The research in this quantitative dissertation study is an effort to close this gap in the vocabulary instruction literature. I investigated the effects of vocabulary instruction that was based on a combination of a strategy and certain types of context clues for deriving word meanings on short- and long-term vocabulary acquisition in fourth-grade students with adequate and poor vocabulary knowledge. The concept of the *fourth-grade slump* described by Chall et al. (1990) gave reason to investigate and focus on fourth graders.

In this chapter, I provide a discussion of the effects associated with the vocabulary instructional intervention used in the present study, beginning with an overview of the study’s major findings. Second, I discuss and interpret these findings in relation to the existing research regarding generative vocabulary interventions. This is followed by a description of the strengths and limitations of the study. In the fourth section of this chapter, I provide some recommendations and implications for future research. I close the chapter with a final conclusion.

Overview of the Results

As discussed in chapters three and four, I used two measurement instruments (GMRT and ECT) to measure the dependent variable (students' vocabulary acquisition) in this dissertation study. Those measures of vocabulary acquisition were administered to participants at three different moments in the study: before providing the vocabulary intervention, immediately after completing the intervention, and three weeks later. Since the individual measures are related to different research questions, the results are presented separately with respect to the major research questions and subquestions and briefly discussed in this section.

Research Question 1

Subquestions of research question one examined whether or not there was a differential increase in students' vocabulary acquisition and maintenance of the learned context clue strategy three weeks post-intervention. More specifically, the emphasis of the first three subquestions in research question 1 is placed on examining the impact of the vocabulary instructional intervention on the performance of the students (students with PVK and AVK) via the ECT measure. Research question 1 is:

1. Does the vocabulary instructional intervention have any effect on the performance of the students on the experimenter-constructed test? If so, for which group of students (students with poor vocabulary knowledge [PVK] and students with adequate vocabulary knowledge [AVK]) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (1A–1D):

Subquestion 1A. Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) overtime? I hypothesize that students in the treatment condition will perform better than students in the control condition on the experimenter-constructed test.

Subquestion 1B. Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the experimenter-constructed test.

Subquestion 1C. Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the experimenter-constructed test.

Subquestion 1D. Is there a difference in retention between short-term learning and long-term learning (Time 2 vs. Time 3) for students in the treatment condition? I hypothesize that the performance of the students will be better in short-term learning (Time 2) in the treatment condition on the experimenter-constructed test.

Results showed that after receiving the vocabulary instructional intervention and regardless of students' groups, there were statistically significant mean differences in the ECT scores at the immediate and delayed posttests across both instructional conditions

(the treatment and control conditions). As predicted, students in the treatment condition performed significantly better than students in the control condition on the ECT.

When comparing the performance of both groups of students (students with PVK and AVK) across both instructional conditions, the findings revealed that there were statistically significant mean differences in the scores of the experimenter-constructed tests at the immediate and delayed posttests for both groups of students. As hypothesized, both groups of students in the treatment condition outperformed their peers with PVK and AVK in the control condition on the ECT measure. However, students with AVK performed significantly better than students with PVK in both instructional conditions (the treatment and control conditions) on the ECT measure.

Subquestion 1D measured the extent to which the students with AVK and PVK retain and apply the learned context clue strategy three weeks after the immediate posttest. If we teach students using contextual analysis strategies, a worthwhile question for a school administrator or classroom teacher would be, “Will it make a difference in student vocabulary learning?” Followed by, “Will students retain the information?” If we spend our time elaborating and delving deeper into this area of vocabulary instruction and strategies, school administrators and classroom teachers will not be pleased with using an intervention that proves helpful in the short term but lacks long-term retention. Thus, it was important to administer a measure that would offer some insight as to whether or not (a) these same students in the treatment groups were retaining the learned context clue strategy, and (b) where students stand in comparison to those uninstructed control-group

students, who received instruction in a traditional manner. Therefore, the ECT measure was used to provide an answer to this subquestion.

The data gathered from the ECT immediate and delayed posttests in this experiment revealed that students in the treatment groups not only learned the context clue strategy but also (a) maintained the knowledge of the newly acquired strategy for three weeks after the completion of the vocabulary instructional intervention, and (b) demonstrated that they could independently apply it to infer word meanings from novel written contexts. However, although the vocabulary intervention positively impacted the students' use of the learned context clue strategy, findings showed that students were more successful in using the strategy on the ECT immediate posttest than on the ECT delayed posttest. To illustrate, the findings demonstrated that there was a small difference in retention as it was a small decrease in the performance of both groups of students from the immediate posttest (short-term learning, Time 2) to the delayed posttest (long-term learning, Time 3) on the ECT measure. Thus, both groups of students performed significantly better in short-term learning compared to their performance in long-term learning. Furthermore, when comparing the performance of students with PVK to their peers with AVK, findings showed that students with AVK in the treatment condition outscored their peers with PVK both times, Times 2 and 3.

Research Question 2

The emphasis of all subquestions in the major research question two is placed on examining the effects of the vocabulary instructional intervention on the performance of the students (students with PVK and AVK) on the GMRT measure. Research question two is:

2. Does the vocabulary instructional intervention have any effect on the performance of the students on the standardized test (Gates-MacGinitie Reading Tests [GMRT])? If so, for which group of students (students with PVK and students with AVK) does the intervention have a significant effect?

This major question is addressed by answering a number of subquestions (2A–2C):

Subquestion 2A. Is there a difference in vocabulary acquisition between students in the two conditions (treatment condition and control condition) at the pretest and delayed posttest? I hypothesize that students in the treatment condition will perform better than students in the control condition on the standardized test.

Subquestion 2B. Is there a difference in vocabulary acquisition between students with PVK in the treatment condition and control condition? I hypothesize that students with PVK in the treatment condition will perform better than students with PVK in the control condition on the standardized test.

Subquestion 2C. Is there a difference in vocabulary acquisition between students with AVK in the treatment condition and control condition? I hypothesize that students with AVK in the treatment condition will perform better than students with AVK in the control condition on the standardized test.

The GMRT, which is a standardized reading test measuring vocabulary knowledge and acquisition, was in a different format than the students were accustomed to. Specifically, in the GMRT, students in Grades 4 through 12 are required to mark their answers in a separate answer document as opposed to students in Grades K through 3, who mark their answers directly in the test booklet. Answering and marking the answers on a separate bubble-answer sheet was found to be a new and difficult task for many student participants, especially students with reading difficulties, including those with learning disabilities (LD). Nevertheless, as hypothesized, after receiving the vocabulary instructional intervention, it was found that—regardless of students’ groups—students in the treatment condition performed significantly better than students in the control condition on the GMRT measure. This finding of the study provides evidence that, even though the GMRT was in an unfamiliar test format for the students, teaching students how to use context clues while reading improves their understanding and ability to derive the meanings of novel vocabulary words in new written contexts.

After comparing the performance of both groups of students (students with PVK and AVK) across both instructional conditions (the treatment and control conditions), the findings revealed that both groups of students in both conditions were not different at the GMRT pretest. As predicted, when examining the performance of both groups of students

after receiving the vocabulary instructional intervention, it was found that only students with PVK in the treatment condition outperformed students with PVK in the control condition on the GMRT delayed posttest. In fact, findings showed that there were no statistically significant differences in the scores of the standardized (GMRT) pretests or delayed posttests between students with AVK in the treatment and control conditions. There was no statistically significant main effect of Condition for students with AVK on their GMRT delayed posttest scores after receiving the vocabulary instructional intervention. This may mean that there was not much room for them to grow, or the test is not sensitive enough to differentiate between students with AVK in the control condition and those in the intervention condition. Therefore, these results fail to reject the null hypothesis that there is no difference in the standardized reading test scores between students with AVK in the treatment and control conditions after receiving the vocabulary instructional intervention. Thus, the standardized (GMRT) measure indicated a clear advantage to students with PVK in the treatment group.

Research Question 3

Research question three examined whether the students' results on both measures of vocabulary knowledge (GMRT and ECT) were consistent across the two instructional conditions (treatment and control conditions). Research question three is:

3. Are the students' results on both measures of vocabulary knowledge (the GMRT and the ECT) consistent across the two conditions (treatment condition and control condition)? I hypothesize that students' results on both measures of vocabulary knowledge across the two conditions will be consistent, which will

confirm that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention.

Both measures of vocabulary knowledge were administered before the beginning of the intervention and at seven weeks. When examining the changes in the students' results on both measures of vocabulary knowledge from the pretest (Time 1) to the delayed posttest (Time 3) across the two instructional conditions, the findings revealed that students in the treatment condition outperformed students in the control condition on both measures of vocabulary knowledge three weeks following the completion of the vocabulary instructional intervention (on the delayed posttest). Additionally, these changes in students' results on both measures of vocabulary knowledge across the two instructional conditions were consistent, which confirms the hypothesis that the change in students' vocabulary acquisition was a result of exposure to the vocabulary instructional intervention. Thus, the results of the present study provided evidence that when students in the treatment groups were exposed to even a short intervention, they demonstrated significantly more positive change in the ability to use context clues to infer and derive the meanings of unknown words from written contexts than did the uninstructed control-group students.

Uniqueness of the Present Study

Before discussing and interpreting the findings of the present study, it is worth mentioning that this dissertation study differs from existing studies regarding the generative vocabulary interventions in some particular ways. First, as discussed in chapter two, none of the reviewed studies on the combination of strategy and context

clues (Buikema & Graves, 1993; Fukkink, 2002) included a delayed posttest to pinpoint the effects of the learned context clues strategy in the long term on students' ability to glean the meanings of unfamiliar words from written contexts. Only in this dissertation study were all student participants in the control groups and treatment groups exposed to two unannounced delayed posttests (ECT and GMRT) three weeks after the immediate posttest was completed. By adding a delayed posttest, this study was able to: (a) compare the influence of the vocabulary instructional intervention on the performance of students in the treatment groups with their peers in the control groups who did not receive the intervention and (b) examine the students' retention of the learned context clues strategy. Because the intervention studies in teaching contextual analysis provide little information about delayed effects, it is not possible to interpret the present findings in relation to the broader literature. However, future research can benefit from this contribution of new information, providing researchers with evidence that there is a difference in retention between short-term learning (measured by the immediate posttest) and long-term learning (measured by the delayed posttest) for students receiving the context clues strategy in the treatment condition.

Second, none of the existing studies on teaching generative vocabulary interventions have ever calculated and compared students' gain scores in two measurement instruments (similar to that employed in the present study, the standardized reading [GMRT] delayed posttest and the ECT delayed posttest) and, then, examined if their results were consistent across the two measures. This was crucial because the consistency between the two measures confirms that the change in students' vocabulary

acquisition was a result of exposure to the vocabulary instructional intervention. In fact, several of the studies examined, especially in Fukkink and de Glopper's (1998) meta-analysis, used the researcher-developed tests as the singular measurement instrument to measure the effect of the instructional interventions provided in these studies.

Additionally, when examining the two measurement instruments (GMRT and ECT)—used in the present study to test students' ability to derive meanings of unfamiliar words in written contexts—to determine whether the higher scores achieved by the treatment groups was due to their performance on a particular test item, correct responses were distributed across several items, suggesting that the effect was not due to a single test item.

Third, the present study included fourth-grade students with LD, which extended the work of Buikema and Graves (1993) and Fukkink (2002). As discussed in chapter two, when conducting an in-depth review of the research literature over the past six decades, only those two experimental studies that taught students a generalized strategy and list of context clues with the intention of improving their efficiency in using context to learn unfamiliar word meanings were found. Thus, the present study extends the work of previous research because in addition to teaching a mixture of certain types of context clues and a cognitive independent word-learning strategy in the intervention, it recruited students with LD as participants in the study.

Lastly, teacher participants in the treatment groups were provided with the materials and training to implement the vocabulary instructional intervention and did so for a period of nine consecutive days. Data of the treatment fidelity showed that the

vocabulary instructional intervention used in the present study is a user-friendly intervention that the teacher participants in the treatment groups could easily learn to implement with high fidelity. As described in chapter three, the teacher participants in the treatment groups were trained twice (individually) for a total of one hour (approximately 30 minutes per session), on how to conduct the vocabulary instructional intervention and how to use the Swivl and iPad devices. These training sessions were reduced from four to two based on Helman et al.'s (2015) recommendation, since teacher participants quickly acquired the target skills. This contrasts with the findings from the existing studies regarding the length of training sessions and professional development, in which the researchers found that teachers need to have at least 20 hours of contact time, including workshops, lectures, and ongoing coaching (e.g., Desimone, 2009; Joyce & Showers, 1982; Leko & Brownell, 2009) to gain knowledge and use it in practice. In this dissertation study, the teacher participants had interacted with me for only a total of one-hour training sessions and informal conversations about the intervention, and nevertheless, they could implement it with high fidelity.

