A TWO-STUDY INVESTIGATION OF RESEARCH ON VOCABULARY STRATEGIES AND THEIR IMPLEMENTATION IN FOURTH GRADE SOCIAL STUDIES CLASSROOMS

A Dissertation

by

ANGELA RENEE SEBESTA HAIRRELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2008

Major Subject: Curriculum and Instruction

A TWO-STUDY INVESTIGATION OF RESEARCH ON VOCABULARY STRATEGIES AND THEIR IMPLEMENTATION IN FOURTH GRADE SOCIAL STUDIES CLASSROOMS

A Dissertation

by

ANGELA RENEE SEBESTA HAIRRELL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Co-Chairs of Committee;	William Rupley
	Deborah Simmons
Committee Members;	Lynn Burlbaw
	Victor Willson
Head of Department;	Dennie Smith

August 2008

Major Subject: Curriculum and Instruction

ABSTRACT

A Two-Study Investigation of Research on Vocabulary Strategies and Their Implementation in Fourth Grade Social Studies Classrooms. (August 2008)

Angela Renee Sebesta Hairrell, B.S., Texas A&M University;

M.Ed., Texas A&M University

Co-Chairs of Advisory Committee: Dr. William Rupley
Dr. Deborah Simmons

Among the multiple dimensions of reading, vocabulary knowledge and strategies are essential to skilled reading. As a result, this two-part dissertation (a) systematically examines the vocabulary intervention research, in both content and methodology, published since 1999, and (b) documents the implementation of evidence-based vocabulary strategies in fourth grade social studies classrooms.

Twenty-four studies were included in the systematic literature review. Results of this study corroborate findings of past studies that several vocabulary strategies have emerged that are effective for increasing students' vocabulary knowledge. Findings further reinforce the National Reading Panel's recommendations regarding the context and magnitude of studies needed. Additionally, results of the analysis of the methodological characteristics of the 24 studies revealed mixed alignment of research methods with standards recommended by educational and research organizations.

A study of 26 fourth grade social studies teachers' use of vocabulary strategies was conducted based on an existing data set acquired as part of a larger professional development study. In that study, teachers were randomly assigned to either a typical practice or professional development group. Analysis of teachers' instructional practice revealed that few of the vocabulary strategies identified in the literature are used in typical fourth grade social studies classrooms. Teachers who received professional development used a wider array of strategies. Controlling for teachers' preknowledge of vocabulary strategy instruction, results of a MANCOVA showed that the professional development group was statistically different from the typical practice group in terms of overall instructional quality, time allotted for vocabulary instruction, and variety of strategies.

Additional analyses were conducted comparing the findings of Durkin's study of comprehension in fourth grade social studies classrooms to the current practices of nine fourth grade social studies teachers. Findings showed little change in teachers' reading comprehension instruction even though the knowledge base of effective instruction has increased in the past 30 years.

DEDICATION

To my husband, Brad and my children, Cullen and Deanna

ACKNOWLEDGEMENTS

I would like to thank my committee chairs, Dr. Deborah Simmons and Dr. William Rupley and my committee members, Dr. Victor Willson and Dr. Lynn Burlbaw.

To Dr. Simmons, thank you for your mentorship and guidance as I have traveled the path from classroom teacher to researcher.

To Dr. Rupley, thank you for your encouragement and input every step of the way.

To Dr. Willson, thank you for helping me understand the new vocabulary of statistics.

To Dr. B, thank you for seeing potential so many years ago and never giving up on me.

A sincere thanks to Lori for "blazing the trail" and always being available.

A special thank you to family and friends who cared for Cullen and Deanna and ensured that they did not miss any important events, or who sent emails of encouragement, or who generally kept me sane. Although I would like to mention each of you and each act of kindness, were I to do so, the acknowledgements would be longer than the actual dissertation!

Most importantly, to Brad, Cullen and Deanna, thank you from the bottom of my heart. Brad, thanks for your encouragement and support over the past three years. Cullen and Deanna, thanks for your love, patience, understanding, and encouragement.

Without each of you this would not have been possible.

NOMENCLATURE

AERA	American Education Research Association
DWM	Degrees of Word Meanings
GMRT	Gates-MacGinitie Reading Test
IES	Institute of Education Sciences
NAEP	National Assessment of Educational Progress
NCLB	No Child Left Behind
NRP	National Reading Panel
WWC	What Works Clearinghouse

TABLE OF CONTENTS

CHAPTER Page

LIST OF FIGURES

LIST OF TABLES

TABLE Page

CHAPTER I

INTRODUCTION

Reading is a complex set of skills that children begin to develop prior to school, refine during their school years, and apply throughout their lives for pleasure, education, and employment. Teaching children to read is one of the most recognized responsibilities of schools and a basic expectation of both parents and children. Recent results of the National Assessment of Educational Progress (NAEP) (Lee, Grigg, & Donahue, 2007) indicated a statistically significant gain in fourth grade students' reading comprehension performance. Among the multiple dimensions of reading, vocabulary knowledge and strategies are essential to skilled reading (Baxter & Reddy, 2007; Biancarosa & Snow, 2006; National Reading Panel, 2000; Snow, 2002). The report of the National Reading Panel (NRP) (2000) highlighted the distinct role of vocabulary knowledge in comprehension by noting that "reading comprehension is a cognitive process…and cannot be understood without examining the critical role of vocabulary learning and instruction in its development" (p. 5-1). However, the Panel unequivocally concluded that the extant vocabulary research knowledge base is insufficient. Concurrent with the interest in vocabulary research is the additional need for methodologically rigorous studies (Gersten, Fuchs, Compton, Coyne, Greenwood, &Innocenti, 2005; Institute of Education Sciences, 2006). While research has indicated a converging set of strategies for increasing vocabulary knowledge, little is known about teachers' actual use

---

This dissertation follows the style of *Reading Research Quarterly.*

of these vocabulary strategies. The difficulty of translating knowledge *of* to knowledge of *how,* or translating research into practice, (Carnine1997; Greenwood & Abbott, 2001; Kauffman, 1996; Kennedy, 1997; Moats, 1999; Robinson, 1998) is a necessary next step.

In a manuscript-style format, I present two individual studies: (a) a systematic literature review of vocabulary interventions reported in the literature since 1999, and (b) an investigation of the use of evidence-based strategies in fourth grade social studies classrooms. This research was designed to examine the quantity and methodological characteristics of vocabulary instructional treatment research and to document to what extent such evidence-based instruction is being implemented. Following the style guidelines in the *Publication Manual of the American Psychological Association, Fifth Edition* (American Psychological Association, 2002*),* I organized this dissertation into five chapters (with Chapters II and III intended to stand alone as manuscripts to be submitted for scholarly publication and Chapter IV providing additional analyses) and two appendices. Chapter I provides an overall introduction to the content that follows.

Chapter II presents the rationale, methods, and results of a systematic literature review of vocabulary intervention research. The purpose of this review is two-fold. First, it updates vocabulary syntheses and catalogues publications since the NRP review of vocabulary research in 2000. Second, this review examines the alignment of research methods employed in experimental and quasi-experimental studies with those proposed by educational organizations and agencies (American Education Research Association Task Force on Reporting of Research Methods in AERA Publications, 2006; Gersten et

al., 2005; Society for Prevention Research, 2004; U.S. Department of Education, Institute of Education Sciences, & National Center for Evaluation and Regional Assistance, 2003; What Works Clearinghouse, 2006).

A major focus of the NRP in 2000 was to identify the state of knowledge about reading instruction, including vocabulary, by conducting a meta-analyses of the extant experimental and quasi-experimental research. At the time, this was the most comprehensive and rigorous review of vocabulary ever conducted evaluating studies from 1978 to 1998. Although the Panel was unable to conduct a meta-analysis, a list of practices that reflected the trends noted in the research literature was developed, including the teaching of vocabulary both directly and indirectly, the importance of multiple exposures, direct instruction, active engagement, multiple strategies, enhanced acquisition through technology and rich contexts (NRP, 2000). The Panel concluded that while there is clearly much more that needs to be known about vocabulary instruction under actual classroom conditions, there is a substantial foundation of empirical knowledge on which to base instructional practice.

The research questions answered in the first article are: (a) What additional evidence from experimental and quasi-experimental vocabulary instructional research, second through eighth grades, is available since the publication of the NRP report? and (b) What are the methodological characteristics of published research studies involving vocabulary instruction in second through eighth grades published since the NRP report?

To complement the investigation of research conducted in vocabulary, the second study focused on teachers' use of strategies in classrooms. Chapter III presents the

results of a study of fourth grade social studies teachers' use of vocabulary instructional strategies. Though effective reading and vocabulary instructional strategies have been documented (Baker, Simmons, & Kame'enui, 1998; Baumann, Kame'enui, & Ash, 2003b; Blachowicz & Fisher, 2000; Jitendra, Edwards, & Sacks, 2004; Kuhn & Stahl, 1998; NRP, 2000; Read, 2004; Stahl & Fairbanks1986), observational studies supporting their use are limited.

An area that is sorely lacking in the research is how to integrate effective vocabulary strategies into content area reading instruction (Biancarosa & Snow, 2006; Snow, 2002). Recent publications have further emphasized the need for investigating content area vocabulary instruction and its resulting impact on reading comprehension and content area learning (Baxter & Reddy, 2007; Scammacca, Roberts, Vaughn, Edmonds, Wexler, Reutebuch, & Torgesen, 2007). Therefore, even with the growing evidence of best practices in some dimensions of reading, little attention has been devoted to developing teacher knowledge of the skills and strategies that promote vocabulary development and enhanced comprehension of informational texts (Snow, 2002).