Moreover, despite the short time period of the vocabulary instructional intervention that was provided to the students (20 to 30 minutes as part of the 50-minute social studies lessons for only nine consecutive days) in the present study, the findings showed that students gained the full benefits of the instructional intervention. Hence, as displayed in chapter four, enough consistency in implementation of the vocabulary instructional intervention occurred to indicate adequate implementation, which suggests that the internal validity of the present study does not seem to be threatened.

Interpretation of the Findings

This dissertation study, which included a diverse representation of students through schools, districts, campuses, male and female participants, and ability, did demonstrate a difference in vocabulary. The results of the present study showed that the students in the treatment groups made excellent use of the context clue strategy they were taught. Indeed, they were able to derive and infer significantly more word meanings from new written contexts than those students in the control groups at both immediate and delayed posttests. Considering that students in the two instructional conditions (treatment and control conditions) in the present study were well matched for initial vocabulary knowledge (as supported by the findings of the standardized reading GMRT pretest) and that the control groups had received no intervention, the difference between the performance of students in the two instructional conditions at the posttests gives an indication of the developed vocabulary instructional intervention effect can have on the deriving ability of students with adequate and poor vocabulary knowledge.

The present evidence also suggests that even brief, direct instruction of word-learning strategies is beneficial for increasing vocabulary knowledge and improving vocabulary acquisition in both groups of students—students with adequate and poor existing vocabulary knowledge. Indeed, the vocabulary instructional intervention that was employed in the present study was effective with just 4.5 hours of classroom teaching. Because it is an independent word-learning strategy, as supported by a significant transfer effect, it has the potential to continue to increase vocabulary knowledge and acquisition ability when teaching has ended, as supported by a significant

retention that was found in the present study. Consequently, this might help students learn more words incidentally as they are encountered in everyday reading, which is the core purpose of the generative vocabulary interventions. Thus, there seems to be emerging evidence for Sternberg's (1987) argument that "if, indeed, most vocabulary is learned from context, then what we most need to do is not to teach vocabulary from context, but to teach students to use context to teach themselves" (p. 97). Importantly, the vocabulary instructional intervention developed in the present study could be extended for use in oral work, in cases where students are not able to read the sentences or passages containing the context clues (Nippold, 2002). For the reasons stated, the vocabulary instructional intervention is recommended for use in schools with students who require extra help developing vocabulary knowledge and acquisition.

The results of the present study are consistent with several extant intervention studies that involved teaching fifth-grade students to use selected context clues (see Askov & Kamm, 1976; Baumann et al., 2002, 2003; Carnine et al., 1984; Hafner, 1965; Jenkins et al., 1989; Patberg et al., 1984) as well as studies involving young students (i.e., students aged 6–8 years old; see Martin-Chang et al., 2007, Nash & Snowling, 2006; Yuen, 2009) and adolescents (see Buikema & Graves, 1993). The present findings also support the results of the meta-analysis conducted by Stahl and Fairbanks (1986), involving studies with adults and children, on the effects of instruction in deriving word meanings from context. Thus, there is accumulating evidence that instruction aimed at enhancing the skill of deriving word meaning from context during reading does have a positive effect.

The findings of this study are also (to some extent) at odds with those of other researchers who have investigated the helpfulness of context in elucidating word meanings (see Beck et al., 1983; Fukkink, 2002; Patberg & Stibbe, 1985; Schatz & Baldwin, 1986; Sternberg, 1987). However, even the small changes in a poor or adequate student's ability to derive word meaning from context can have the potential impact of improving the deriving ability on their annual vocabulary acquisition during reading. Students cannot be expected to give a dictionary-like definition of an unknown word after just one encounter (Fukkink et al., 2001). Certainly, once a meaning is derived, it is not necessarily learned, but deriving the meaning may be the first step toward learning it. Thus, it follows that improved deriving ability could have considerable effects on vocabulary learning (Jenkins et al., 1989).

Contribution to the Broader Research

This dissertation study makes three contributions to the existing research on vocabulary instruction and independent vocabulary learning strategies, since research in the related literature is still limited. First, the findings of the present study provide evidence that deriving word meaning from context appears amenable to instruction. This finding parallels the findings of the majority of the meta-analyses conducted on the effects of instruction in deriving word meanings from context (e.g., Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998; Stahl & Fairbanks, 1986). Second, findings support the efficacy of teaching contextual analysis to derive the meanings of novel words in conjunction with subject-matter texts, and such word-learning instruction does not impede or degrade students' content learning.

Third, both students with adequate and poor vocabulary knowledge, including those with LD, are able to learn how to use context to derive the meanings of unfamiliar words when they have additional support. This aligns with the findings of Nash and Snowling's (2006) study, in which student participants with poor vocabulary knowledge showed positive outcomes in improving their vocabulary knowledge and demonstrated that they could independently use the newly acquired strategy to derive meanings from written context three months following the intervention. Furthermore, as discussed in chapter two, researchers found that:

- Students' growth in vocabulary can be best accounted for by independent reading (Harris et al., 2011; Karbalaei et al., 2012; Nagy et al., 1987).
- Students whose vocabularies are most in need of support are less likely to be able to get information from context (Beck et al., 1983; Beck et al., 2002; Graves, 2006; Schatz & Baldwin, 1986).
- Deriving word meaning from context is a complex process that is susceptible to errors at several points (Baumann et al., 2003; Fukkink, 2002, 2005; Fukkink & de Glopper, 1998).
- Students with reading difficulties, especially students with LD, often have minimal to no training in deriving meanings for unfamiliar words using context (Fukkink & de Glopper, 1998; Kuhn & Stahl, 1998). As a result, they often have fragmented knowledge of words, particularly a narrower understanding of word features (Jitendra et al., 2004; Swanson et al., 1999; Swanson & Vaughn, 2010).

These findings emphasize that teaching students how to use context to derive word meanings is quite important. The purpose of teaching strategies to improve the learning of word meanings from context is to help students learn more words incidentally as they are encountered in everyday reading. Doing so should, in turn, lead to a larger vocabulary over time as students read texts containing unknown words (Kuhn & Stahl, 1998). Hence, providing even brief, direct, and explicit vocabulary instruction in various cognitive and metacognitive strategies to assist students in independently finding the meaning of unknown words during reading, similar to that employed in this dissertation study, represent one of the best solutions to address the critical obstacles to enhanced vocabulary development for students. This is particularly applicable to students with reading difficulties, including those with LD. Chapter one includes detailed discussion of explicit instruction; for learning more about critical obstacles to enhanced vocabulary development, see Jitendra et al. (2004) and Stahl and Nagy (2006).

According to Karbalaee et al. (2012), if students are explicitly taught how to use context as a vocabulary learning strategy, their ability to learn words independently may increase. Explicit instruction might also enhance students' reading comprehension ability. In their series of studies, Williams and colleagues (2004, 2007, 2009) show that explicit instruction is helpful to facilitate student reading comprehension. Moreover, the statistically significant results that were reported in this dissertation study by both groups of students (PVK and AVK) in the treatment condition suggest that instruction in context clue strategies can improve the ability of students with reading difficulties, including those with LD, to use context to derive the meanings of unfamiliar words. The findings of

this dissertation study are among the very few studies that have reported statistically significant results that were supported by and pointed to the benefit of the context clues as a teaching strategy.

In summary, numerous studies have supported the efficacy of teaching students to employ contextual analysis as a word-learning strategy, and results of the current investigation, at least in part, reinforce this trend. Indeed, the findings of the present study make an important contribution to what is known about training students to use context. These findings also enable us to conclude, with some degree of confidence, that training in this skill appears to be worthwhile, at least for students who have vocabulary deficits, while at the same time highlighting areas for future research.

Limitations and Strengths of the Present Study

The findings of the present study need to be placed in the context of several limitations. First, although a quasi-experimental design was employed in the present study, only 59 students in three urban cities of a southern state participated in this study. This small sample size is potentially a limiting factor, although small sample sizes are not unusual in intervention studies that target students experiencing reading difficulties (e.g., Goerrs et al., 1999; Leong, Simmons, & Izatt-Gambell, 1990). However, it would be hard to generalize the findings of this investigation to all students, especially students with reading difficulties, across the country given the small sample size. Moreover, the present study might not be generalized beyond fourth-grade classrooms. More research could be conducted in a broader geographical context where a larger sample size with a wider range of abilities might be available. In their meta-analysis, Fukkink and de Gloppe

(1998) suggested that a total sample size of 136 would be needed for a statistical power of .80 in the word derivation studies; the statistical power can also be raised by using covariance analysis. Therefore, as discussed in chapter three, the advantages of a large sample size to interpret significant results are that it allows a more precise estimate of the treatment effect, and it usually is easier to assess the representativeness of the sample and to generalize the results (Biau et al., 2008).

Second, I created all of the lessons in the vocabulary instructional intervention that were used in the present study. While I created the lessons in the instructional intervention using evidence-based practices found in the literature for teaching the context clues strategy, and had these lessons reviewed by experienced colleagues, important questions remain about the ability of other teachers or researchers to create effective vocabulary instructional interventions. This is an important question to be answered by future research.

Third, to make the vocabulary instructional intervention in the present study more practical and relevant, I took nonfiction texts from the fourth-grade social studies curriculum to use as part of the instructional intervention. Hence, results are restricted to vocabulary instruction infused into social studies curriculum and materials.

Finally, because the main focus of the present study was to examine the impact of the vocabulary instructional intervention on students' vocabulary acquisition, unfortunately the potential effects of the vocabulary instruction on students' reading comprehension ability was not measured. Indeed, Level 4 of the Fourth Edition of the GMRT was used in the present study as one of the measurement assessments. Each test

level (e.g., Level 4) of this standardized norm-referenced test was designed for a given grade (Grade 4) and intended to be given three times a year (Fall, Winter, Spring) to assess students' reading progress, specifically in the areas of vocabulary and reading comprehension. Despite the fact that the GMRT was used in the pretest as a general benchmark of students' vocabulary knowledge prior to the study and in the delayed posttest, given three weeks after the immediate posttest, the main purpose of using the GMRT in the present study was to examine if there was any growth in students' vocabulary acquisition ability as a result of exposure to the vocabulary instructional intervention.

It is accepted that the time and effort spent instructing students in context clues to improve their ability in inferring unfamiliar words in written contexts stems from a desire for both better reading comprehension and faster vocabulary development. Researchers have asserted that nearly all academic learning requires comprehension of text whether in science, social studies, or even mathematics (Knight, Browder, Agnello, & Lee, 2010). Nagy and Scott (2000) described word meanings as making up as much as 70–80% of comprehension. As discussed in chapter two, only three of the previous studies reviewed included a measure of reading comprehension (Baumann et al., 2002, 2003; Hafner, 1965), and none of these studies showed significant gains, although both of Baumann et al.'s studies conceded possible measurement issues with their true/false comprehension questions. As a result, in the present study, it was vital to assess students' progress not only in vocabulary but also in reading comprehension. However, given the limited scope of this experiment with respect to the duration of the study (approximately seven weeks),

growth on the comprehension subtest of the GMRT standardized measure would likely be impossible. To illustrate, it is unrealistic to expect an immediate, direct comprehension transfer effect for instruction only in contextual analysis, particularly in a short-term experiment that does not include the components of comprehension instruction and wide reading.

In spite of these limitations, the present study reinforces and extends the limited extant empirical base regarding the efficacy of teaching selected types of context clues alone (Askov & Kamm, 1976; Baumann et al., 2002; Baumann, Edwards et al., 2003; Buikema & Graves, 1993; Carnine et al., 1984; Hafner, 1965; Jenkins et al., 1989; Patberg et al., 1984) and cognitive independent word-learning strategies and certain types of context clues in combination (Buikema & Graves, 1993; Fukkink, 2002). Thus, there is support for the traditional practice of teaching early- to upper-elementary students to employ generative vocabulary interventions to decipher the meanings of novel words while reading in written contexts.

Recommendations and Implications for Future Research

The present study has brought about possibilities for future research extending this study. One possibility might be to broaden this particular study using more classrooms and developing long-term curriculum into a full school year to evaluate the outcomes with an extended period of intervention. Many researchers agree that vocabulary instruction and its application in the classroom are critical (e.g., Fukkink, 2002); however, it is also important that more studies are completed in classroom settings that allow investigation of current practices and test the value of future practices.

Additionally, I am interested to see in future research the results of similar intervention studies with a larger focal group over a longer period of time. Indeed, the length of the intervention and frequency of the lessons in the present study were in line with other related studies (Askov & Kamm, 1976; Baumann et al., 2002; Baumann, Edwards et al., 2003; Fukkink, 2002; see also several of the studies examined in the meta-analysis conducted by Fukkink & de Glopper, 1998). However, if there is a difference over a nine-day period as shown in the present study, what might the results show over a longer period of time? The value of using contextual analysis strategies and emphasizing vocabulary in the classroom was proven in the present study to be effective for students and worth pursuing further studies in the future.

Additionally, since this study focused on social studies curriculum and materials, a future study involving other content areas such as science, ELA, and mathematics might provide insight into the effect of combined vocabulary instructional intervention in additional content areas. A primary goal of educational research is to solve problems that are relevant to student achievement. This requires interventions that are research based to find their way into classrooms. To increase the likelihood of research translation to practice, the research content and curriculum should reflect the requirements of the classroom. Thus, comparing the outcomes in multiple content areas could assist in drawing conclusions about similarities and differences in vocabulary strategy instruction using different content.

In fact, after completing the present study, I am left with a few questions. As strategy instruction should become part of content area instruction, and teaching

strategies takes up content instructional time, how much time should be dedicated to effectively teaching vocabulary strategies? Also, as research of vocabulary instruction in content areas is relatively new for content-area teachers, how long must a strategy be taught for students to automatically incorporate a strategy into their toolbox of strategies? Further research involving a large group of participants, with a wider range of abilities over a longer period of time, might help answer these research questions. Moreover, it is important from a methodological point of view, that in future studies random assignment of students to instructional conditions is applied, which can further increase internal validity.

The lessons in the vocabulary instructional intervention created for this study are scripted lessons, meaning they are highly structured lessons. In this method of teaching, the teacher is expected to read the lesson scripts verbatim. It is a form of direct instruction meant to guide teachers in order to sustain consistency in teaching the vocabulary instructional intervention by teachers in the treatment groups. Thus, when replicating the present study, the treatment fidelity will be a critical issue for the successful implementation of these scripted lessons in the vocabulary instructional intervention. The lack of high fidelity of implementation to the vocabulary instructional interventions embedded in the scripted lessons might lead to the lack of instructional effect. Consequently, the instructional intervention may lead to no or small improvement in students' understanding of the use of context clues as a strategy to determine the meaning of unfamiliar words in written contexts, which is the goal of the vocabulary instructional intervention in this dissertation study.

Even interventions that research proves can have a robust positive impact on student learning must be put into practice every day in the way developers intended because “no program—no matter how sound it is—can have impact if its essential elements are not used” (Yap, Aldersebaes, Railsback, Shaughnessy, & Speth, 2000, p. 32; see also Protheroe, 2008). In simple terms, an intervention or approach that is effective in some settings can be ineffective in others if the way it is implemented takes it far away from its original design. For example, in 2008, O’Donnell and Lynch conducted a study to examine fidelity of implementation to inquiry-based science units’ instructional strategies as a moderator of curriculum unit effectiveness. The researchers found positive effects on student achievement only when teachers used inquiry-based materials and when there was “high fidelity of implementation to the instructional strategies embedded in the materials” (p. 2). Accordingly, it may behoove future researchers to pay closer attention to the high fidelity of implementation to their vocabulary instructional interventions embedded in the scripted lessons.

Furthermore, future studies could utilize multiple measures—including observations, interviews, and surveys—for obtaining information about teachers’ understanding and use of independent-word learning strategies for teaching contextual analysis approaches to students with reading difficulties, including those with LD. As suggested by Jenkins et al. (1989), qualitative rather than quantitative changes may be called for if the instruction is to enhance students’ derivational skills further. For example, training in deriving word meaning could include instruction and practice in locating clues outside of the sentence that contains the target word.

Additional research is warranted to explore more thoroughly the potential effects of contextual analysis instruction on text comprehension. In particular, more sensitive comprehension measures are required, and it would be useful to examine the relationships among instruction in contextual analysis and other components of a balanced vocabulary program such as wide reading and explicit instruction in comprehension-critical vocabulary. The NRP report (2000) identified explicit instruction as one of the most important methods of teaching vocabulary.

Nine consecutive days of intervention in the present study is a brief period of time to demonstrate a change in terms of comprehension. The findings of the existing studies regarding contextual analysis strategies call for teachers to provide various strategies to struggling readers repeatedly over an extended time that they can utilize automatically on their own, to increase vocabulary knowledge and comprehension in content areas to narrow the gap in achievement from non-struggling readers. Thus, students with poor vocabulary knowledge may not automatically infer the meanings of new words from context but can be taught to do so, though their new word-learning strategies may require considerable reinforcement.

The use of a comprehension component in vocabulary instructional intervention studies that would offer additional questions over an extended time to demonstrate greater reliability, validity, and gains in progress is essential. The research literature does not necessarily predict a direct comprehension effect from teaching word-learning strategies. Nevertheless, we know that vocabulary knowledge is highly predictive of students' subsequent reading comprehension (e.g., Cunningham & Stanovich, 1997), and

several studies have demonstrated a direct link between teaching vocabulary and comprehension (e.g., Beck et al., 1983; Carlson et al., 2013; Chall et al., 1990; Freebody & Anderson, 1983; Kame'enui & Baumann, 2012; Lesaux & Kieffer, 2010; Oakhill et al., 2015; Reed et al., 2016; Rupley et al., 1998; Wagner & Meros, 2010). Thus, future research is required to elucidate whether teaching word-learning strategies positively influences reading comprehension for all students in the classroom, including students with LD, and if so, how. In addition, because the lessons in the vocabulary instructional intervention that were used in the present study were provided on consecutive days, further research is needed to evaluate the efficacy of teaching students contextual analysis strategies on separate days.

Future experimental studies should also incorporate a measure to evaluate the transfer of deriving word meaning to incidental word learning abilities. Incidental word learning not only involves meaning derivation but also memorization of word form and meaning. Thus, determining word meaning from context integrates with other skills in the incidental word learning process. As discussed in chapter two, it was expected that instruction would lead to improving the skill of deriving word meaning from context, which was followed by increased incidental word learning, and would result in accelerated vocabulary growth (Fukkink, 2002; Jenkins et al., 1989). This triple relationship has only been partially explored. Therefore, it could be useful to bridge the gap between word meaning derivation and incidental word learning to investigate their assumed relationship that creates the “sound and persuasive rationale” (Jenkins et al., 1989, p. 218) that underlies, or at least partly motivates, word derivation studies.

Research shows that students learn on average about 3000 words each year (Nagy & Anderson, 1984; Nagy & Herman, 1987). Nagy et al. (1987) estimated that incidental learning from written context represents about a third of this annual growth. Hence, future research needs to refine the instructional strategies and effects as well as the transfer of learning to typical reading tasks.

Another avenue for further study would be investigating whether the procedures developed for use in the present study with fourth-grade students could be adapted for use with secondary school students. Fukkink et al. (2001) stated that “the ability to derive word meaning from context depended on grade and concreteness of concepts” (p. 477). They also observed that deriving word meanings from written context is a significant source of vocabulary expansion for students at all ages. Therefore, it could be informative to implement a similar study at a higher grade level and focus on the knowledge of contextual analysis strategies in middle and high school students to see what their level of understanding is past the *fourth-grade slump*. What have students learned that will allow them to use these contextual analysis strategies to comprehend the complex vocabulary they encounter in later grades? Is there evidence to support that these students had already been introduced to multiple strategies to improve word knowledge and understanding in the earlier grades?

Finally, chapter two’s thorough investigation of the available literature revealed a need for further studies into the practices of teachers in the area of vocabulary instruction and interventions for their students who have reading difficulties, including those with LD. The literature review showed that teachers understand what research-based strategies

are, but those practices are not being utilized as needed within the classroom setting. Some researchers refer to this occurrence as *a research-to-practice gap* (Cook & Cook, 2013; Cook, Smith, & Tankersley, 2011; Slavin, 2002). It was hypothesized that the “gap” was due to the under-use of professional educational journals, which contain the most recent research for strategies in the area of vocabulary instruction. Most professional journals present empirical data using statistical analysis and interpretation and are written typically for professors and fellow researchers, not classroom teachers. The information contained in those journals may not be easily translated into practice in classrooms. Current research has shown some of the effective strategies that special education researchers can use to improve the dissemination of their findings and evidence-based practices (EBPs) in special education. These strategies include making dissemination simple and providing effective professional development.

An important strategy that special education researchers can apply in order to make their ideas useful to practice is keeping the writing and dissemination simple. Articles that are overly detailed or contain inaccessible terminology are difficult to remember or to apply. Special education teachers, whose jobs require quick decision-making in multiple contexts, need clear and concise messages they can use to guide their behavior in a variety of situations. When people are faced with more information than they can readily process, they often focus on only one particular part of that information (e.g., ease of implementation, availability of materials) or fail to focus on anything at all. This is clearly consistent with the tenets of cognitive load theory, which posits that “human working memory simply is not able to process many elements” (Van

Merriënboer & Sweller, 2010, p. 86). When writing includes overly-detailed information or jargon, it becomes easy to miss the main ideas of the content. In fact, Zikmund-Fisher, Fagerlin, and Ubel (2010) stated that “less can be more” when reporting information on EBPs (p. 661). Thus, limiting the amount of information presented in dissemination can enhance the meaning of that information for practitioners. In sum, to effectively disseminate information on EBPs in special education, researchers must present only the most critical and convincing information and evidence regarding a given practice.

Providing intensive professional development (PD) is considered to be another effective strategy for improving the dissemination of research findings and EBPs in special education. Effective PD increases teachers’ knowledge, skills, and attitudes related to new practices, which in turn should lead to new changes in instruction, thus leading to improved student learning (Whitworth & Chiu, 2015). Klinger, Vaughn, Arguelles, Hughes, and Leftwich (2013) illustrated that traditional approaches to PD are generally insufficient for affecting meaningful, long-term changes because they typically entail training teachers to implement new practices through brief one-time workshops. However, effective PD is a *continuous* process that is more comprehensive, intensive, and supportive. This ensures that teachers understand and can implement core components of new practices with fidelity, adapt the practices to fit their specific contexts, and sustain them over time in real-world conditions (Darling-Hammond, Hyler, & Gardner, 2017). Additionally, effective PD includes active learning opportunities, such as modeling, coaching, or discussing implementation efforts and problems (Darling-Hammond et al., 2017). Klinger et al. (2013) suggest that when researchers provide

effective PD, both researchers and practitioners can benefit. For example, practitioners can benefit from ongoing support provided by the researchers or developers of new practices, and at the same time, researchers can learn much about how practices work in real time scenarios. This can lead to development of even stronger innovations with greater potential for successful dissemination and use of EBPs.

Conclusion

It clearly appears that there is a need for further research on vocabulary instruction for students with reading difficulties and disabilities, especially for research in context clues and strategies for independent word learning, such as the one described in this dissertation study. The available research in this area is not as recent as it needs to be. The education field is in a constant state of change. To reflect this change, educational researchers—particularly reading researchers—need to continue to study the most effective methods for teaching vocabulary to all students, including those who have poor vocabulary knowledge. When students learn to apply the most common types of context clues, they may be better able to read and understand new and complex words. By providing students with strategies, such as the Outside-In strategy, for attacking words in context, teachers may promote independence in reading. If teachers across subject areas apply these approaches, students may learn to generalize the learning and habitualize the use of effective strategies for learning and remembering words. Students may experience an encouraging sense of accomplishment and further motivation as they begin to interact cognitively with words throughout the school day.

The findings of this dissertation study provide evidence that the vocabulary instructional intervention was effective and suggest that teaching students how to use context clues while reading—even brief, direct, and explicit vocabulary interventions—improves their understanding and ability to derive the meanings of novel vocabulary words in new written contexts. Thus, I believe that the situation is somewhat different now regarding “the very bad news” and “paucity of research” on teaching vocabulary “in instructional contexts.” Much work remains to be done, but I hope to have provided some insight into the vocabulary tricks related to how classroom teachers can provide students instruction in contextual analysis strategies.

REFERENCES

- Aladjem, D. K., & Borman, K. M. (2006, April). *Summary of findings from the national longitudinal evaluation of comprehensive school reform*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.
- Ali, S., & Rafi, M. (2016). Learning disabilities: Characteristics and instructional approaches. *International Journal of Humanities Social Sciences and Education*, 3(4), 111–115.
- Almasi, J., & Fullerton, S. (2012). *Teaching strategic processes in reading* (2nd ed.). New York, NY: Guilford Press.
- Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, 3(3), 21–34.
- Alzahrane, T. S. (2016). Using a technology to strategically support at-risk students' self-monitoring of on-task behaviors in general education classrooms. *National Social Science Technology Journal*, 6(1), 1–12.
- Alzahrane, T. S. (2019). *Labeling is not the issue: Advancing the benefits of labeling children with learning disabilities when response to intervention is implemented*. Unpublished manuscript.
- Amato-Zech, N. A., Hoff, K. E., & Doepke, K. J. (2006). Increasing on-task behavior in the classroom: Extension of self-monitoring strategies. *Psychology in The Schools*, 43(2), 211–221.
- American Psychiatric Association (APA). (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.
- Anderson, R. C., & Nagy, W. E. (1991). Word meanings. In R. Barr, M. Kamil, P. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research: Volume II* (pp. 690–724). New York, NY: Longman.
- Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator*, 16(4), 14–47.