The purpose of this study is to evaluate whether and to what degree evidence-based vocabulary strategies are used in fourth grade social studies instruction. Furthermore, this study attempts to determine whether professional development in evidence-based vocabulary strategies increases the quantity and quality of teachers' use of these strategies. Research questions addressed are: (a) What evidence-based vocabulary strategies do fourth grade teachers use during social studies instruction? To

what extent are these strategies used during social studies instruction? (b) Does professional development in evidence-based vocabulary strategies result in differential use in the quantity and/or quality of strategies by teachers during social studies instruction?

Chapter IV provides a supplemental descriptive analysis of the data reported in Chapter III. An initial purpose of this dissertation was to document changes in social studies vocabulary practice in fourth grade social studies classrooms. A question answered is: How do the findings in the present study compare to Durkin's findings 30 years ago regarding the amount of time devoted to vocabulary instruction in fourth grade social studies classrooms?

Chapter V summarizes the research questions and findings of the two studies. Chapter V is followed by appendices that provide further detail on the instruments used in this dissertation. Appendix A includes the abstraction form used in the systematic literature review, while Appendix B contains the modified strategy use instrument used in the study.

CHAPTER II

THE STATE OF VOCABULARY RESEARCH: A SYSTEMATIC LITERATURE REVIEW FROM 1999 -2007

Vocabulary knowledge is one of many factors strongly associated with reading competence, particularly in the upper elementary grades and beyond where the majority of text read is informational (Anderson & Freebody, 1981; Carney, Anderson, Blackburn, & Blessing, 1984; Hart & Risley, 2003; Hirsch, 2003; Kame'enui, Carnine, & Freschi, 1982; Stahl & Fairbanks, 1986). Understanding the meanings of words and their relation to text comprehension and reading achievement has been the focus of considerable correlational and causal research. For example, Cunningham and Stanovich (1997) found correlations ranging from .55 through .85 between vocabulary knowledge and reading comprehension. McKeown, Beck, Omanson, and Perfetti (1983) documented the benefits of vocabulary instruction and practice in multiple contexts, and its relationship to improved reading comprehension. More recently, Cromley and Azevedo's (2007) direct and inferential mediation model found that vocabulary was one of the largest contributors to ninth grade students' reading comprehension. Despite these converging findings, recent research syntheses highlighted that the body of research supporting vocabulary intervention was insufficient in quantity and quality (Baumann et al., 2003b; Jitendra et al., 2004; National Reading Panel, 2000).

The National Reading Panel (NRP) report in 2000 played a prominent role in profiling vocabulary as a core component of reading instruction. Concurrent with the

recognized importance of vocabulary to reading comprehension, this report noted the inadequacy of the vocabulary research base. Multiple reviews have summarized research findings on vocabulary acquisition and instruction (Baker et al., 1998; Baumann et al., 2003b; Harmon, Hedrick, & Wood, 2005; Jitendra et al., 2004; Kuhn & Stahl, 1998; Read, 2004; Swanborn & de Glopper, 1999) with a variety of foci and methodologies. Although the research base to support specific vocabulary instruction may be insufficient, the importance of vocabulary to reading comprehension and academic achievement has been well-documented, and the need for reliable knowledge of how vocabulary is effectively taught and acquired is critical.

Inherent in reliable findings are rigorous standards of high-quality research. Numerous researchers and national entities have published guidelines for conducting quality research studies (American Education Research Association Task Force on Reporting of Research Methods in AERA Publications, 2006; Gersten et al., 2005; Society for Prevention Research, 2004; U.S. Department of Education et al., 2003; What Works Clearinghouse, 2006). These standards for research are lofty goals; however, their intent is aimed at establishing and maintaining research rigor.

The call for more vocabulary research that likewise addresses the standards for higher-quality research was initiated with the NRP report and reinforced by professional organizations, such as the Institute of Education Sciences (IES). The purpose of this present study is to examine research published since the NRP review and to analyze it according to intervention content and research methods. First, summarizing research reviews from 1998 - 2004 provides context for the analyses. Second, recommendations

for rigorous research are synthesized. Third, using the synthesized research standards and summarized vocabulary research findings, vocabulary studies from 1999 to 2007 will be described.

What We Know About Vocabulary Instruction

The most widely disseminated and comprehensive review of vocabulary instruction was the NRP Report (2000). The Panel identified 50 studies published from 1979-1998 that met the inclusion criteria. Unfortunately, due to this small number of studies, the Panel was unable to conduct a meta-analysis. Criteria included experimental or quasi-experimental studies, reports of research, studies focused on the English language and implemented with English speaking groups, and excluding studies dealing exclusively with learning disabled or other special populations. Forty-two of the studies (84%) were conducted in third-sixth grades and six (12%) were conducted with informational texts. As a result of their findings, nine implications for vocabulary practice were recommended.

Even with this foundation of vocabulary instructional knowledge, gaps in evidence-based vocabulary instruction exist, which led the Panel to conclude that research in school settings is still needed. While numerous vocabulary instructional strategies have been studied, few studies have focused on the efficacy of any one particular strategy in a variety of contexts. Thus, the robustness of any one given vocabulary strategy has to be called into question. Additionally, vocabulary knowledge at different age and ability levels is inconclusive. Two gaps identified in NRP cause serious threats to external validity: (a) the lack of studies conducted under typical

classroom conditions with large number of students, and (b) a dearth of studies using vocabulary instructional strategies with expository text. Identified directions for future research, as noted by the Panel, included evaluation of vocabulary knowledge, documentation of instructional effects at various grade and achievement levels, use of technology and multi-media approaches for instruction, integration of vocabulary and comprehension instruction, determination of the most effective combinations of strategies, and evaluation of the efficacy of various types of professional development.

Recent Reviews of Vocabulary Research

In addition to the NRP's report, six reviews and one meta-analysis of vocabulary instruction were published between 1998 and 2004 (Baker et al., 1998; Baumann et al., 2003b; Harmon et al., 2005; Jitendra et al., 2004; Kuhn & Stahl, 1998; Read, 2004; Swanborn & de Glopper, 1999). The single meta-analysis conducted by Swanborn and de Glopper (1999) examined incidental word learning. Kuhn and Stahl (1998) synthesized the research of learning words from context, whereas Baker and colleagues (1998) identified advances in the research on vocabulary development for diverse learners. In *Research on Vocabulary: Voltaire Redux*, Baumann, Kame'enui, and Ash (2003) categorized vocabulary strategies by their use: strategies for teaching specific words, and strategies to learn words independently. Other vocabulary reviews have focused on either more restricted populations or topics. For example, Read (2004) examined studies in second language learners vocabulary instruction since 1999, while Harmon, Hedrick, and Wood (2005) identified several effective strategies for students struggling with content area texts. The importance of choosing an instructional method

based on the instructional goals and needs of the individual student was highlighted by Jitendra, Edwards, and Sacks (2004).

Of the seven reviews, five were nonquantitative literature reviews and two employed quantitative methods (Jitendra et al., 2004; Swanborn & de Glopper, 1999). Three of the studies were published before 2000. The stated purposes of the reviews ranged from gathering evidence regarding the efficacy of a specific strategy with a particular population to a historical review of intervention studies. Three syntheses focused on struggling readers or those with learning disabilities and one focused on second language learners. The time span of reviewed studies ranged from three to 30 years. Additionally, two authors identified the specific criteria for selection of articles. No reviews discussed the research methodologies used in the studies cited. Table II.1 presents specific details of each review.

Support for Vocabulary Instruction

One goal of the NRP (2000) was the creation of a taxonomy of vocabulary instruction to benchmark classroom practice. Unfortunately, this was not possible due to both the uniqueness of the studies available and the multi-dimensional nature of vocabulary instruction. Recently, the nature of skills and strategies has been discussed (Afflerbach, Pearson, & Paris, 2008), with skills being those abilities that become automatic and strategies being those that are consciously employed when needed. For the purpose of this study, strategies are identified as vocabulary instructional practices that students can consciously use to learn the meanings of new and unfamiliar vocabulary words. The ultimate goal is that as students practice these strategies, they

Table II.1.
Summary of Post-NRP Vocabulary Literature Reviews and Their Findings

Review	Year	Type	Purpose/Range of search	No. of studies	Population	Selection	Findings
Baker, Simmons, & Kame'enui	1998	Synthesis	Vocabulary development as it relates to diverse learners 1977-1991	23	Diverse learners	Not explicit	Semantic mapping/ features analysis Keyword method Computer-assisted instruction
Baumann, Kame'enui, & Ash	2003	Review	Vocabulary research and intervention studies since the 1970s 1963-1996	138 (not reported)	All learners	Not explicit	Word specific strategies Mnemonics Preteaching Provide partial knowledge Listening Reading independently Writing Morphemic and contextual analysis Student responsibility Use of resources
Harmon, Hedrick, & Wood	2005	Review	Effective vocabulary strategies for struggling students in the content area 1981-2001	46 (not reported)	Students reading below grade level	Not explicit	Independent reading Use of trade books Contextual-base Self-selection of words Explicit instruction Multiple exposures Avoidance of drill and Structural analysis Staff development

Table II.1 cont.

Review	Year	Type	Purpose/Range of search	No. of studies	Population	Selection	Findings
Jitendra, Edwards, & Sacks	2004	Review-used some statistics (M, SD, ES)	Summarize vocabulary interventions 1978-1996	19/27 interventions	Students with learning disabilities	Published experimental, quasi-experimental, or single-subject studies Subjects identified LD Elementary-High School students Addressed vocabulary meaning Measured vocabulary outcomes	Goals and needs of the learner determine most effective instructional method Direct instruction Practice Computer-aided instruction
Kuhn & Stahl	1998	Synthesis	Using context to learn word meanings 1976-1996	14	All learners	Examined contextual analysis	Students can learn to use context to derive word meanings No evidence that instruction in contextual analysis increase incidental word learning.

Table II.1 cont.

Review	Year	Type	Purpose/Range of search	No. of studies	Population	Selection	Findings
Read	2004	Review	Vocabulary instruction since 1999 1999-2002	29 (not reported)	Second language learners	Not explicit	Keyword method Semantic features analysis Direct instruction Constant time delay Activity-based
Swanborn& de Glopper	1999	Meta-analysis	Meta-analysis of incidental word learning studies 1985-1995	20	Excluded second language learners	Assessed from reading in the mother tongue No attention drawn to vocabulary Context not intentionally transparent Words encountered in only one text Provided sufficient statistical information	Found a mean effect size of logit (p) = -1.70 The higher the grade the more words learned incidentally The higher the reading ability the more words learned incidentally Assessments sensitive to partial word learning showed higher gains Students learn more words when the ratio of text to words is higher

develop into skilled readers moving from thoughtful application to automatic use. The following section provides definitions and the evidence base of vocabulary strategies drawn from the NRP report and more recent reviews and, which are, subsequently, included in this review (see also Table II.2).

Evidence Base of Specific Vocabulary Strategies. This section reviews identified salient components of effective vocabulary instruction, such as contextual and morphological analysis, semantic analyses, mnemonics, explicit instruction, incidental word learning, repeated exposures, computer-assisted instruction, and combinations of the above strategies.

Contextual analysis has the largest empirical base of all vocabulary strategies. Contextual analysis is the use of clues within the context of the text to derive word meanings. This strategy may be either explicitly taught or occur incidentally (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003; Baumann & Kame'enui, 2003; Edwards, Font, Baumann, & Boland, 2003; Fukkink & de Glopper, 1998; Harmon et al., 2005; Kuhn & Stahl, 1998; Stahl & Fairbanks, 1986; Swanborn & de Glopper, 1999, 2002).

Morphology is the study of word parts. By breaking words into meaningful parts (root and affixes), the meaning can often be inferred (Biemiller, 2003b; White, Sowell, & Yanagihara, 1989). Once taught, morphological skills can help students learn words independently. Morphology can be used in conjunction with contextual analysis to provide further clues to unlocking the meaning of words (Baumann, Edwards et al., 2003; Baumann, Font, Edwards, & Boland, 2005).

Table II.2.
Evidence Base for Individual Vocabulary Instructional Strategies

	Definition/Examples	**References**
Mnemonics	Strategies that teach students the meaning of vocabulary words primarily through the use of a keyword or memory strategy.	NRP, Baker et al., 1998; Baumann et al., 2003a; National Reading Panel, 2000; Read, 2004
Explicit vocabulary instruction	Teacher-provided definitions, but also extend to teacher- and/or student-guided activities that combine multiple strategies that provide a rich understanding of the word	Baumann et al., 2003b; Fukkink & de Glopper, 1998; Harmon et al., 2005; Jitendra et al., 2004; NRP, 2000; Read, 2004
Incidental word learning	Story book reading, read-alouds, and independent reading where words are learned through everyday exposure	Harmon et al., 2005; NRP, 2000; Swanborn & de Glopper, 1999
Repeated practice or multiple exposures	A concerted effort to provide students with multiple opportunities to encounter and use targeted vocabulary words.	Baumann et al., 2003a; Baumann & Kame'enui, 2003; Beck et al., 2002; Bryant et al., 2003; Harmon et al., 2005; Jitendra et al., 2004; NRP, 2000; Stahl & Fairbanks, 1986; Swanborn & de Glopper, 1999
Computer assisted instruction	Use of computers or other forms of multimedia as part of the vocabulary intervention	Baker et al., 1998; Jitendra et al., 2004; NRP, 2000; Read, 2004
Multiple strategies	Many vocabulary interventions combine several strategies; therefore it is difficult to decipher the unique effect of each component.	Bryant et al., 2003; NRP, 2000; Snow, 2002

Table II.2 cont.

	Definition/Examples	References
Contextual analysis	Using context clues to derive word meanings. These strategies may be explicitly taught or occur incidentally.	Baumann et al., 2003a; Baumann & Kame'enui, 2003; Edwards, Font, Baumann, & Boland, 2003; Fukkink & de Glopper, 1998; Harmon et al., 2005; Kuhn & Stahl, 1998; Stahl & Fairbanks, 1986; Swanborn & de Glopper, 1999, 2002
Morphological analysis	Strategies to learn new words through study of word parts.	Baumann et al., 2003a; Baumann, Font, Edwards, & Boland, 2005; Biemiller, 2003b; White, Sowell, & Yanagihara, 1989
Semantic analyses	Strategies that attempts to help students categorize new vocabulary in meaningful ways, often through graphic representations	Baker et al., 1998; Bryant, Goodwin, Bryant, & Higgins, 2003; Carnine, Crawford, Harniss, Hollenbeck, & Miller, 2002; Jitendra et al., 2004; Read, 2004

Semantic analyses for teaching students new vocabulary are designed to enable students to categorize new vocabulary in meaningful ways, often using graphic organizers. Semantic mapping, where students map the target word with other words that categorically share the same meaning, is one example of this strategy used in research studies (Baker et al., 1998; Bryant, Goodwin, Bryant, & Higgins, 2003; Carnine, Crawford, Harniss, Hollenbeck, & Miller, 2002; Jitendra et al., 2004; Read, 2004). Other semantic analyses strategies are charts, such as semantic features analysis and graphic illustrations, to emphasize the meaningful relationships between words (Bryant et al., 2003; Heimlich & Pittelman, 1986; Jitendra et al., 2004).

Mnemonics include strategies that teach students the meaning of vocabulary words primarily through the use of a memory strategy, such as a keyword (Baker et al., 1998; Baumann, Edwards et al., 2003; NRP, 2000; Read, 2004). Many studies report

positive learning outcomes when students use memory association strategies to learn new words.

Explicit vocabulary instruction has repeatedly been shown to be an important principle of vocabulary instruction (Baumann, Kame'enui et al., 2003; Fukkink & de Glopper, 1998; Harmon et al., 2005; Jitendra et al., 2004; NRP, 2000; Read, 2004). Explicit instruction can include teacher-provided definitions and extend to teacher-directed activities that combine multiple strategies in scaffolded situations that are aimed at providing a rich and deep understanding of the word's meaning.

Incidental word learning included studies where no specific strategy to learn words was taught, but vocabulary growth was measured. Storybook reading, read-alouds, and independent reading are examples of strategies that are classified as incidental word learning (Harmon et al., 2005; NRP, 2000; Swanborn & de Glopper, 1999).

Repeated practice or multiple exposures occur when there is a concerted effort to provide students with multiple opportunities to encounter and use identified vocabulary words. Repeated exposures could be planned through games, repeated readings, and discussions (Baumann, Edwards et al., 2003; Baumann & Kame'enui, 2003; Beck, McKeown, & Kucan, 2002; Bryant et al., 2003; Harmon et al., 2005; Jitendra et al., 2004; NRP, 2000; Stahl & Fairbanks, 1986; Swanborn & de Glopper, 1999).

Computer-assisted instruction utilizes computers or other forms of multimedia as part of vocabulary instruction (Baker et al., 1998; Jitendra et al., 2004; NRP, 2000;

Read, 2004). Examples include computer programs to introduce and practice targeted vocabulary words or the use of video to introduce and learn new words.

Multiple strategies occur when several vocabulary strategies are used in combination. Several studies examined interventions using multiple strategies; however, it is difficult to determine the unique effect of each component (Bryant et al., 2003; NRP, 2000; Snow, 2002). Common strategy combinations found in the literature are contextual and morphological analysis, the explicit instruction strategies, or the development of packages of vocabulary instruction that includes multiple components.

As Afflerbach et al. (2008) highlighted, once vocabulary strategies are taught, students can then apply them independently when comprehension of the text breaks down. As a result of the more recent vocabulary research, there is evidence that vocabulary knowledge can be taught, and there are strategies that can be effectively used to enhance vocabulary acquisition. Based on this knowledge, this study examines vocabulary studies from 1999-2007 to identify the most recent scholarship regarding vocabulary instruction.

Converging Criteria to Evaluate the Rigor of Experimental and Quasi-experimental Research

The current research climate calls for studies that not only address issues relevant to the needs of the field, but that also meet the stringent guidelines set forth by researchers. Professional organizations and education agencies alike have put forth criteria to enhance and maintain the likelihood that current and future research reflect these standards.

For this review, I relied on two prominent sources, Gersten and colleagues (2005) and What Works Clearinghouse (WWC) (2006), for the foundational information of synthesized methodological characteristics. Gersten et al. (2005) developed standards for special education research based on recent publications in the research methodology field, Cooper and Hedges (1994), and other research syntheses conducted in special education. The purpose of these criteria was to evaluate whether a strategy or practice met the standards and could be classified as evidence-based. Using stringent research standards and a lengthy review process, WWC (2006) provided guidelines to determine the evidentiary strength of programs. According to these guidelines, programs reviewed are assigned one of three labels: meet evidence standards, meet evidence standards with reservations, or do not meet evidence standards.

Providing additional support for the synthesis of methodological characteristics were the American Educational Research Association (AERA) (2006), the Society for Prevention Research (SPR) (2004), and the U.S. Department of Education (USDE) (2003). AERA provided publication guidelines for empirical social science research as a framework of expectations in eight general areas, concluding that reports should clearly convey the logic of the research project from development to outcomes, and results and conclusions are justified by the evidence presented the evidence presented justifies results and conclusions. In *Standards of Evidence: Criteria for Efficacy, Effectiveness, and Dissemination* (Society for Prevention Research, 2004), a set of standards to identify effective programs for replication and dissemination with the goal of increasing confidence in and use of effective practices was presented. Published by the U.S.

Department of Education, *Identifying and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide* (2003) was designed to allow practitioners to identify and consequently implement evidence-based programs. High importance was placed on randomized, controlled trials in determining effectiveness. While not only providing guidelines to classify interventions as backed by "strong" or "possible" evidence of effectiveness, they also provided clear definitions of each characteristic.