- Anglin, J. M. (1993). *Vocabulary development: A morphological analysis*. Chicago, IL: Society for Research in Child Development.
- Archer, A. L., & Hughes, C. A. (2011). *Explicit instruction: Effective and efficient teaching*. New York, NY: Guilford Press.
- Askov, E. N., & Kamm, K. (1976). Context clues: Should we teach children to use a classification system in reading? *Journal of Educational Research* 69(9), 341–344.
- Baumann, J. F., Edwards, E. C., Boland, E., Olejnik, S., & Kame'enui, E. J. (2003). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meaning. *American Educational Research Journal*, 40(2), 447–494.
- Baumann, J. F., Edwards, E. C., Font, G., Tereshinski, C. A., Kame'enui, E. J., & Olejnik, S. (2002). Teaching morphemic and contextual analysis to fifth-grade students. *Reading Research Quarterly*, 37(2), 150–176.
- Baumann, J. F., & Kame'enui, E. J. (1991). Research on vocabulary instruction: Ode to Voltaire. In J. Flood, J. M. Jenson, D. Lapp, & J. R. Squire (Eds.), *Handbook of research on teaching the English language arts* (pp. 604–632). New York, NY: Macmillan.
- Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003). Research on vocabulary instruction: Ode to Voltaire redux. In J. Flood, D. Lapp, J. R. Squire, & J. M. Jensen (Eds.), *Handbook on research teaching the English language arts* (2nd ed., pp. 752–785). Mahwah, NJ: Erlbaum.
- Beck I. L., & McKeown M. G. (1991). Conditions of vocabulary acquisition. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), *Handbook of reading research* (pp. 789–814). Mahwah, NJ: Erlbaum.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction* (1st ed.). New York, NY: Guilford Press.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2013). *Bringing words to life: Robust vocabulary instruction* (2nd ed.). New York, NY: Guilford Press.
- Beck, I. L., McKeown, M. G., & McCaslin, E. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal*, 83(3), 177–181.

- Beck, I. L., McKeown, M. G., & Omanson, R. C. (1984, April). *The fertility of some types of vocabulary instruction*. Paper presented at the meeting of the American Educational Research Association, New Orleans, LA.
- Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 147–163). Hillsdale, NJ: Erlbaum.
- Bender, W. N. (2001). *Learning disabilities: Characteristics, identification, and teaching strategies* (4th ed.). Boston, MA: Allyn and Bacon.
- Berne, J. I., & Blachowich, C. L. Z. (2008). What reading teachers say about vocabulary instruction: Voices from the classroom. *The Reading Teacher*, 62(4), 314–323.
- Berninger, V. W., & Richards, T. L. (2002). *Brain literacy for educators and psychologists*. Amsterdam, Netherlands: Academic Press.
- Bianco, M., Bressoux, P., Doyen, A., Lambert, E., Lima, L., & Pelleng, C. (2012). Impact of early code-skill and oral comprehension training on reading achievement in first grade. *Journal of Research in Reading*, 35(4), 427–455.
- Biau, D. J., Kernéis, S., & Porcher, R. (2008). Statistics in brief: The importance of sample size in the planning and interpretation of medical research. *Clinical Orthopaedics and Related Research*, 466(9), 2282–2288.
- Biemiller, A. (1999, April). *Estimating vocabulary growth on ESL children with and without listening comprehension instruction*. Paper presented at the annual conference of the American Educational Research Association, Montreal, Quebec.
- Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology*, 98(1), 44–62.
- Blachowicz, C. L. Z., & Fisher, P. J. (2000). Vocabulary instruction. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 503–523). Mahwah, NJ: Erlbaum.
- Blachowicz, C. L. Z., & Fisher, P. J. (2004). Vocabulary lessons. *Educational Leadership*, 61(6), 66–69.
- Blachowicz, C. L. Z., & Fisher, P. J. (2006). *Teaching vocabulary in all classrooms* (3rd ed.). Upper Saddle River, NJ: Pearson.

- Blewitt, P., Rump, K., Shealy, S., & Cook, S. (2009). Shared book reading: When and how questions affect young children's word learning. *Journal of Educational Psychology, 101*(2), 294–304.
- Bowe, F. (2005). *Making inclusion work*. Upper Saddle River, NJ: Pearson.
- Boyle, J. R. (2008). Reading strategies for students with mild disabilities. *Intervention in School and Clinic, 44*(1), 3–9.
- Bravo, M. A., & Cervetti, G. N. (2008). Teaching vocabulary through text and experience in content areas. In A. E. Farstrup & S. J. Samuels (Eds.), *What research has to say about vocabulary instruction* (pp. 130–149). Newark, DE: International Reading Association.
- Bromley, K. (2007). Nine things every teacher should know about words and vocabulary instruction. *Journal of Adolescent & Adult Literacy, 50*(7), 528–536.
- Buikema, J., & Graves, M. (1993). Teaching students to use context cues to infer word meanings. *Journal of Reading, 36*(6), 450–457.
- Butler, S., Urrutia, K., Buenger, A., Gonzalez, N., Hunt, M., & Eisenhart, C. (2010). *A review of the current research on vocabulary instruction: A research synthesis*. Retrieved from <https://www2.ed.gov/programs/readingfirst/support/rmcfinal1.pdf>
- Cain, K. (2007). Deriving word meanings from context: Does explanation facilitate contextual analysis? *Journal of Research in Reading, 30*(4), 347–359.
- Çakici, D. (2017). An overview of metacognitive strategies in reading comprehension skill. *The Journal of Academic Social Science Studies, 57*, 67–82.
- Caldwell, J., & Leslie, L. (2010). Thinking aloud in expository text: Processes and outcomes. *Journal of Literacy Research, 42*(3), 308–340.
- Carlo, M., August, D., McLaughlin, B., & Snow, C. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*(2), 188–189.
- Carlson, E., Jenkins, F., Li, T., & Brownell, M. (2013). The interactions of vocabulary, phonemic awareness, decoding, and reading comprehension. *The Journal of Educational Research, 106*(2), 120–131.
- Carnine, D., Kame'enui, E. J., & Coyle, G. (1984). Utilization of contextual information in determining the meaning of unfamiliar words. *Reading Research Quarterly, 19*(2), 188–204.

- Carr, J., & Bertrando, S. (2012). *Teaching English learners and students with learning difficulties in an inclusive classroom: A guidebook for teachers*. San Francisco, CA: WestEd.
- Chall, J. S., Conard, S. S., & Harris-Sharples, S. (1991). *Should textbooks challenge students?: The case for easier or harder textbooks*. New York, NY: Teachers College Press.
- Chall, J. S., Jacobs, V. A., & Baldwin, L. E. (1990). *The reading crisis: Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Child Trends Databank. (2014). Learning disabilities. Retrieved from <https://www.childtrends.org/?indicators=learning-disabilities>
- Cobb, C., & Blachowicz, C. (2014). *No more "look up the list" vocabulary instruction*. Portsmouth, NH: Heinemann.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd. ed.). New York, NY: Psychology Press.
- Common Core State Standards Initiative. (n.d.). College and career readiness anchor standards for language. Retrieved from <http://www.corestandards.org/ELA-Literacy/CCRA/L/>
- Cook, B. G., & Cook, S. (2013). Unraveling evidence-based practices in special education. *The Journal of Special Education, 47*(2), 71–82.
- Cook, B. G., Gerber, M. M., & Semmel, M. I. (1997). Are effective school reforms effective for all students? - The implications of joint outcome production for school reform. *Exceptionality, 7*(2), 77–96.
- Cook, B. G., Smith, G. J., & Tankersley, M. (2011). Evidence-based practices in education. In S. Graham, T. Urban, & K. Harris (Eds.), *APA educational psychology handbook* (Vol. 3). Washington, DC: American Psychological Association.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson.
- Council for Exceptional Children (CEC). (2014). *Council for Exceptional Children standards for evidence-based practices in special education*. Retrieved from <https://www.cec.sped.org/~media/Images/Standards/CEC%20EBP%20Standards%20cover/CECs%20Evidence%20Based%20Practice%20Standards.pdf>

- Coyne, M., McCoach, D., & Kapp, S. (2007a). Vocabulary intervention for kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disability Quarterly*, *30*(2), 74–88.
- Coyne, M., McCoach, D., Loftus, S., Zipoli, R., & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal*, *110*(1), 1–18.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.
- Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945.
- Cunningham, A. E., Zibulsky, J., Stanovich, K. E., & Stanovich, P. J. (2009). How teachers would spend their time teaching language arts: The mismatch between self-reported and best practices. *Journal of Learning Disabilities*, *42*(5), 418–430.
- Cutting, L. E., Materek, A., Cole, C. A., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia*, *59*(1), 34–54.
- Daalen-Kapteijns, M. van., Schouten-van Parreren, C., & de Glopper, K. (1997). *The training of a word learning strategy: Results in process and product* (Rep. No. 463). Amsterdam, Netherlands: SCO-Kohnstamm Institute.
- Dale, E., & O'Rourke, J. (1981). *The living word vocabulary*. Chicago, IL: World Book-Childcraft International.
- Darling-Hammond, L., Hyster, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Palo Alto, CA: Learning Policy Institute.
- Denton, C. A., & Al Otaiba, S. A. (2011). Teaching word identification to students with reading difficulties and disabilities. *Focus on Exceptional Children*, *43*(7), 1–16.
- Denton, C. A., Enos, M., York, M. J., Francis, D. J., Barnes, M. A., Kulesz, P. A., Fletcher, J. M., ... Carter, S. (2015). Text-processing differences in adolescent adequate and poor comprehenders reading accessible and challenging narrative and informational text. *Reading Research Quarterly*, *50*(4), 393–416.
- Denton, C. A., & Vaughn, S. (2008). Reading and writing intervention for older students with disabilities: Possibilities and challenges. *Learning Disabilities Research & Practice*, *23*(2), 61–62.

- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199.
- Dickinson, D. K., Cote, L., & Smith, M. W. (1993). Learning vocabulary in preschool: Social and discourse contexts affecting vocabulary growth. In C. Daiute & W. Damon (Ed.), *The development of literacy through social interaction: New directions for child and adolescent development, Number 61* (pp. 67–78). San Francisco, CA: Jossey-Bass.
- Drummond, K., Chinen, M., Duncan, T. G., Miller, H. R., Fryer, L., Zmach, C., & Culp, K. (2011). *Impact of the Thinking Reader software program on grade 6 reading vocabulary, comprehension, strategies, and motivation* (NCEE 2010–4035). Washington, DC: National Center for Education Evaluation and Regional Assistance.
- Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a matthew effect. *Journal of Speech, Language, and Hearing Research*, 58(3), 853–864.
- Durkin, D. (1979). What classroom instruction has to say about reading comprehension instruction. *Reading Research Quarterly*, 14(4), 481–533.
- Ebbers, S. M., & Denton C. A. (2008). A root awakening: Vocabulary instruction for older students with reading difficulties. *Learning Disabilities Research & Practice*, 23(2), 90–102.
- Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341–366.
- Elleman, A., Steacy, L., Olinghouse, N., & Compton, D. (2017). Examining child and word characteristics in vocabulary learning of struggling readers. *Scientific Studies of Reading*, 21(2), 133–145.
- Eller, R. G., Pappas, C. C., & Brown, E. (1988). The lexical development of kindergartners: Learning from written context. *Journal of Reading Behavior*, 20(1), 5–24.
- Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly*, 14(2), 174–187.