Research that is rigorous and relevant is a professional standard for researchers today. Common themes from publications of education and research agencies are the importance of randomization and clear reporting of the methods, such as comparability of groups and attrition. Related to this, the necessity of effect size reporting was expressed in the majority of the guidelines. Complete descriptions of the instructional methods and the sample are important to determine the effectiveness of the program and populations who may reap the most benefits. While all studies may not meet all standards, these guidelines can provide a starting point for evaluating present studies and increasing the rigor of vocabulary research in future studies.

Statement of the Problem and Research Questions

The purpose of this systematic literature review was to (a) catalogue publications since NRP (2000) in the areas of vocabulary instructional research, and (b) examine the alignment of research methods employed in experimental and quasi-experimental studies with those proposed by educational organizations and agencies.

The research questions addressed were:

1) What additional evidence from experimental and quasi-experimental vocabulary instructional research, second through eighth grades, is available since the publication of the National Reading Panel report (1999-2007)?

2) What are the methodological characteristics of published research studies involving vocabulary instruction in second through eighth grades published since the NRP report (1999-2007)?

Method

This systematic literature review catalogued and described empirical studies published from 1999-2007 that examined vocabulary instructional strategies and their impact on student vocabulary knowledge and acquisition. Studies prior to 1998 were included in the NRP vocabulary synthesis; therefore, it was determined that 1999 was an appropriate starting point for this literature review. Procedures used in this study are described below with the following subheadings: inclusion/exclusion criteria, data sources and search strategy, outcome of searches, data extraction and synthesis, and data analysis.

Inclusion/Exclusion Criteria

Criteria for inclusion in this systematic literature review were drawn primarily from those used by NRP (2000).

Inclusions criteria:

(1) published in a peer-reviewed journal (English only),

(2) defined vocabulary as word knowledge, not word recognition

(3) examined empirically vocabulary interventions,

(4) conducted in grades 2-8 or ages 8-14,

(5) published between January 1999 and November 2007, and

(6) conducted with native English speakers.

Exclusion criteria:

(1) included in NRP,

(2) focused only on second language learners or on languages other than English,

(3) used qualitative methods only,

(4) defined as theoretical or editorial pieces,

(5) unpublished doctoral dissertations or presentations, and

(6) published in books or handbooks of research.

If studies used a mixed methods design (a combination of quantitative and qualitative methods), they remained in the sample.

Data Sources and Search Strategy

Eight databases of published studies were searched -- Academic Search Premier, Education Full Text (Wilson), Linguistics + Language, PsycINFO 1872-current, Social Sciences Full Text, Education: A SAGE Full-Text Collection, ISI Web of Knowledge, and Google Scholar. Search terms included “vocabulary instruction”, “vocabulary learning”, and specific vocabulary strategies (e.g., context clues, explicit vocabulary instruction, incidental word learning). To ensure a comprehensive search, author searches were also conducted of those in previously identified studies. Reference lists of

reviewed studies and available literature reviews were also hand-searched for additional publications. This search identified 21 articles meeting the established criteria. Because three of the articles reported more than one study, 24 studies comprised the final sample.

Outcome of Searches

Initial searches resulted in 402 possible articles on vocabulary instruction that were obtained through electronic and hand searches. Using the above-mentioned inclusion and exclusion criteria, an initial screening was conducted based on title and abstract. Based on the initial screening, 36 articles were then admitted to the full-review process, which included a critical reading of the entire article and abstraction of data. Under this review, 15 articles were found to not meet one or more of the inclusion criteria, resulting in the final sample of 21 articles. Twenty percent *(n* = 80) of the total possible articles (*n* = 402) were double-screened with an inter-rater reliability of 1.0.

Data Extraction and Synthesis

Data abstraction categories were developed for two dimensions: (a) methodological characteristics, and (b) vocabulary instruction (both delivery features and strategies). Synthesizing the recommendations of WWC (2006) and Gersten and colleagues (2005), a list of methodological characteristics was developed. To support decisions regarding the inclusion of characteristics, The Coalition for Evidence-Based Policy (U.S. Department of Education, Institute of Education Sciences, & National Center for Evaluation and Regional Assistance, 2003) the *Standards of Evidence* by the Society for Prevention Research (2004), and AERA's Standards for Reporting on Empirical Social Science Research in AERA Publications (2006) were used. The

resulting synthesized methodological characteristics used to catalogue and describe the studies in the present literature review are listed in Table II.3.

Table II.3.
Synthesized Methodological Characteristics of Research Standards from Five Sources

Design Characteristics	1	2	3	4	5
What was the design of the study?	X	X		X	X
Did the study include a control group or a comparison group?	X		X	X	
Was the method for assigning participants to study conditions random assignment, stratified random assignment, or nonrandom?	X	X	X	X	X
Was the method of randomization reported?		X		X	X
Sample Characteristics					
Was there sufficient evidence on the relative characteristics of the sample to generalize to a population?			X	X	X
What was the total sample size? What were the group sizes? What was the number of classrooms (if appropriate)?		X		X	X
Was there sufficient evidence to determine that the sample is comparable across conditions through 1) demographics and/or 2) key predictor measures? Were groups matched? Or were differences statistically controlled for?	X	X	X	X	X
Was attrition reported? If severe, was the issue addressed?	X	X	X	X	
Intervention Characteristics					
Was a sufficient description of interventionists/teachers provided?	X	X	X	X	
Was the intervention clearly described and specified? Was the comparison condition described and specified?	X		X	X	X
Was fidelity and fidelity quality to the intervention reported?	X		X		
Statistical Characteristics	X	X	X	X	
Were multiple measures used to evaluate both generalizable results and proximal results?					

Table II.3 cont.

Statistical Characteristic cont.	1	2	3	4	5
Was the effect of the intervention measured beyond the immediate posttest?	X		X	X	
Did the measures used have evidence of reliability?	X	X	X		X
Were the measures used a valid measure of the target construct?	X		X	X	X
Was inter-rater reliability reported?	X	X	X		
Was the unit of analysis appropriate for the research question?	X	X	X		
What statistical methods were employed in the analysis?	X	X		X	X
Were effect sizes reported and how were they calculated ?	X	X		X	X
Was the reporting of the results done in a clear, coherent manner?	X				X

(1) Gersten et al. (2005).
(2) WWC (2006).
(3) Society for Prevention Research. (2004).
(4) Coalition for Evidence-based Policy. (2003).
(5) AERA (2006).

Methodological characteristics were categorized by study design, sample, intervention, and statistical analyses. Under study design, the primary question was what design methodologies were used in the study to ensure the validity of findings--specifically the use of random assignment of subjects and comparison and/or control groups. The sample description served dual purposes. First, for a study to have adequate power, the sample size must be large enough. Second, initial group differences must be identified and accounted for in the analyses. The description of the sample must provide sufficient information to compare the study sample to similar populations and for replication. Attrition must also be considered to determine if there was differential attrition between groups. Intervention characteristics included descriptions of the

treatment and the comparison or control treatment, including the length, setting, and provider of the treatment. Adequate descriptions of these characteristics allow for comparisons across studies. Statistical analyses characteristics related to the study design, data collection, statistical analysis, and reporting of results. Information abstracted included the types of measures commonly administered, common statistical analyses conducted, and the reporting of effect sizes. Because of its prominence in IES and WWC publications, the match between unit of randomization and unit of analyses was also recorded.

The initial coding was refined after a sample of the studies was coded, resulting in a better understanding of the available literature and the methodological characteristics commonly found. For example, it became apparent that simply coding assignment as random or nonrandom was too general, and it was necessary to code more descriptively to include stratified random samples. To ensure reliability across coders, descriptions for each methodological characteristic were drawn from information available from WWC (2006) and used to create a definitions document and codebook. Coders were trained in the use of the abstraction instrument and then completed the abstraction of two practice articles. Any disagreements in coding were discussed until agreement was reached.

Inter-rater reliability was determined by two raters using double coding of 30% (n = 8) of the final sample. Inter-rater reliability by category was as follows: study design, .97; sample characteristics, .92; intervention characteristics, 1.0; data collection methods, .97; data analysis, 1.0; findings, 1.0; and an overall inter-rater reliability of .97.

Full agreement on the abstracted articles was reached through discussion between the raters. Additionally, articles coded early in the investigation were coded again at the end to determine reliability over time (July 2007 - November 2007), resulting in no discrepancies.

All abstracted data were entered into a matrix (see Table II.4 for an abridged version) to facilitate the identification of patterns and themes across studies.

Data Analysis

The goals of the analyses were to document and describe the most recent vocabulary findings as well as the methodological characteristics of research studies available regarding vocabulary instruction since the NRP publication. A meta-analysis was beyond the scope of this current project. However, because of the importance of effect sizes in high quality research to document both statistical and practical impact of the treatments, effect sizes are included as reported in the original studies. Additionally, to facilitate comparisons among studies, standardized mean effect sizes were computed using the formulas in Lipsey and Wilson (2001). The standardized mean effect size was calculated from information provided in the research articles using the following formula.

$$ES_{sm} = \frac{\bar{X}_{G1} - \bar{X}_{G2}}{S_p}$$

In studies with an experimental and a control group, post-test means were used to determine the effect size. In pretest–post-test designs without control groups, the pretest mean and posttest mean were used. Standardized mean effect sizes were not calculated

Table II.4.
Abridged Research Matrix of Study Characteristics and Summary of Findings of 24 Vocabulary Studies in the Systematic Literature Review

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Apthorp, (2006).	Pre/Post Random	3	226	-EI -MS -ME -MA -SEM -RA -VIS	Classroom with teacher	20 min 5 days a week for 24 weeks	NR for sight vocabulary	Site A: 1.78** Site B: 1.57**	-Positive effects on vocabulary growth at one site; not replicated at the second site.
Baumann, Edwards, Boland, Olejnik, & Kame'enui, (2003).	Pre/Post Stratified Random	5	157	-EI -MS -CA -MA	Classroom with teacher	15 min per class for 25 days	For morph: η^2=.01 to .42 For context: η^2=.002 to .04	For morph: 1.18* For context: 0.15*	-Explicit lessons on vocabulary enhanced student learning and recall of vocabulary. -MA provided a tool for independent word learning. -Mixed results on the transferability of CA.

Table II.4 cont.

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Baumann, Edwards, Font, Tereshinski, Kame'enui, & Olejnik, (2002).	Pre/Post Random	5	88	-EI -MS -ME -CA -MA	Classroom with teacher	50 min per lesson for 12 lessons	For MA: *d*=.12 to 1.32 For CA: *d*= -.01 to .87 MA over CA: *d*= -.99 to 1.58	Morphemic Group:-0.11 for CA and 2.56 for MA Context Group: -0.12 for CA and 0.16 for MA For MA/CA group: 1.30 for CA and 1.66 for MA	-Effect for inferring word meaning is immediate for MA. -Effect of combining MA and CA are just as powerful as alone. -Effects degrade with time.
Boulware-Gooden, Carreker, Thornhill, & Joshi, (2007).	Pre/Post Non-random	3	119	-MS -MCS	Classroom with teacher	30 min 5 days a week for 5 weeks	T v. C *d*=.16	1.93*	-Use of MCS was more effective than traditional vocabulary instruction.
Cain, (2007).	Post test only Stratified random	7-8 yrs	45	-CA	With researcher	1 session	For Group: η^2_p=.14	Unable to calculate	-Explanations for inferred definitions lead to improved performance.

Table II.4 cont.

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Carlisle, Fleming, & Gudbrandsen, (2000). 2 studies	Pre/Post Non-random	4 8	42 45	-IWL	Classroom with teacher	4th: 2 class periods each week for 5 weeks 8th: 18 class periods	NR	Grade 4 topical: 1.04* Grade 8 topical: 1.2*	-Learning and retention of topical words significantly improved. -8th graders increased problem solving ability associated with growth in topical word knowledge.
Curtis, & Longo, (2001).	Pre/Post Non-random	6, 7, 8	NR	-EI -MS -ME	Classroom with teacher	45 min 5 times per week for 6 weeks	NR	Unable to calculate	-Students showed improvement on analysis of word relationships, responses to readings, use of words in speaking and writing, and recognition of word meanings.

Table II.4 cont.

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Fore III, Boon, &Lowrie, (2007).	Single subject/ multi-baseline Non-random	7	6	-MS -SEM	With Researcher	20 min twice per week for 3–5 weeks	NR	14.98*	-Concept model increased vocabulary learning
Lubliner, & Smetana, (2005).	Pre/Post Not reported	5	77	-MS -ME -CA -MA -SEM -RA MCS	Classroom with teachers and researcher	45 min twice per week for 12 weeks	For Title 1 students: *d*=.19 to 1.03 For Above Average: *d*= -.12 to .49	Above average v. low: 0.27*	-Implementation of a multi-component vocabulary intervention is possible in classrooms. -Brief lessons in vocabulary can improve student performance.
Nash, &Snowling, (2006).	Pre/Post Non-random	7-8 yrs	71	-MS -CA -MA -SEM -EI	With researcher	30 min twice per week for 6 weeks	For context: *d*=3.17	3.83*	-CA was more effective than definitions alone. -SEM made vocabulary more durable. -EI helped students use context.

Table II.4 cont.

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Stevens, (2003).	Pre/Post Matched	6, 7, 8	3916	-MS -RA -Writing -Pairs	Classroom with classroom teacher	4 months	ES=.33	0.066**	-Students' achievement improved. -Cooperative learning and writing led to more student engagement.
Twyman, McCleery, & Tindal, (2006).	Repeated measures Not reported	8	45	-MCS	Classroom with teacher	46 min per day for 5 weeks (21 sessions total)	*d*=1.94	0.81*	-Students learned vocabulary and were able to use new words in problem solving activities.
Uberti, Scruggs, &Mastropieri, (2003).	Pre/Post Not reported	3	74	-MS -ME -VIS -MNE -DEF	Classroom with teacher	1 session	NR	Picture v. Definition: -0.68* Keyword v. Definition: 0.89*	-Effective for 3rd grade students learning new words. -May be most helpful for students with learning disabilities.

Table II.4 cont.

Reference	Design	Grade	Sample Size	Strategy	Instruction Setting	Duration	Reported ES	Standardized ES	Findings
Veerkamp, Kamps, & Cooper, (2007).	Reversal Non-random	6	71	-DEF -ME	Classroom with classroom teacher	197 min per week for 19 weeks	NR	Unable to calculate	-Class-wide peer tutoring improved student performance.
Xin, &Rieth, (2001).	Pre/Post Random	4, 5, 6	76	-MS -VIS	Classroom with teacher	3-30 min 3 times per week for 6 weeks	NR	0.32*	-Video-based vocabulary instruction was more effective than traditional dictionary methods.

*Note: *researcher measure, **standardized measure, NR= Not Reported*
CA=contextual analysis, DEF=definitions, EI=explicit instruction, IF=instructional feedback, IWL=incidental word learning, MA=morphological analysis, MCS=metacognitive strategies, ME=multiple exposures, MNE=mnemonics, MS=multiple strategies, RA=read aloud, VIS=visual

Table II.5.
Summary of Study Findings Organized by Vocabulary Strategy

Strategy	No. of Studies	Age/Grade Levels	Effect Sizes	Practical Significance and Future Directions
Contextual Analysis	13	2, 3, 4, 5, 6	η^2=.002 to .035 *d*= -.01 to 3.17	Results for this strategy are positive in a variety of grade levels. Continued research needed on the extent of instruction needed for effects and what is appropriate for various ability levels.
Morphological Analysis	4	3, 5	*d*= -1.00 to 1.58	Results for this strategy show differential effects for grade levels studied.
Semantics	8	2, 3, 5, 6, 7	*d* = 0.55 to 3.17	This strategy was one component of successful interventions.
Mnemonics	2	3, 4, 6	*d*= 2.85	Small number of studies make it impossible to generalize.
Explicit Instruction	5	4, 5, 6, 7, 8	NR	Results of most recent scholarship support findings of NRP.
Incidental Word Learning	1	4, 8	NR	Some words can be learned incidentally in science classrooms.
Repeated Exposures	10	3, 5, 6, 7, 8	NR	Results of most recent scholarship support findings of NRP.
Multiple strategies	15	2, 3, 5, 6, 7, 8	*d*= -0.12 to 2.85	Successful intervention strategies employed multiple strategies.

Table II.5 cont.

Strategy	No. of Studies	Age/Grade Levels	Effect Sizes	Practical Significance and Future Directions
Computer Assisted/ Multi-media	1	4, 5, 6	NR	Only one study makes it difficult to generalize.
Instructional Feedback	1	4	NR	Only one study and an extremely small sample size limit generalizability.
Metacognitive strategies	3	7, 8	*d*= .161 to 1.94	Small number of studies make it difficult to generalize.
Definitions	3	2, 3, 6	NR	Most commonly used as a comparison group.
Visual Cues	2	3, 4	$\eta^2_p = .55$	Small number of studies make it difficult to generalize.
Writing	2		NR	Small number of studies make it difficult to generalize.

Note: NR = not reported specifically for the identified strategy

for studies without control or comparison groups. Standardized mean effect sizes calculated are reported in Table II.4.

Results

Findings of Vocabulary Strategies

The goal of the research question one was to report available empirical evidence of vocabulary instruction published since 1999. Findings are reported by strategy and summarized in Table II.5.

Contextual analysis was the most common vocabulary instruction identified. Of the 13 studies, five were conducted by different research teams and eight were conducted by three different research teams. Research was conducted in second through sixth grade in both reading and social studies classrooms. Effect sizes ranged from d = -.01 to 3.17. Findings generally favored contextual analysis instruction for increasing vocabulary learning. In addition to learning words using contextual analysis, several researchers have looked at the nature of the contextual clues in the text. Cain, Oakhill, and Elbro (2003) and Cain, Oakhill, and Lemmon (2004) found that the location of the clue can positively impact vocabulary inference.

Morphological analysis was considered in four studies using both expository and narrative text (Baumann, Edwards, Boland, Olejnik, & Kame'enui, 2003a; Baumann, Edwards, Font, Tereshinski, Kame'enui, & Olejnik, 2002; Lubliner & Smetana, 2005; Nash & Snowling, 2006). Research was conducted in third and fifth grade classrooms. Two studies were conducted in social studies classrooms. Reported effect sizes ranged from d = -1.00 to 1.58 in studies comparing morphological analysis to other strategies.

When morphological analysis was a component of the vocabulary treatment, effect sizes ranged from *d* = -.12 to 1.74. In general, findings for morphological analysis were positive. Baumann et al. (2002) found that when contextual and morphological analysis were combined, effect sizes for morphology over the combined effect was *d* = 0.59 and for context was *d* = 0.46

Semantic analyses, focusing on the relationships of words, were included in seven of the 24 studies. Semantic analyses included discussions, word definitions, characteristics of the word, examples and non-examples, contextual examples, visuals, graphic organizers, oral discussions, and listing of related words. Findings on the use of semantic analyses were mixed. Eight studies in second through seventh grades reported effect sizes from *d* = 0.55 to 3.17. One study showed differential differences depending on the initial achievement level of students (Apthorp, 2006), with students who had initially low vocabulary scores making larger gains, while another study showed preliminary evidence that semantic representations are more “durable” (Nash & Snowling, 2006, p. 349), especially for those children with limited vocabulary knowledge.