- Endres, T., & Renkl, A. (2015). Mechanisms behind the testing effect: An empirical investigation of retrieval practice in meaningful learning. *Frontiers in Psychology, 6*, art. 1054.
- Epler-Brooks, P. (2018). *Instructional strategies in general education and putting the Individuals with Disabilities Act (IDEA) into practice*. Hershey, PA: IGI Global.
- Fischer, U. (1994). Learning words from context and dictionaries: An experimental comparison. *Applied Psycholinguistics, 15*(4), 551–574.
- Fisher, D., & Frey, N. (2008). *Better learning through structured teaching: A framework for the gradual release of responsibility*. Alexandria, VA: ASCD.
- Fisher, G. (2001). Cross age tutoring: Alternatives to the reading resource room for struggling adolescent readers. *Journal of Instructional Psychology, 28*(4), 234–240.
- Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. *Annals of Dyslexia, 54*(2), 304–331.
- Fletcher, J. M., Lyon, R. G., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention*. New York, NY: Guilford Press.
- Ford-Connors, E., & Paratore, J. R. (2015). Vocabulary instruction in fifth grade and beyond: Sources of word learning and productive contexts for development. *Review of Educational Research, 85*(1), 50–91.
- Freebody, P., & Anderson, R. C. (1983). Effects of vocabulary difficulty, text cohesion, and schema availability on reading comprehension. *Reading Research Quarterly, 18*(3), 277–294.
- Fresch, M. J. (2008). *An essential history of current reading practices*. Newark, DE: International Reading Association.
- Friend, M. (2014). *Special education: Contemporary perspectives for school professionals* (4th ed.). Upper Saddle River, NJ: Pearson.
- Friend, M. (2017). *Special education: Contemporary perspectives for school professionals* (5th ed.). Upper Saddle River, NJ: Pearson.
- Fukkink, R. G. (2002). Effects of instruction on deriving word meaning from context and incidental word learning. *L1: Educational Studies in Language and Literature, 2*(1), 37–57.

- Fukkink, R. G. (2005). Deriving word meaning from written context: A process analysis. *Learning and Instruction, 15*(1), 23–43.
- Fukkink, R. G., Blok, H., & de Glopper, K. (2001). Deriving word meaning from a written context: A multicomponential Skill. *Language Learning: A Journal of Research in Language Studies, 51*(3), 477–496.
- Fukkink, R. G., & de Glopper, K. (1998). Effects of instruction in deriving word meaning from context: A meta-analysis. *Review of Educational Research, 68*(4), 450–469.
- Gambrell, L. B., Morrow, L. M., & Pressley, M. (2007). *Best practices in literacy instruction* (3rd ed.). New York, NY: Guilford Press.
- Gargiulo, R. M., & Bouck, E. C. (2017). *Special education in contemporary society: An introduction to exceptionality* (6th ed.). Los Angeles, CA: SAGE.
- Gersten, R., Fuchs, L., Williams, J., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 71*(2), 279–320.
- Gilbert, S. L. (2009). *An examination of the effects of computer assisted instruction on reading comprehension* (Master's thesis). Retrieved from ProQuest Dissertations & Theses database. (Accession No. 1465864)
- Goerss, B. L., Beck, I. L., & McKeown, M. G. (1999). Increasing remedial students' ability to derive word meaning from context. *Reading Psychology, 20*(2), 151–175.
- Graves, M. F. (1986). Vocabulary learning and instruction. In E. Z. Rothkopf & L. C. Ehri (Eds.), *Review of research in education, Volume 13* (pp. 49–89). Washington, DC: American Educational Research Association.
- Graves, M. F. (2006). *The vocabulary book: Learning and instruction*. New York, NY: Teachers College Press.
- Graves, M. F. (2016). *The vocabulary book: Learning and instruction* (2nd ed.). New York, NY: Teachers College Press.
- Graves, M. F., August, D., & Mancilla-Martinez, J. (2012). *Teaching vocabulary to English language learners*. New York, NY: Teachers College Press.
- Graves, M. F., Juel, C. F., Graves, B. B., & Dewitz, P. (2011). *Teaching reading in the 21st century: Motivating all learners* (5th ed.). Boston, MA: Pearson.

- Gunning, T. (2013). *Creating literacy instruction for all students* (8th ed.). Boston, MA: Pearson.
- Hafner, L. E. (1965). A one-month experiment in teaching context aids in fifth grade. *Journal of Educational Research* 58(10), 472–474.
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–334.
- Harris, M., Schumaker, J., & Deshler, D. (2011). The effects of strategic morphological analysis instruction on the vocabulary performance of secondary students with and without disabilities. *Learning Disability Quarterly*, 34(1), 17–33.
- Hasselbring, T. S., & Bausch, M. E. (2006). Assistive technologies for reading. *Educational Leadership*, 63(4), 72–75.
- Helman, A. L., Calhoun, M. B., & Kern, L. (2015). Improving the science vocabulary of English language learners with reading disabilities. *Learning Disabilities Quarterly*, 38(1), 40–52.
- Henchy, A. M. (2013). *Review and evaluation of reliability generalization research* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (Accession No. 1566489385)
- Hennebry, M., Rogers, V., Macaro, E., & Murphy, V. (2017). Direct teaching of vocabulary after listening: Is it worth the effort and what method is best? *Language Learning Journal*, 45(3), 282–300.
- Herman, P. A., & Dole, J. (1988). Theory and practice in vocabulary learning and instruction. *The Elementary School Journal*, 89(1), 43–54.
- Higgins, E. L., Raskind, M. H., Goldberg, R. J., & Herman, K. L. (2002). Stages of acceptance of a learning disability: The impact of labeling. *Learning Disability Quarterly*, 25(1), 3–18.
- Hirsch, J. E. D. (2003). Reading comprehension requires knowledge—of words and the world: Scientific insights into the fourth-grade slump and stagnant reading comprehension. *American Educator*, 27(1), 10–31.
- Honig, B., Gutlohn, L., & Diamond, L. (2012). *Teaching Reading Sourcebook* (2nd ed.). Novato, CA: Arena Press.

- Howell, D. (2013). *Statistical methods for psychology* (8th ed.). Belmont, CA: Wadsworth Cengage Learning.
- Hu, H. C. M. (2013). The effects of word frequency and contextual types on vocabulary acquisition from extensive reading: A case study. *Journal of Language Teaching and Research, 4*(3), 487–495.
- Hu, H. C. M., & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System, 45*(1), 27–38.
- Hu, M., & Nation, I. S. P. (2000). Vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403–430.
- Huckin, T. N., & Zhendong, J. (1986). Inferring word-meaning from context: A study in second language acquisition. *Proceedings of the Eastern States Conference on Linguistics, 3*(4), 271–280.
- Hudson, R. F., Isakson, C., Richman, T., Lane, H. B., & Arriaza-Allen, S. (2011). An examination of a small-group decoding intervention for struggling readers: Comparing accuracy and automaticity criteria. *Learning Disabilities Research & Practice, 26*(1), 15–27.
- Hughes, C., Golas, M., Cosgriff, J., Brigham, N., Edwards, C., & Cashen, K. (2011). Effects of a social skills intervention among high school students with intellectual disabilities and autism and their general education peers. *Research & Practice for Persons with Severe Disabilities, 36*(1/2), 46–61.
- Individuals with Disabilities Education Improvement Act of 2004, P.L. 108-446, 20 U.S.C. § 1400 et seq.
- Irvin, J. (2001). Assisting struggling readers in building vocabulary and background knowledge. *Voices from the Middle, 8*(4), 37–43.
- Iwai, Y. (2016). The effect of explicit instruction on strategic reading in a literacy methods course. *International Journal of Teaching and Learning in Higher Education, 28*(1), 110–118.
- Izquierdo, I., Olea, J., & Abad, F. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema, 26*(3), 395–400.
- Jenkins, J. R., Matlock, B., & Slocum, T. A. (1989). Two approaches to vocabulary instruction: The teaching of individual word meanings and practice in deriving word meaning from context. *Reading Research Quarterly 24*(2), 215–235.

- Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*(3), 299–322.
- Jitendra, A. K., & Gajria, M. (2011). Reading comprehension instruction for students with learning disabilities. *Focus on Exceptional Children, 43*(8), 1–16.
- Johnson, K. M., & McCabe, P. P. (2003). Review of the Gates-MacGinitie 4th edition, forms S and T. In B. S. Plake, J. C. Impara, & R. S. Spies (Eds.), *Fifteenth mental measurements yearbook*. Lincoln, NE: Buros Center for Testing.
- Joseph, L., & Schisler, R. (2009). Should adolescents go back to the basics? *Remedial and Special Education, 30*(3), 131–147.
- Joshi, R. M. (2005). Vocabulary: A critical component of comprehension. *Reading & Writing Quarterly, 21*(3), 209–219.
- Joyce, B., & Showers, B. (1982). The coaching of teaching. *Educational Leadership, 40*(1), 4–10.
- Justice, L. M., Meier, J., & Walpole, S. (2005). Learning new words from storybooks: An efficacy study with at-risk kindergartners. *Language, Speech & Hearing Services in Schools, 36*(1), 17–32.
- Kaluger, G., & Kolson, C. J. (1978). *Reading and learning disabilities*. Columbus, OH: Merrill.
- Kame'enui, E. J., & Baumann, J. F. (2012). *Vocabulary instruction: Research to practice* (2nd ed.). New York, NY: Guilford Press.
- Kamil, M. L., Borman, G. D., Dole, J., Kral, C. C., Salinger, T., & Torgesen, J. (2008). *Improving adolescent literacy: Effective classroom and intervention practices: A practice guide* (NCEE 2008-4027). Washington, DC: National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/adlit_pg_082608.pdf
- Karbalaei, A., Amoli, F. A., & Tavakoli, M. M. (2012). The effects of explicit teaching of context clues at undergraduate level in EFL and ESL context. *European Online Journal of Natural and Social Sciences, 1*(3), 68–77.
- Kavale, K. A., & Forness, S. R. (1996). Social skill deficits and learning disabilities: A meta-analysis. *Journal of Learning Disabilities, 29*(3), 226–237.

- Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing, 21*(8), 783–804.
- Kieffer, M. J., & Lesaux, N. K. (2012). Development of morphological awareness and vocabulary knowledge in Spanish-speaking language minority learners: A parallel process latent growth curve model. *Applied Psycholinguistics, 33*(1), 23–54.
- Kim, Y. S., Wagner, R. K., & Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading, 15*(4), 338–362.
- King-Sears, M., Swanson, C., & Mainzer, L. (2011). TECHnology and literacy for adolescents with disabilities. *Journal of Adolescent & Adult Literacy, 54*(8), 569–578.
- Klinger, J. K., Vaughn, S., Arguelles, M. E., Hughes, M. T., & Leftwich, S. A. (2013). Collaborative strategic reading: “Real-world” lessons from classroom teachers. *Remedial and Special Education, 25*(5), 291–302.
- Knight, V. F., Browder, D., Angello, B., & Lee, A. (2010). Academic instruction for students with severe disabilities. *Focus on Exceptional Children, 42*(1), 1–14.
- Kranzer, K. G. (1988). *A study of the effects of instruction on incidental word learning and on the ability to derive word meaning from context* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (Accession No. 303721192)
- Kratochwill, T., Hitchcock, J., Horner, R., Levin, J., Odom, S., Rindskopf, D., & Shadish, W. (2013). Single-Case intervention research design standards. *Remedial and Special Education, 34*(1), 26–26.
- Kucan, L., & Beck, I. (1997). Thinking aloud and reading comprehension research: Inquiry, instruction, and social interaction. *Review of Educational Research, 67*(3), 271–299.
- Kuhn, M. R. (2005). A comparative study of small group fluency instruction. *Reading Psychology, 26*(2), 127–146.
- Kuhn, M. R., & Stahl, S. A. (1998). Teaching children to learn word meanings from context: A synthesis and some questions. *Journal of Literacy Research, 30*(1), 119–138.

- Laerd Statistics (2015). Three-way mixed ANOVA (with one within-subjects and two between-subjects factors) using SPSS Statistics. *Statistical tutorials and software guides*. Retrieved from <https://statistics.laerd.com/>
- Laing, S., & Kamhi, A. (2002). The use of think-aloud protocols to compare inferencing abilities in average and below-average readers. *Journal of Learning Disabilities, 35*(5), 437–448.
- Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans to thinking machines* (pp. 316–323). Clevedon, England: Multilingual Matters.
- Learning Disabilities Association of America. (n.d.). History. Retrieved from <https://ldaamerica.org/about-us/history/>
- Lee, A. M. I. (2018). Instructional intervention: What you need to know. Retrieved from <https://www.understood.org/en/learning-attention-issues/treatments-approaches/educational-strategies/instructional-intervention-what-you-need-to-know>
- Lee, J., Grigg, W., & Donahue, P. (2007). *The nation's report card: Reading 2007* (NCES-2007-496). Washington, DC: National Center for Education Statistics. Retrieved from <https://nces.ed.gov/nationsreportcard/pubs/main2007/2007496.asp>
- Leko, M. M., & Brownell, M. T. (2009). Crafting quality professional development for special educators: What school leaders should know. *Teaching Exceptional Children, 42*(1), 64–70.
- Lenhart, J., Lenhard, W., Vaahtoranta, E., & Suggate, S. (2018). Incidental vocabulary acquisition from listening to stories: A comparison between read-aloud and free storytelling approaches. *Educational Psychology, 38*(5), 596–616.
- Leong, C. K., Simmons, D. R., & Izatt-Gambell, M. (1990). The effect of systematic training in elaboration on word meaning and prose comprehension in poor readers. *Annals of Dyslexia, 40*(1), 192–215.
- Lerner, J. W. (2000). *Learning disabilities: Theories, diagnosis, and teaching strategies*. Boston, MA: Houghton Mifflin.
- Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal, 47*(3), 596–632.