Mnemonic*s* to increase vocabulary knowledge was used in two studies. Mnemonics has been compared to definition instruction (Uberti, Scruggs, & Mastropieri, 2003) and contextual instruction (Jones, Levin, Levin, & Beitzel, 2000). Jones et al. (2000) compared the mnemonic condition to the context condition resulting in effect sizes of *d* = 2.85 and 2.04 for immediate and delayed recall, respectively, concluding that "mnemonically instructed students performed better than non-mnemonic semantic-

context students on all dependent measures, including recall of story information" (Jones et al., 2000, p. 260).

Explicit instruction was a component of five studies. Generally, explicit vocabulary instruction was most often combined with contextual analysis. A common finding was that explicit, teacher-led lessons on vocabulary enhanced student learning and recall of vocabulary (Baumann et al., 2003a; Cain et al., 2004). The statistically significant results of one study situated in social studies may provide initial evidence that explicit instruction may be more critical in expository text where the number of content-concept laden words is denser (Armbruster & Anderson, 1988). Unfortunately, the studies in this review did not compare explicit instruction to non-explicit instruction.

Incidental word learning was the focus of one study (8.3%) (Carlisle, Fleming, & Gudbrandsen, 2000) situated in fifth and eighth grade science classrooms during five-week units. While no effect sizes were reported, χ^2 results for fifth grade students showed significant differences of understanding between the pretest and post-test for science topical words at all ability levels. Corroborating the findings of other studies, the benefit to below average students was not as great.

Multiple exposures and multiple strategies are principles of vocabulary instruction often incorporated into vocabulary interventions. Ten of the 24 studies (41.7%) used multiple exposures to vocabulary words as a component of the vocabulary treatment. Additionally, multiple strategies were used in two-thirds of the studies (16 of the 24 studies). Because every study used a different combination and dosage of strategies, it was difficult to deconstruct the treatment and discern the most salient

practices for increasing vocabulary learning. Since these features were not studied in isolation, the discrete results for multiple exposures and multiple strategies are difficult to determine.

Several instructional strategies not previously identified in NRP (2000) or vocabulary research reviews were found in this review. These include multimedia presentation, instructional feedback, metacognitive strategies, definitions, and the use of pairs.

Multimedia presentation was used in one study (Xin & Rieth, 2001), which utilized video presentations to enhance vocabulary instruction. Students receiving vocabulary instruction anchored in video presentation statistically outperformed traditional dictionary instruction. While several studies were identified in the initial search that involved computer assisted instruction or multimedia presentation, upon closer inspection these studies involved second language learners or students whose native language was not English.

Instructional feedback regarding vocabulary knowledge in social studies was used in one study with a small sample size (n = 3): this study concluded that teachers' instructional feedback alone was insufficient for vocabulary learning (Ross & Stevens, 2003).

Metacognitive strategies, used in three studies (Boulware-Gooden, Carreker, Thornhill, & Joshi, 2007; Lubliner & Smetana, 2005; Twyman, McCleery, & Tindal, 2006), is instruction "to help children monitor comprehension of words and internalize and implement word-learning strategies to increase comprehension of natural texts"

(Lubliner & Smetana, 2005, p. 165). Results of studies using metacognitive strategies were generally favorable, with reported effect sizes of $d = 0.161$ to 1.94.

Definitions were used for comparison in vocabulary studies (Nash & Snowling, 2006; Uberti et al., 2003; Veerkamp, Kamps, & Cooper, 2007; Xin & Rieth, 2001). Gardner (2007) studied the difference between original definitions and revised definitions, finding that revised definitions were more effective than original definitions for all levels of readers, thus supporting the findings of prior researchers.

Three studies examined the use of groups or pairs with middle school students. In general, students who worked in small groups or pairs learned more vocabulary and were more confident in their knowledge (Jones et al., 2000; Stevens, 2003; Veerkamp et al., 2007).

Expository text was incorporated in eight studies (33.3%) (one math, four science, three social studies), using a variety of strategies. Tywman, McCleery, & Tindal (2006) found that students who received instruction aimed at increasing domain conceptual knowledge performed statistically better in vocabulary, possibly as a result of higher quality instruction and more time devoted to instruction. Support for multi-strategic instruction especially in the content areas, including "focused discussion of the meanings of key terms and a concerted effort to bridge initial understandings of word meanings and their scientific uses might help students learn the ideas and information of the unit" (Carlisle et al., 2000, p. 207).

Findings on the Methodological Characteristics of the Studies

Research question two focused on describing the methodological characteristics of recently published research studies involving vocabulary instructional strategies in second grade and above. First, a description of study characteristics (including design, sample, intervention, and statistical characteristics) will be provided, followed by a summary of research findings organized by study characteristics.

Description of Study Characteristics. The 24 studies reviewed were published in 18 different journals, and listed 48 different authors. *Education and Treatment of Children* and the *Journal of Educational Psychology* had the most articles (two each). Seven studies were published in 2007, which is as many as were published from 2004 to 2006.

Design characteristics included the research paradigm, design, and selection of groups. Due to the inclusion/exclusion characteristics, only experimental or quasi-experimental studies were included. Five identified studies used a mixed methods design. The most common designs were pretest–post-test comparisons (n = 16) or post-test only (n = 4). Randomized assignment was used by ten studies, yet only one reported the method of randomization. Three of the 24 studies were conducted at multiple sites.

The sample sizes of these studies varied considerably, ranging from three to 3,916 students, with one study not reporting the sample size. Seven studies had sample sizes greater than 100. Teacher numbers varied from one to 49. The grade levels of students in these studies ranged from second through eighth grade, with over two-thirds of the studies conducted in third through sixth grade. Of the studies analyzed, 20.8% (n

= 5) reported characteristics of the sample, such as race, gender, and economic standing. Six of the 24 (25%) studies included students with learning disabilities. Using the sample characteristic information found in the articles, ten (41.7%) were extensive enough to enable the reader to replicate the sample. Eight of the studies reported attrition.

Half the studies described both the intervention and either the comparison or control group, and two-thirds provided a description of the interventionist(s). In nine studies (37.5%), the treatment was delivered by the researcher outside of regular classroom instruction. Treatment was provided by both the classroom teacher and researcher in one study. In the remaining 14 studies (58%), treatment was provided by the classroom teacher. The length of intervention varied greatly, from one instructional session to one school year (36 weeks), with sixteen studies lasting longer than four weeks (66.7%). An important aspect of a study is the fidelity of implementation of the prescribed treatment. Fidelity is defined as the extent to which the expected intervention is implemented (Gresham, MacMillan, Beebe-Frankenberger, & Bocian, 2000), while the quality of fidelity refers to how well the intervention was implemented. Although fidelity of implementation was reported in 37.5% (n = 9) of the studies, none reported the quality of fidelity.

Of the 24 studies analyzed, over half used both standardized and non-standardized measures (n = 13, 54.2%), with eight using delayed post-tests. The most common measures used were the Gates-McGinite Reading Test (GMRT) (n = 7) and researcher-developed measures (n = 19).

Groups were comparable or differences were controlled for in 11 of the 24 studies (45.83%). The level of analysis matched the level of assignment in 19 of the 24 studies (80%), although one study addressed the mismatch in assignment in the discussion.

The most common statistical analyses employed were bivariate statistics such as, analysis of variance (ANOVA) and Chi-square. One study used hierarchical linear modeling. The most common statistical tests employed were t-tests, two-way ANOVA, and post hoc tests (n = 6 for each). This may be due to the relatively small sample sizes in the studies. Fourteen studies (58.3%) reported effect sizes. Of these 14, five used Cohen's *d*, one reported η^2, three reported Cohen's *d* and η^2, one reported η_p^2, and four reported other measures of effect size.

Effects by Study Characteristics

Effects by Study Design. Several study designs are represented in the corpus of studies with the most common ones being pretest–post-test, post-test only, and repeated measures. It would be expected from the research standards guidelines that studies combining pretest–post-test design with random assignment would provide the most valid results.

Sixteen studies used a pretest–post-test design. Of these studies, seven also employed random assignment. Standardized effect sizes were unable to be calculated for two of the studies because of the number of comparison groups (Baumann et al., 2002; Gardner, 2007). Effect sizes for vocabulary treatments in studies using randomized assignment ranged from small to large (ES range: 0.01 – 4.93, M = 1.29, SD = 1.11).

The largest effect was found for mnemonics over context instruction (Jones et al., 2000). Small effects were found for a multi-strategy treatment (Nelson & Stage, 2007). Effect sizes that compared interventions to comparison conditions were smaller when compared to no-treatment controls versus treatment groups ($F(1, 23) = 7.73$, $p < .05$).

Post-test only design was used in three publications reporting four studies (Cain, 2007; Cain et al., 2003; Cain et al., 2004). While these studies were generally single sessions with a researcher and effect sizes were not available across all studies, findings were primarily positive.

Repeated measures design was used in two studies. Fore et al. (2007) showed very strong effects with the repeated measures study of six students (from 3.7% at the pretest to 71.5% at the posttest, ES = 14.98). Although the results were not statistically significant, this could be due to the small number of studies in the sample.

Likewise, Twyman et al. (2006) showed strong effects at the final point between treatment and control conditions (ES = 0.81)

Effects by Treatment Design. The vocabulary studies reviewed either compared single strategies or tested the impact of a vocabulary treatment utilizing multiple strategies. As a result, an examination of effects by type of treatment was deemed appropriate.

Single strategy studies were uncommon in the current sample of studies. As noted earlier, contextual analysis (n = 5) was the most commonly studied vocabulary strategy (Baumann et al., 2003a; Baumann et al., 2002; Cain, 2007; Cain et al., 2003; Cain et al., 2004). Standardized effect sizes were favorable for the use of context (ES

range: 0.15 - 1.04). Incidental word learning was measured in one study, while morphological analysis, multiple exposures, and the use of feedback were the focus of one study each. The use of video (Xin & Rieth, 2001) in combination with other vocabulary strategies such as class discussions and illustrative sentences was shown to be effective.