- Leung, C. (2008). Preschoolers' acquisition of scientific vocabulary through repeated read-aloud events, retellings, and hands-on science activities. *Reading Psychology, 29*(2), 165–193.
- Lewis, J. (2009). *Essential questions in adolescent literacy: Teachers and researchers describe what works in classrooms*. New York, NY: Guilford Press.
- Lipson, M. Y., & Lang, L. B. (1991). Not as easy as it seems: Some unresolved questions about fluency. *Theory into Practice, 30*(3), 218–227.
- Lipson, M. Y., & Wixson, K. K. (1991). *Assessment & instruction of reading disability: An interactive approach*. New York, NY: HarperCollins.
- Lockavitch, J. (2010, July 15). Ten critical facts from vocabulary research. *Failure Free Reading*. Retrieved from <https://www.failurefreeonline.com/educator/>
- Loftus, S. M., & Coyne, M. D. (2013). Vocabulary instruction within a multi-tier approach. *Reading and Writing Quarterly, 29*(1), 4–19.
- López, L. M., & Greenfield, D. B. (2004). The cross-language transfer of phonological skills of Hispanic head start children. *Bilingual Research Journal, 28*(1), 1–18.
- López, P., Torrance, M., Rijlaarsdam, G., & Fidalgo, R. (2017). Effects of direct instruction and strategy modeling on upper-primary students' writing development. *Frontiers in Psychology, 8*(1), 1–10.
- Lublimer, S., & Smetana, L. (2005). The effects of comprehensive vocabulary instruction on Title I students' metacognitive word-learning skills and reading comprehension. *Journal of Literacy Research, 37*(2), 163–200.
- Lyon, G. R., & Chhabra, V. (2004). The science of reading research. *Educational Leadership, 61*(6), 13–17.
- MacGinitie, W. H., MacGinitie, R. K., Maria, K., Dreyer, L. G., & Hughes, K. E. (2008). *Manual for scoring and interpretation: Levels 4–6 (Forms S and T)* (4th ed.). Rolling Meadows, IL: Riverside.
- Mainali, B. P. (2013). Higher order thinking in education. *Academic Voices: A Multidisciplinary Journal, 2*(1), 5–10.
- Malmgren, K. W., & Trezek, B. J. (2009). Literacy instruction for secondary students with disabilities. *Focus on Exceptional Children, 41*(6), 1–12.

- Manzo, A. V., Manzo, U. C., & Thomas, M. M. (2006). Rationale for systematic vocabulary development: Antidote for state mandates. *Journal of Adolescent & Adult Literacy*, 49(7), 610–619.
- Maria, K., & Hughes, K. E. (2008). *Gates-MacGinitie reading tests - Technical report supplement (Forms S and T, 2005–2006 Renorming)* (4th Ed.). Rolling Meadows, IL: Riverside.
- Marksheffel, N. (1966). *Better reading in the secondary school: Principles and procedures for teachers*. New York, NY: Ronald Press.
- Martin-Chang, S., Levy, B., & O’Neil, S. (2007). Word acquisition, retention, and transfer: Findings from contextual and isolated word training. *Journal of Experimental Child Psychology*, 96(1), 37–56.
- Matsuda, L. G. (1987). *A historical examination of vocabulary development in the schools, 1900–1985* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (Accession No. 303469928)
- May, A. L., & Stone, C. A. (2010). Stereotypes of individuals with learning disabilities: Views of college students with and without learning disabilities. *Journal of Learning Disabilities*, 43(6), 483–499.
- McCambridge, J., Witton, J., & Elbourne, D. (2014). Systematic review of the Hawthorne Effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, 67(3), 267–77.
- McClintock, B., Pesco, D., & Martin-Chang, S. (2014). Thinking aloud: Effects on text comprehension by children with specific language impairment and their peers. *International Journal of Language & Communication Disorders*, 49(6), 637–648.
- Melekoglu, M. A. (2011). Impact of motivation to read on reading gains for struggling readers with and without learning disabilities. *Learning Disability Quarterly*, 34(4), 248–261.
- Miller, G. A., & Gildea, P. M. (1987). How children learn words. *Scientific American*, 257(3), 94–99.
- Milsom, A., & Glanville, J. L. (2010). Factors mediating the relationship between social skills and academic grades in a sample of students diagnosed with learning disabilities or emotional disturbance. *Remedial & Special Education*, 31(4), 241–251.

- Misulis, K. (1999). Making vocabulary development manageable in content instruction. *Contemporary Education*, 70(2), 25–30.
- Moats, L. C. (2005). *Language essentials for teachers of reading and spelling: The speech sounds of English: Phonetics, phonology, and phoneme awareness*. Longmont, CA: Sopris West Educational Services.
- Mondria, J. A., & Wit-De Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, 12(3), 249–267.
- Montelongo, J., Hernández, A., Herter, R., & Cuello, J. (2011). Using cognates to scaffold context clue strategies for Latino ELS. *The Reading Teacher*, 64(6), 429–434.
- Nagy, W. E. (2005). Why vocabulary instruction needs to be long-term and comprehensive. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 27–44). Mahwah, NJ: Erlbaum.
- Nagy, W. E., & Anderson, R. (1984). How many words are there in printed school English? *Reading Research Quarterly*, 19(3), 304–330.
- Nagy, W. E., Anderson, R. C., & Herman, P. A. (1987). Learning word meanings from context during normal reading. *Reading Research Quarterly*, 24(2), 237–270.
- Nagy, W., Berninger, V. W., & Abbott, R. D. (2006). Contributions of morphology beyond phonology to literacy outcomes of upper elementary and middle school students. *Educational Psychology*, 98, 134–147.
- Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19–36). Hillsdale, NJ: Erlbaum.
- Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 269–284). Mahwah, NJ: Erlbaum.
- Nash, H., & Snowling, M. (2006). Teaching new words to children with poor existing vocabulary knowledge: A controlled evaluation of the definition and context methods. *International Journal of Language and Communication Disorders*, 41(3), 335–354.
- National Assessment of Educational Progress. (2017). *The condition of education: 2017*. Washington, DC: Author.

- National Center for Education Statistics. (2018a). Children and youth with disabilities. Washington, DC: National Assessment of Educational Progress (NAEP). Retrieved from https://nces.ed.gov/programs/coe/indicator_cgg.asp
- National Center for Education Statistics. (2018b). Reading performance. Washington, DC: NAEP. Retrieved from https://nces.ed.gov/programs/coe/indicator_cnb.asp
- National Center for Learning Disabilities (NCLD). (2013). High school graduation. Retrieved from <http://www.nclld.org/images/content/files/hill-briefs/hs-grad-brieftemplate.pdf>
- National Governors Association Center for Best Practices, & Council of Chief State School Officers. (2010). *Common core state standards for English language arts standards: Kindergarten–grade 12*. Retrieved from <http://www.corestandards.org/ELA-Literacy/>
- National Joint Committee on Learning Disabilities (NJCLD). (2008). Adolescent literacy and older students with learning disabilities. *Learning Disability Quarterly*, 31(4), 211–218.
- National Reading Panel (2000). *Report of the National Reading Panel: Teaching children to read*. Bethesda, MD: National Institute of Child Health and Human Development. Retrieved from http://www.nichd.nih.gov/publications/nrp/upload/smallbook_pdf.pdf
- Nelson, N. W., & Van Meter, A. M. (2006). *Finding the words: Vocabulary development for young authors*. Austin, TX: Pro-Ed.
- Nelson, R. N., & Stage, S. A. (2007). Fostering the development of vocabulary knowledge and reading comprehension through contextually-based multiple meaning vocabulary instruction. *Education & Treatment of Children*, 30(1), 1–22.
- Nippold, M. A. (2002). Lexical learning in school-age children, adolescents, and adults: A process where language and literacy converge. *Journal of Child Language*, 29(2), 474–478.
- No Child Left Behind Act of 2001, P.L. 107-110, 20 U.S.C. § 6319 (2002).
- Noordzij, M., Tripepi, G., Dekker, F., Zoccali, C., Tanck, M., & Jager, K. (2010). Sample size calculations: Basic principles and common pitfalls. *Nephrology Dialysis Transplantation*, 25(5), 1388–1393.

- North Dakota Department of Public Instruction. (2010). *Fidelity of implementation*. Retrieved from <https://www.nd.gov/dpi/uploads/60/June8Section3SpecialtopicsRevised3.18.13.docx.pdf>
- Oakhill, J., Cain, K., & Elbro, C. (2015). *Understanding and teaching reading comprehension: A handbook*. New York, NY: Routledge.
- O'Donnell, C. L., & Lynch, J. L. (2008, March). *Fidelity of implementation to instructional strategies as a moderator of science curriculum unit effectiveness*. Paper presented at the annual meeting of the 2008 National Association of Research in Science Teaching. Baltimore, MD.
- Ogle, D., & Blachowicz, C. L. Z. (2002). Beyond literature circles: Helping students comprehend informational texts. In C. C. Block & M. Pressley (Eds.), *Comprehension instruction: Research-based best practices* (pp. 259–274). New York, NY: Guilford.
- O'Neill, R. E., McDonnell, J. J., Billingsley, F. F., & Jenson, W. R. (2011). *Single case research designs in educational and community settings*. Upper Saddle River, NJ: Pearson.
- Oswald, D., Sherratt, F., & Smith, S. (2014). Handling the Hawthorne effect: The challenges surrounding a participant observer. *Review of Social Studies, 1*(1), 53–73.
- Patberg, J. P., Graves, M. F., & Stibbe, M. A. (1984). Effects of active teaching and practice in facilitating students' use of context clues. In J. A. Niles & L. A. Harris (Eds.), *Changing perspectives in research in reading/language processing and instruction* (pp. 146–151). Rochester, NY: National Reading Conference.
- Patberg, J. P., & Stibbe, M. A. (1985, December). *The effects of contextual analysis instruction on vocabulary learning*. Paper presented at the annual meeting of the National Reading Conference, San Diego, CA.
- Pierce, C. A., Block, R. A., & Aguinis, H. (2004). Cautionary note on reporting Eta-Squared values from Multifactor ANOVA designs. *Educational and Psychological Measurement, 64*(6), 916–924.
- Protheroe, N. (2008). The impact of fidelity of implementation in effective standards-based instruction. *Principal, 88*(1), 38–41.

- Pullen, P. C., Lane, H. B., Ashworth, K. A., & Lovelace, S. P. (2011). Learning disabilities. In J. M. Kauffman & D. P. Hallahan (Eds.). *The handbook of special education* (pp. 187–197). New York, NY: Routledge.
- Pullen, P. C., Tuckwiller, E. D., Konold, T. R., Maynard, K. L., & Coyne, M. D. (2010). A tiered intervention model for early vocabulary instruction: The effects of tiered instruction for young students at risk for reading disability. *Learning Disabilities Research & Practice, 25*(3), 110–123.
- RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND Education.
- RAND Reading Study Group. (2004). A research agenda for improving reading comprehension. In R. B. Ruddell & N. J. Unrau (Eds.), *Theoretical models and processes of reading* (pp. 720–754). Newark, DE: International Reading Association.
- Reed, D. K., Petscher, Y., & Foorman, B. R. (2016). The contribution of vocabulary knowledge and spelling to the reading comprehension of adolescents who are and are not English language learners. *Reading and Writing, 29*(4), 633–657.
- Richek, M. A. (2006). It's easy, it's fun...it's vocabulary. *New England Reading Association Journal, 42*(2), 19–23.
- Roser, N., & Juel, C. (1982). Effects of vocabulary instruction on reading comprehension. In J. S. Niles & L. A. Harris (Eds.), *New inquiries into reading research. Thirty-first yearbook of the National Reading Conference*. (pp. 110–118). Albany, NY: National Reading Conference.
- Rowe, M., Ozuru, Y., O'Reilly, T., & McNamara, D. S. (2008). Where's the difficulty in standardized reading tests: The passage or the question? *Behavior Research Methods, 40*(4), 1001–1015.
- Rupley, W. H. (2009). Introduction to direct/explicit instruction in reading for the struggling reader: Phonemic awareness, phonics, fluency, vocabulary, and comprehension. *Reading & Writing Quarterly: Overcoming Learning Difficulties, 25*(2/3), 119–124.
- Rupley, W. H., Blair, T. R., & Nichols, W. D. (2009). Effective reading instruction for struggling readers: The role of direct/explicit teaching. *Reading & Writing Quarterly, 25*(2/3), 125–138.
- Rupley, W. H., Logan, J. W., & Nichols, W. D. (1998). Vocabulary instruction in a balanced reading program. *Reading Teacher, 52*(4), 336–346.

- Rupley, W. H., & Nichols, W. D. (2005). Vocabulary instruction for the struggling reader. *Reading & Writing Quarterly*, 21(3), 239–260.
- Schatz, E., & Baldwin, R. (1986). Context clues are unreliable predictors of word meanings. *Reading Research Quarterly*, 21(4), 439–453.
- Sénéchal, M., & Cornell, E. M. (1993). Vocabulary acquisition through shared reading experiences. *Reading Research Quarterly*, 28(4), 360–374.
- Shippen, M. E., Houchins, D. E., Steventon, C., & Sartor, D. (2005). A comparison of two direct instruction reading programs for urban middle school students. *Remedial and Special Education*, 26(3), 175–182.
- Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences of the United States of America*, 113(10), 2621–2624.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational Researcher*, 31(7), 15–21.
- Smith, R. W., Osborne, L. T., Crim, D., & Rhu, A. H. (1986). Labeling theory as applied to learning disabilities: Survey findings and policy suggestions. *Journal of Learning Disabilities*, 19(4), 195–202.
- Spencer, M., Quinn, J., & Wagner, R. (2014). Specific reading comprehension disability: Major problem, myth, or misnomer? *Learning Disabilities Research & Practice*, 29(1), 3–9.
- Stahl, K. A., & Bravo, M. A. (2010). Contemporary classroom vocabulary assessment for content areas. *The Reading Teacher*, 63(7), 566–578.
- Stahl, S. A. (2005). Four problems with teaching word meanings (and what to do to make vocabulary an integral part of instruction). In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 95–115). Mahwah, NJ: Erlbaum.
- Stahl, S. A., & Fairbanks, M. M. (1986). The effects of vocabulary instruction: A model-based meta-analysis. *Review of Educational Research*, 56(1), 72–110.
- Stahl, S. A., & Hayes, D. (2013). *Instructional models in reading*. Hoboken, NJ: Taylor and Francis.
- Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings*. Mahwah, NJ: Erlbaum.

- Stahl, S. A., & Shiel, T. G. (1999). Teaching meaning vocabulary: Productive approaches for poor readers. In California State Board of Education (Ed.), *Read all about it! Readings to inform the profession* (pp. 291–321). Sacramento, CA: Editor.
- Stanovich, K. E. (1986). Has the learning disabilities field lost its intelligence? *Journal of Learning Disabilities*, 22(8), 487–492.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89–105). Hillsdale, NJ: Erlbaum.
- Sternberg, R. J., & Powell, J. S. (1983). Comprehending verbal communication. *American Psychologist*, 38(8), 878–893.
- Stillwell, R., Sable, J., & Plotts, C. (2011). *Public school graduates and dropouts from the common core of data: School year 2008–09* (NCES 2011-312). Washington, DC: National Center for Education Statistics. Retrieved from <http://nces.ed.gov/pubsearch>
- Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A metaanalysis. *Review of Educational Research*, 69(3), 261–285.
- Swanson, E. A., & Vaughn, S. (2010). An observation study of reading instruction provided to elementary students with learning disabilities in the resource room. *Psychology in the Schools*, 47(5), 481–492.
- Swanson, H. L., Mink, J., & Bocian, K. M. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: More alike than different? *Journal of Educational Psychology*, 91(2), 321–333.
- Sze, S. (2009). The effects of assistive technology on students with disabilities. *Journal of Educational Technology Systems*, 37(4), 419–429.
- Teng, F. (2016). Incidental vocabulary acquisition from reading-only and reading-while-listening: A multi-dimensional approach. *Innovation in Language Learning and Teaching*, 12(3), 274–288.
- Tilstra, J., McMaster, K., Broek, P. V. D., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading*, 32(4), 383–401.
- Tolman, C. (2005). Working smarter, not harder: What teachers of reading need to know and be able to teach. *Perspectives*, 31(4), 15–23.

- Tomesen, M., & Aarnoutse, C. (1998). Effects of an instructional programme for deriving word meanings. *Educational Studies* 24(1), 107–128.
- Torgesen, J. K. (2002). The prevention of reading difficulties. *Journal of School Psychology*, 40(1), 7–26.
- Torgesen, J. K., Houston, D. D., Rissman, L. M., Decker, S. M., Roberts, G., & Vaughn, S. (2007). *Academic literacy instruction for adolescents: A guidance document from the Center on Instruction*. Portsmouth, NH: RMC Research Corporation.
- University of Oregon Center on Teaching and Learning (n.d.). Big ideas in beginning reading. Retrieved from http://reading.uoregon.edu/big_ideas/
- Van Merriënboer, J., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies. *Medical Education*, 44(1), 85–93.
- Vaughn, S., Gersten, R., & Chard, D. (2000). The underlying message in LD intervention research: Findings from research syntheses. *Exceptional Children*, 67(1), 99–114.
- Vaughn, S., Wanzek, J., Wexler, J., Barth, A., Cirino, P. T., Fletcher, J. M., & Francis, D. J. (2010). The relative effects of group size on reading progress of older students with reading difficulties. *Reading and Writing: An Interdisciplinary Journal*, 23(8), 931–956.
- Wagner, R. K., & Meros, D. (2010). Vocabulary and reading comprehension: Direct, indirect, and reciprocal influences. *Focus on Exceptional Children*, 43(1), 1–10.
- Walters, J. (2004). Teaching the use of context to infer meaning: A longitudinal survey of 11 and 12 vocabulary research. *Language Teaching*, 37(4), 243–252.
- Wei, X., Blackorby, J., & Schiller, E. (2011). Growth in reading achievement of students with disabilities, ages 7 to 17. *Exceptional Children*, 78(1), 89–106.
- Weiser, B. (2013). Effective vocabulary instruction for kindergarten to 12th grade students experiencing learning disabilities. *Council for Learning Disabilities*. Retrieved from <https://council-for-learning-disabilities.org/effective-vocabulary-instruction-for-kindergarten-to-12th-grade-students-experiencing-learning-disabilities>
- White, T., Graves, M., & Slater, W. (1990). Growth of reading vocabulary in diverse elementary schools: Decoding and word meaning. *Journal of Educational Psychology*, 82(2), 281–290.

- Whitworth, B. A., & Chiu, J. L. (2015). Professional development and teacher change: The missing leadership link. *Journal of Science Teacher Education, 26*(2), 121–137.
- Wiersma, W., & Jurs, S. G. (2009). *Research methods in education: An introduction* (9th ed.). Boston, MA: Pearson.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. London, England: Edward Arnold.
- Williams, J. P., Hall, K. M., & Lauer, K. D. (2004). Teaching expository text structure to young at-risk learners: Building the basics of comprehension instruction. *Exceptionality, 12*(3), 129–144.
- Williams, J. P., Nubla-Kung, A. M., Pollini, S., Stafford, K. B., Garcia, A., & Snyder, A. E. (2007). Teaching cause-effect text structure through social studies content to at-risk second graders. *Journal of Learning Disabilities, 40*(2), 111–120.
- Williams, J. P., Stafford, K. B., Lauer, K. D., Hall, K. M., & Pollini, S. (2009). Embedding reading comprehension training in content-area instruction. *Journal of Educational Psychology, 101*(1), 1–20.
- Wong, B. Y. L., Harris, K. R., Graham, S., & Butler, D. L. (2003). Cognitive strategies instruction research in learning disabilities. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 383–402). New York, NY: Guilford.
- Woolfolk, A. (2010). *Educational Psychology* (11th ed.). Upper Saddle River, NJ: Pearson.
- Wright, W. D., & Wright, P. D. (2009). *Special Education Law*. Hartfield, VA: Harbor House Law Press.
- Yap, K., Aldersebaes, I., Railsback, J., Shaughnessy, J., & Speth, T. (2000). *Evaluating whole-school reform efforts: A guide for district and school staff* (2nd ed.). Portland, OR: Northwest Regional Educational Laboratory.
- Yuen, J. (2009). *Context clue detectives: Empowering students with a self-learning device through the teaching of context clues* (Unpublished master's thesis). University of California.
- Zepeda, C., Richey, J., Ronevich, P., & Nokes-Malach, T. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An In Vivo study. *Journal of Educational Psychology, 107*(4), 954–954.

Zikmund-Fisher, B., Fagerlin, A., & Ubel, P. (2010). A demonstration of “less can be more” in risk graphics. *Medical Decision Making: An International Journal of the Society for Medical Decision Making*, 30(6), 661–71.

APPENDIX A

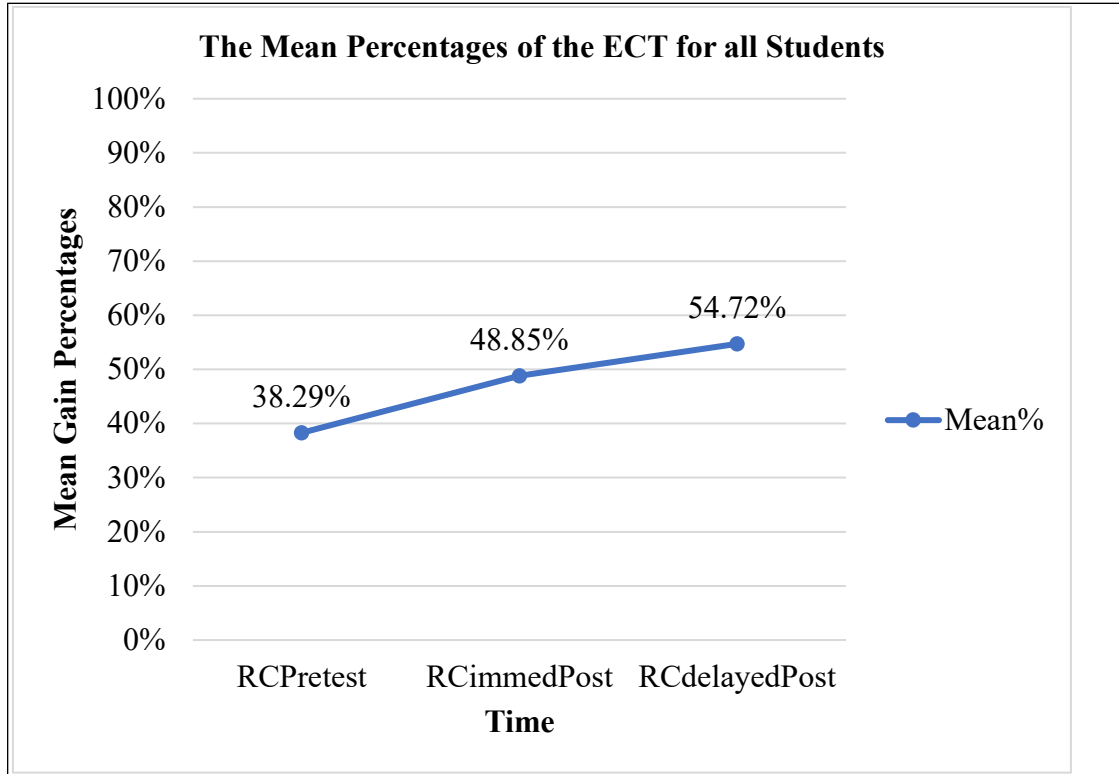
OVERVIEW OF THE PILOT STUDY RESULTS, EFFECT SIZE, AND REQUIRED SAMPLE SIZE

Overview of the Pilot Study Results

The Experimenter-Constructed Tests

All student participants in the pilot study were exposed to an immediate posttest to compare their performance before and after receiving the proposed vocabulary instructional intervention. When analyzing the results of the experimenter-constructed test, there was an increase in the performance of the students after receiving the vocabulary intervention. Figure A.1 demonstrates that when looking at the mean percentages of the experimenter-constructed test for all student participants, the results of the experimenter-constructed immediate posttest showed a dramatic increase of more than 10% (38.29% to 48.85%) when students were exposed to the vocabulary intervention for 20 to 30 minutes as part of their 50-minute social study lessons for *only* five consecutive days. To examine the retention of the vocabulary strategy, students took long-term follow up tests three weeks after the immediate posttest was completed. Unexpectedly, the results of the experimenter-constructed delayed posttest showed that the students not only retained the context clues strategy, but their performance increased by approximately 6% (see Figure A.1). It is worth mentioning that the total student involvement in the pilot study was *only* 6 hours and 35 minutes spread over the six-week time period.