Instructional packages, vocabulary treatments that used multiple evidence-based strategies, were more common in the reviewed body of literature (19 of the 24 studies). Although standardized effect sizes varied greatly (ES range: 0.01 to 14.98), overall, packages of vocabulary instruction were effective, though not statistically more effective than single subject studies. Common elements of the vocabulary treatments were a mix of delivery features, such as explicit instruction and multiple exposures, opportunities for supervised practice and specific vocabulary learning strategies like contextual or morphological analysis and graphic organizers.

Effects by Duration of Intervention. As reported above, the duration of many of the treatment were mixed, with five consisting of only one session, and 16 lasting longer than a month. Generally, findings did not statistically differ in regard to the treatment length. For studies that were longer than four weeks, standard effect sizes ranged from 0.07 to 14.98 (M = 2.42, SD = 4.76), while shorter treatments ranged from -0.68 to 4.93 (M = 1.48, SD = 1.38).

Effects by Setting. The setting included the person providing treatment and the location of that treatment. Treatments were delivered by the researcher, the classroom teacher, or a combination of the two within classroom instruction or in a pull-out setting.

Of the 24 studies in this literature review, one used a combination of researcher and teacher in the classroom (Lubliner & Smetana, 2005), nine were delivered by the researcher alone, and the remaining ones were based in the classroom, with the classroom teacher providing instruction. Effect sizes were unable to be calculated in half of the studies delivered by the researcher (Cain, 2007; Cain et al., 2003; Gardner, 2007; Ross & Stevens, 2003). Although there was great variance in those studies with standardized effect sizes (ES: 1.04, 3.83, 4.93, and 14.98), all were positive supporting vocabulary instruction. When both the classroom teacher and researcher presented an extended multi-strategy treatment, they found that Title I students benefitted more from instruction. Effects of classroom treatments delivered by the teacher, while not as great as those delivered by the researcher, varied less with standardized effect sizes ranging with variances of 10.82 and 1.50 respectively.

Effects by Measures. There is evidence that effect size may be influenced by the type of measure used. Standardized tests often have smaller effects than researcher-developed measures, though the difference was not statistically significant. Apthorp (2006) showed gains at one site when using the GMRT (standardized effect sizes of 1.78 and 1.57). However, gains were not shown at a second site. Nelson and Stage (2007) reported effect sizes on GMRT as high as 0.28. Researcher-developed measures, most commonly used for pretest-posttest comparisons showed the largest standardized effect sizes, as high as 14.98 (Fore III et al., 2007). Baumann et al. (2003) used numerous researcher-developed assessment in contextual and morphological analysis with effect sizes ranging from $\eta^2 = .002$ to .42. Reliability on all measures was in the acceptable

range (.75-.92); therefore, reliability was not compromised. While standardized measures were often administered at the beginning of studies, few reported post scores on these same measures. Post measures were almost exclusively on researcher-developed measures. Two studies reported pre-post effects for both standardized and research-developed measures to compare the treatment results based on the type of measure, with statistically significant results obtained from both types of measures.

Discussion

Two distinct tasks were accomplished in this systematic literature review. First, a comprehensive summary of the empirical literature on vocabulary instruction published since NRP (2000) was conducted to determine advances in the field and gaps in current knowledge. Second, the research methodologies of studies conducted from 1999-2007 were compared to research standards proposed by Gersten and colleagues (2005), WWC (2006), the Coalition for Evidence-Based Policy (2003), the Society for Prevention Research (2004), and AERA (2006).

Several notable findings result from this literature review. The important role of vocabulary instruction as an essential component of reading instruction is further illustrated. Findings in this review corroborate those of past reviews and meta-analyses—that several vocabulary strategies have emerged that are effective for increasing students' vocabulary knowledge. The recommendations for future research directions in the NRP (2000) report are being heeded, with more studies situated in actual classrooms, across grade levels, and with students with various achievement levels. While instruction in vocabulary strategies is shown to have a positive impact on

vocabulary learning, with reported effect sizes as high as $d = 3.17$, the wide variation among studies exemplifies the multi-dimensional character of vocabulary acquisition and the numerous factors influencing the acquisition and maintenance of vocabulary learning. Presley, Disney, and Anderson (2007) recognized the need for research in packages of vocabulary instruction. Recent research studies report the impact of vocabulary instructional packages on student learning, with 15 of the 24 studies in this literature review examining vocabulary instruction utilizing multiple strategies.

Considering the calculated standardized effect sizes, students receiving vocabulary instruction outperformed those who did not receive instruction. This is consistent with the findings of others (Beck, McKeown, & Kucan, 2002; Ebbers & Denton, 2008); that vocabulary instruction shows an increase in vocabulary knowledge. In practical terms, some vocabulary instruction is better than none at all.

To summarize the findings of this literature review regarding vocabulary instructional strategies, while second through eighth grades were represented in this literature review, the majority were in fifth grade, and only one study was conducted in second grade. Contextual analysis, repeated exposures, and semantic strategies studies were the most common and in the widest grade range. Conversely, only one study was identified using alternative presentations, such as computer-assisted or video. Technology's use to enhance vocabulary acquisition and maintenance is a present gap in the knowledge base. Strategies beginning to show promise, but in need of more evidence are the use of pairs, visual cues, or writing activities to build vocabulary knowledge, as well as the use of metacognitive strategies. While one-third of the studies used

expository text, not all strategies or grade levels were represented. Research informing vocabulary instructional practice in the content areas is needed, especially in fourth grade and above where learning from text gains importance.

While effective vocabulary instruction is paramount, the rigor of the research from which these results are obtained is of equal importance. Four broad areas of methodological characteristics were described. These were design characteristics, sample characteristics, intervention characteristics, and statistical characteristics. In general, experimental and quasi-experimental studies were common, yet the use of randomization is not yet consistent. This finding are similar to the analysis of educational intervention research in four educational psychology journals by Hsieh, Acee, Chung, et al. (2005) concluding that the percentage of published articles using randomization has declined since 1983. An interesting finding is the disparity of sample sizes in the identified studies (ranging from three to 3,916 students). Although sample descriptions in the studies generally are insufficient for replication studies, adequate information was generally provided to determine if groups are comparable across conditions (n = 13). A methodological strong point is the intervention descriptions, with all studies providing sufficient descriptions of the strategies used and how they were used. The number of studies (58%) conducted in the context of the classroom is another positive finding of this review. Reporting of fidelity of implementation as well as the quality of that implementation are areas of need. Of the studies in this review, only nine reported fidelity of implementation and not one addressed the quality of the implementation. Statistical characteristics is the last general category of methodological

characteristics. In the vocabulary instructional literature, standardized measures are often used for initial grouping and achievement levels, but changes in achievement are most often documented with researcher-developed measures only. This strong dependence on researcher-developed measures may be indicative of the need for accurate measures of vocabulary knowledge, which was also addressed in the NRP (2000) report. While the majority of studies reported effect sizes, numerous methods are used to calculate these results. Hsieh et al. (2005) found that the reporting of effect sizes increased when journals adopted editorial policies that report their reporting. The level of analysis in the vocabulary studies is often limited to t-tests and two-way ANOVA. In sum, while the body of research is generally approaching the rigorous research standards in the areas of design and intervention characteristics, the areas of sample and statistical characteristics have room for improvement.

Based on the findings, recommendations for future research are discussed in the following section. Vocabulary instruction can increase vocabulary knowledge; yet, the best program of instruction, including the type and amount of vocabulary instruction to improve students' vocabulary, is still elusive. Several studies have also shown differential effects for different populations, especially those with varying initial vocabulary levels. For this reason, part of the vocabulary instruction blueprint must also consider student characteristics. Currently, most studies are conducted with small to medium sample sizes, lacking the robustness needed to have respectable external validity. Furthermore, studies designed to determine the most appropriate amount of time to devote to vocabulary instruction, whether daily or weekly, are needed. The

limited implementation length of studies seems contradictory to what is known about vocabulary; i.e., vocabulary growth is incremental over time and requires multiple exposures. Long-term studies to measure not only the acquisition, but also the maintenance of vocabulary are desirable. While theoretically, it can be hypothesized which strategies facilitate long term vocabulary growth, there is a lack empirical confirmation of these ideas.

One quandary in vocabulary research is the best method of measurement of vocabulary acquisition. In this review, over half of the studies used both standardized and researcher-developed measures. Yet only two studies used standardized measures as both pretest and post-test measures. The measurement of vocabulary is a hurdle in vocabulary research because one must account for the multi-dimensional nature of vocabulary. Because of the conceptual nature of content vocabulary, the depth of understanding of a word may be just as important as the sheer number of words known. As evidenced in the varied effect sizes, measures sensitive to small increments in vocabulary learning are needed. Measurements that capture the multi-dimensional aspects of vocabulary knowledge are also needed. Effects on standardized measures are difficult to obtain, while researcher-developed measures may not be sensitive to the transferability of the learned vocabulary.

A component of vocabulary that is often ignored is the retention of word knowledge over time. Eight studies utilized delayed post-testing, thus knowledge of which vocabulary instructional strategies support long-term learning of words is mostly unknown. Therefore; measures must be sensitive to the breadth, as well as depth of word

knowledge, while also determining the difference in short-term learning and long-term maintenance.

The next generation of vocabulary research must commit to advancing the field through the use of best research design, practices, analysis, and reporting. Improvement of the quality of vocabulary research necessitates that researchers be conscientious in reporting and provide adequate descriptions of the participants, interventionists, and intervention. Establishment of a line of testable findings cannot be achieved through replication inquiries unless all characteristic features of studies are reported. The use of randomized intervention studies is needed to advance the field (Pressley & Harris, 1994; Shavelson & Towne, 2002). Researchers must strive to eliminate a primary source of confusion when comparing studies, which is the lack of a clear set of terminology for instructional practices and research methodological characteristics. While the advent of WWC has helped to standardize research terms, it is not widespread. For example, the concepts of comparison groups, control groups, and treatment groups are often used interchangeably. Reporting of studies must share both practical and statistical implications and effects. For practitioners, the reporting of effect sizes can help them determine the practical impact of the intervention.