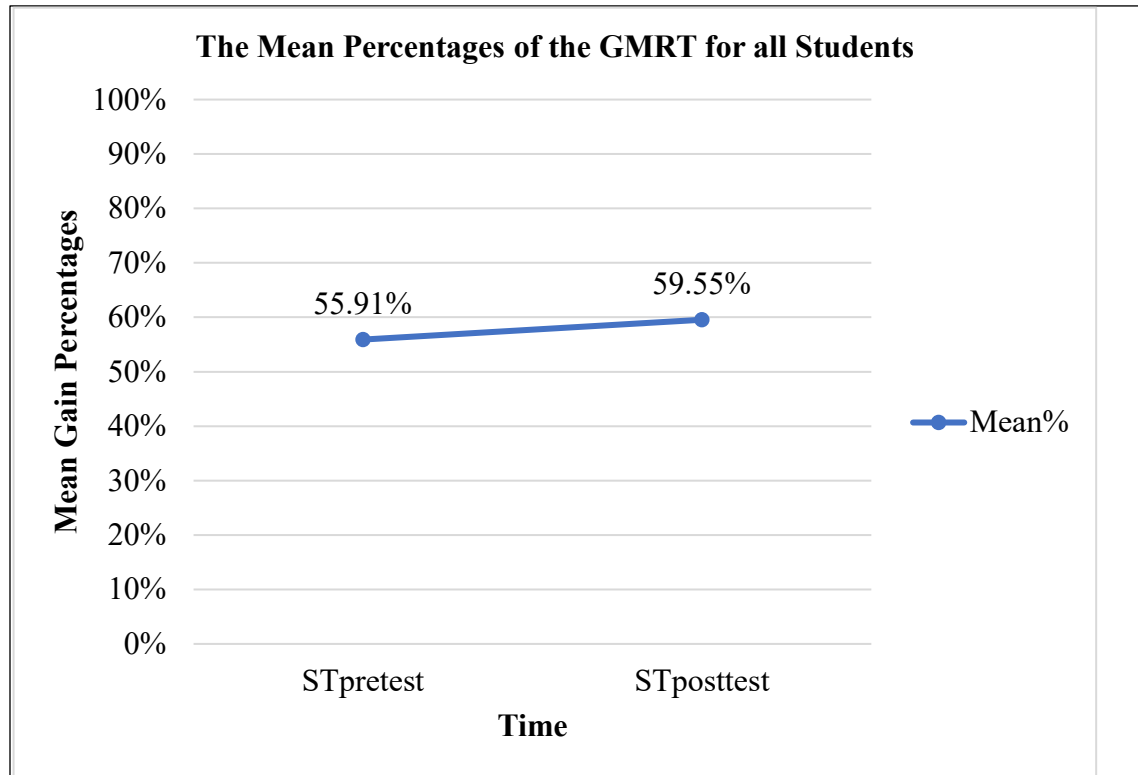
Figure A.1 The Mean Percentages of the Experimenter-Constructed Tests for All Students



The Standardized Test (GMRT)

This standardized test is designed to assess student reading levels throughout the course of their education. The Gates-MacGinitie Reading Tests (GMRT), Fourth Edition, was used in the delayed posttest after the immediate posttest in this pilot study to see if there was any growth in the students' reading ability, specifically in vocabulary. When conducting Level 4 of the GMRT, the results of the mean percentages of the standardized test for all students revealed that there was an increase (4%) in their reading ability, both in vocabulary and comprehension (see Figure A.2). Although this reading growth was small, achieving this increase in such a short time (five weeks) was a crucial finding.

Figure A.2 The Mean Percentages of the Standardized Tests for All Students



Effect Size

When using the *a priori* statistical power (with one tail, power = 0.80, alpha = 0.05) for detecting a significant difference and finding the effect size (the statistical power) of the vocabulary instructional intervention in the pilot study, the results showed that the effect size of the vocabulary intervention was Cohen's $d = 0.79$. This is considered a moderate effect size (Cohen, 2009).

***Important Note:** According to Cohen's d :

$d = 0.20$ is a small effect

$d = 0.50$ is a moderate effect

$d = 0.80$ is a large effect

The Required Sample Size for the Present/Formal Study

The results of the pilot study indicated a Cohen's d effect size of 0.79 with a corresponding eta-squared effect size of 0.138. This is considered a moderate effect size for the proposed intervention on the performance of students without disabilities (Cohen, 2009). Subsequently, a power analysis using this effect size applied to a three way-mixed ANOVA with two between-subjects factors and one within-subject factor with three time points indicated that a minimum sample size of 40 participants will result in power of 0.80, at the nominal alpha level of .05.

APPENDIX B

AN EXAMPLE OF THE LESSON PLANS USED IN THE VOCABULARY INSTRUCTIONAL INTERVENTION

DAY 5: Antonym and Synonym Context Clues

Standards and Key Concepts	
Unit of Study	Government Functions (Week 10)
Grade Level	Fourth Grade
Essential Vocabulary	commits, fine, govern, rules
Objective(s)	<ul style="list-style-type: none"> ▪ Learn the Skill – Students will learn how to use the antonym and synonym context clues to understand unfamiliar words. ▪ Practice the Skill – Students will practice using the antonym and synonym context clues with several sentences and paragraphs to understand unfamiliar words.
Materials Needed	<ul style="list-style-type: none"> • Whiteboard or smart board • Computer • PowerPoint presentation (Slides 25 - 34) • Classroom projector • Markers • Project R-3 + R-4 • A-4 & 7 • Copies of R-3 + R-4 + A-4 & 7
Lesson Implementation	
I Do (Teacher-Led Practice)	<p><u>PowerPoint Content:</u> The presentation includes the definitions and examples of the antonym and synonym context clues. Also included are several practice sentences and paragraphs that can be used as additional practice with the class as guided or independent practice.</p>

	<p>Note:</p> <ul style="list-style-type: none"> - Teacher should talk through each slide in detail. Enrich the presentation with your own teaching of the material. - Make sure that each student has copies of R-3 and R-4. It will be referenced to throughout the PowerPoint presentation. <ul style="list-style-type: none"> ➤ Project R-3: Use it as a reference for each context clue type. Focus on the signal words. ➤ Project R-4: Use it as a reference for the steps of how to use each context clue. - Slides 25–26 & 30–31: Teacher should model the exercises in these slides about the antonym and synonym context clues types by highlighting the signal words and how meaning was attained using the “Think Aloud” strategy.
<p>We Do (Guided Practice)</p>	<ul style="list-style-type: none"> - Slides 27–28, & 32: Teacher works with the students to solve the exercises provided for each context clue type in these slides. - Slides 29 & 33: Students work with a partner to see if they can use the antonym and synonym context clues in the passages provided to figure out the meanings of unfamiliar words. Teacher needs to provide support when needed.
<p>You Do (Independent Practice)</p>	<ul style="list-style-type: none"> - A-4 Text: Our Government - Slides 34: Students can continue to practice independently using context clues to figure out the meanings of unfamiliar words in the passage provided (A-4).

APPENDIX C

EXAMPLES OF THE QUESTION ITEMS INCLUDED IN BOTH EXPERIMENTER-CONSTRUCTED INSTRUMENTS

Part 1. Multiple-Choice Recognition Test

For each of the following sentences, choose the correct definition of the word in *italics*.

(1) Margaret had *ambivalent* feelings about attending the concert. She knew she'd enjoy the music, but she didn't really want to go out in the rainy weather.

- a. mixed b. angry c. distrustful d. pleasant

(2) Regular exercise is *beneficial* to your body, but too much exercise is not good for you.

- a. helpful b. harmful c. pleasant d. hurtful

(3) My little brother is *gaunt*—just so thin and bony!

- a. fat b. joyful c. disobey d. slim

(4) Brother was quiet and well-behaved, but Sister was as *obstreperous* as a bucking mule.

- a. fast b. difficult c. pretty d. calm

(5) The dark clouds looked *ominous*; therefore, Mr. Tejada decided it was best that we cancel the trip.

- a. nice b. threatening c. spread d. become smaller or fewer
in size

Part 2. Meaning Recall Test

What do these words in italics mean?

(1) At that time, Uncle Roger was experimenting with a *peripatetic* existence, which is to say he moved constantly, never staying long enough to have an official address.

peripatetic:

.....

(2) Though some students are *aloof*, others pay attention to everything.

aloof:

.....

(3) Lou was sent to the *haberdashery* to find a new suit. He needed to wear one for his uncles' wedding.

haberdashery:

.....

(4) If the meeting begins at 10 AM, as planned, we should *adjourn* by 4:30.

adjourn:

.....

(5) That morning, the weather seemed *propitious* for our coming field trip: the sun was out, temperatures were in the 70s, and the humidity was low.

propitious:

.....

APPENDIX D

TEACHER FIDELITY CHECKLIST

Teacher Fidelity Checklist

Teacher: _____ Date: _____
 Observer: _____ Lesson Day#: _____
 School: _____

Directions: As you observe the intervention lesson, please write an “X” in the “Yes” box located next to each step if the behavior was observed. Place an “X” in the “No” column if the behavior was not observed. Once the lesson is complete, add the total number of “Xs” in the “No” and “Yes” columns separately. Please add additional comments in the box provided below the observable behaviors section.

Note: If the step/behavior is not applicable, write N/A in the “Yes” column and do not include it in the calculation of fidelity.

Observable Teacher Behaviors	Yes	No
TEACHER-LED PRACTICE (I Do)		
Teacher has the appropriate instructional materials/technology prepared for the lesson (e.g., computer, classroom projector, PowerPoint presentation slides, markers, copies of reference sheets, copies of context clues graphic organizer, worksheets, etc.)		
Teacher begins the class with a review of the previous lesson.		
Teacher provides each student with a copy of the materials needed (e.g., text, graphic organizer, worksheets, reference sheets [R-3/4], etc.)		
Teacher checks if students have copies of Project R-3 and R-4 sheets and reminds students to use them to figure out the meaning of unfamiliar words. <i>Note: Teacher has students turn to these reference sheets throughout the PowerPoint presentation. Project R-3 will be used as a reference for each type of context clues (focusing on the signal words). Project R-4 will be used as a reference for the steps of how to use each context clue.</i>		

Teacher uses practice sentences or paragraphs in the PowerPoint slides as an activating strategy.		
Teacher explains and talks through each PowerPoint slide in detail and enriches the presentation with his/her own teaching of the materials.		
Teacher reads the practice sentences or paragraphs aloud and then demonstrates examples of the thinking process that good and skilled readers use to understand how using context clues can improve comprehension.		
Using the “think aloud” cognitive strategy, teacher models the exercises in the PowerPoint slides about the type of context clues that are covered in the lesson by highlighting signal words and explaining how the meanings of unfamiliar words were attained.		
Teacher explains and models how to use the context clues graphic organizer sheet using the practice sentences provided in the PowerPoint slides.		
Teacher uses the context clues graphic organizer sheet OR draws it next to at least one of the practice sentences or paragraphs on the whiteboard or smart board and uses it.		
Teacher uses time of the lesson effectively and efficiently. <i>Note: This includes teacher following all the steps or activities suggested in the lesson plan.</i>		
GUIDED PRACTICE (We Do)		
Teacher has students turn to the guided practice activities and exercises for the lesson.		
Teacher uses several practice sentences and/or short paragraphs provided in the PowerPoint slides as additional practice with the class.		
Teacher collaboratively works with the students to solve the practice sentences and/or paragraphs provided for each context clue type in the PowerPoint slides.		

Teacher circulates throughout the classroom and has students work in small groups or in pairs (monitoring progress) on the activities/worksheets of the lesson.		
If students are able to find the meanings of unfamiliar words in the practice sentences and/or paragraphs provided, the teacher asks them how they figured out the meanings.		
If unable to determine the meaning of a word(s), teacher shows students (or has classmates share) how they used clues from the context to find the meanings of the unfamiliar words.		
INDEPENDENT PRACTICE (You Do)		
Teacher has students turn to the independent practice worksheet or text and instructs them to complete it but to ask for help if needed.		
Teacher circulates throughout the classroom and monitors students' work to answer questions or provide assistance when needed.		
Teacher reviews students' answer sheets once they have finished and then corrects students' errors as needed.		
Total number of behaviors observed (Y or "Yes" column)		
Total number of behaviors not observed (N or "No" column)		

Teacher Fidelity Total:

 /20 points Total Y/Total Y + N x 100 = %