The multi-dimensional nature of vocabulary calls for the use of high levels of statistical analysis and the use of statistical models to determine which factors contribute most to vocabulary learning. Through the use of hierarchical linear modeling and structural equation modeling, the many factors that contribute to vocabulary learning may be discovered.

Conclusion

The recommendations of NRP are as relevant today as they were in 2000. For practitioners, this literature review emphasizes the need for vocabulary instruction to be an integral part of reading instruction as students begin to use reading as a learning tool. For researchers, this study informs the next steps for vocabulary instruction research topics and methodology. Also, future directions for research topics and methodologies may be drawn from the summary provided.

CHAPTER III

A STUDY OF VOCABULARY INSTRUCTION IN FOURTH GRADE SOCIAL STUDIES CLASSROOMS

By fourth grade, students are assumed to have sufficient reading skill to decode words, comprehend passages, and learn from text. Recent findings from the National Assessment of Educational Progress (NAEP) (Lee et al., 2007), however, do not confirm this assumption because almost two-thirds of fourth graders struggled with on-level reading tasks. As students enter the later elementary years and continue into middle school and high school, content area text plays a larger role in their education. At the same time, students transition from learning to read to using reading as a skill to obtain information (Chall, 1996).

One critical factor required for text comprehension is vocabulary knowledge (Anderson & Freebody, 1981; Mezynski, 1983; Schatschneider, Buck, Torgesen, Wagner, Hassler, Hecht, & Powell-Smith, 2004; Stahl & Fairbanks, 1986), a factor that contributes significantly to the difficulty of content area text (Baxter & Reddy, 2007; Biancarosa & Snow, 2006; Hirsch, 2003). As students transition from third to fourth grade, they are often challenged by new vocabulary coupled with new concepts (Armbruster & Gudbrandsen, 1986). Chall, Jacobs, and Baldwin (1990) found that although students in second and third grade performed on level, in fourth grade, word meaning scores were approximately one year behind, growing to a gap of more than two years by seventh grade for all readers. It is evident that vocabulary plays a significant role

in reading comprehension, and, consequently, is an integral part of both the reading and content area curricula.

The strategies and skills required of content area text are numerous (de Leon, 2002), while at the same time students receive little exposure to this type of text prior to the later elementary years (Duke, 2000; Moss & Newton, 2002). Content area texts require students to decode and understand large words, connect prior knowledge with new ideas, and summarize and organize information in a genre where the content, vocabulary, and syntax are unfamiliar. Not only do content area texts in general differ from narrative text, but each content area (e.g., science, social studies, mathematics) has its own set of vocabulary and common text structures (Heller & Greenleaf, 2007). In short, students are accountable for learning new information from a text that is more difficult in reading level, vocabulary, content, and organization. The increased difficulty of the text may be one explanation for the fourth grade slump (Chall et al., 1990).

Vocabulary and Social Studies Instruction and Learning

A primary source used to teach social studies is the textbook. Unfortunately, social studies textbooks have been labeled inconsiderate (Armbruster & Anderson, 1988), bombardments of unfamiliar concepts (Harmon et al., 2005), and vocabulary- and conceptually dense (Blachowicz & Fisher, 2000). Since students must understand 90 - 95% of the words in a text to adequately comprehend (Nagy & Scott, 2000), difficulty in comprehension of social studies texts can be attributed, in part, to the high density of unfamiliar vocabulary (Baumann et al., 2003a; Biemiller, 2003a).

In content area text, as the number of unfamiliar words increases, the nature of the words also changes. Content area vocabulary is highly specialized and cognitively challenging, requiring students to use critical reading skills (Gardner, 2004) and sophisticated decoding skills to access information. Often referred to as academic vocabulary, these words carry much of the content load. Additionally, students may have not previously seen or heard many of these words, and many have multiple meanings. For example, in a typical 100-word narrative passage, students can encounter ten unknown words and still maintain adequate comprehension. Unknown words in content area texts are generally concept words (e.g., democracy, independence, government, treaty), so with only 90% accuracy, both reading comprehension and content learning can be greatly compromised. When the complexity of the texts outpaces students' vocabulary knowledge and reading comprehension skills, content learning decreases. There is evidence that this relationship may be reciprocal; as students' content knowledge improves, so does their ability to comprehend texts (Schatschneider et al., 2004). Knowledge of a subject implies knowledge of the subjects' vocabulary; therefore, vocabulary knowledge may be the link that connects reading comprehension and learning from content area texts. The next step then is to determine effective instructional strategies to build students' vocabulary knowledge as they begin to encounter texts that are conceptually and structurally challenging.

Evidence of Effective Vocabulary Strategies

A converging set of evidence-based vocabulary strategies has been established over the past 15 years to enhance vocabulary learning. The evidence is gathered from a

collection of primary studies, research reviews, meta-analyses, and book chapters. The majority of studies involved students in third grade to high school, with many focused in the content areas (Baker et al., 1998; Baumann et al., 2003a; Bryant et al., 2003; Harmon et al., 2005). From this base of evidence, several strategies have either been applied to content area texts or show promise as a means to enhance vocabulary learning. Each will be discussed in the following section.

Explicit Instruction

Explicit vocabulary instruction has been shown to be an effective delivery method of vocabulary instruction. Explicit instruction is defined as the intentional explanation and modeling of information or procedures in order to help students acquire information or strategies. The importance of explicit vocabulary instruction of key vocabulary in enhancing students' acquisition of word meaning has been documented in several research syntheses (Baumann et al., 2003b; Fukkink & de Glopper, 1998; Harmon et al., 2005; Jitendra et al., 2004; NRP, 2000; Read, 2004). Focusing only on content area vocabulary strategies, Harmon et al. (2005) reported that explicitly preteaching key words may be especially critical to students' learning in social studies.

Contextual Analysis

Contextual analysis has the largest empirical base of all vocabulary instructional interventions (Baumann et al., 2003a; Baumann et al., 2005; Baumann & Kame'enui, 2003; Edwards, Font, Baumann, & Boland, 2004; Fukkink & de Glopper, 1998; Harmon et al., 2005; Stahl & Fairbanks, 1986). Contextual analysis is the use of clues within the text to derive word meanings. These strategies may be explicitly taught or occur

incidentally. Comparing students who received contextual analysis instruction to those who did not resulted in large effect sizes (mean effect sizes: 0.49 to 0.87) for lesson dependent measures (Baumann et al., 2005). The meta-analysis by Fukkink and de Glopper (1998) showed significant positive effects of contextual analysis instruction, with a generalized effect size of 0.57.

Morphological Analysis

Morphology is the study of "the structure of words in terms of morphemes…that is, prefixes, roots, and suffixes" (Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003). By breaking words into meaningful parts (root and affixes), the meaning can often be inferred. Furthermore, since many content area words share common Greek and Latin roots, morphological analysis can support vocabulary growth (Harmon et al., 2005). Comparing students who received morphemic instruction to those who did not resulted in effect sizes ranging from $d = 0.30$ to 1.32 on lesson dependent words and $d = 0.12$ to 1.01 on measures of transfer words (Baumann et al., 2002).

Morphological analysis is sometimes used in conjunction with contextual analysis to provide a more comprehensive strategy to access word meaning. Baumann et al. (2002) compared fifth grade students in four groups (morphemic only, context only, morphemic-context, and instructed control). Results indicated both practically and statistically positive results when students received either morphological or contextual instruction in isolation or in combination when compared to the control group ($d = 0.30$ to 1.32). Situated in fifth grade social studies, Baumann et al. (2003a) found that when taught morphological and contextual strategies in combination, students were able to use

these strategies to learn new words when compared to the textbook vocabulary group. Descriptive data from written questionnaires and interviews revealed increased content learning as a result of this vocabulary treatment. Teachers commented that students saw the connection between reading class instruction and social studies instruction, while students shared that the strategies helped them understand the text better and know the words when they were reading.

One important advantage of contextual and morphological analysis is the assumption that these strategies transfer to different contexts. Once learned, students can apply these strategies independently in a variety of situations. While the question of transferability is still empirically unproven, evidence is beginning to emerge that this may be the case (Baumann et al., 2003a; Baumann et al., 2002; Martin-Chang, Levy, & O'Neil, 2007).

Multiple Exposures

Multiple exposures to target vocabulary, or repeated practice, are designed to reinforce learning and occurs when students have multiple opportunities to encounter and use identified vocabulary words. Repeated exposures can be planned through games, activities, repeated readings, technology applications, or discussions. Numerous research reviews support the practice of repeated exposures to learn new vocabulary for all age levels, from elementary (Beck et al., 2002; NRP, 2000) to junior high and high school (Bryant et al., 2003; Harmon et al., 2005; NRP, 2000). Mean effect sizes for multiple exposures ranged from 1.32 to 2.33 in Stahl and Fairbanks' meta-analysis (1986).

Semantic Organizers

Semantic organizers are designed to help students categorize new vocabulary in meaningful ways (Heimlich & Pittelman, 1986). These most often take the form of graphic organizers, such as semantic maps or semantic features analysis. Numerous studies and reviews have supported the use of semantic organizers to enhance vocabulary and content instruction (Best, Dockrell, & Braisby, 2006; Bos & Anders, 1990; Heimlich & Pittelman, 1986; Moats, 2004; Pittelman, Heimlich, Berglund, & French, 1991). In a review of vocabulary strategies, Bryant et al. (2003) reported that students who utilized semantic strategies outperformed students who only received definition instruction on both immediate and delayed post tests. Jitendra et al. (2004) found a mean effect size of 1.10 for semantic strategies based on their review of 10 studies. Follow up analyses in seven studies also showed a large effect size (ES = 0.94), indicating that semantic organizational strategies enabled students to maintain word meaning knowledge.

Mnemonic Strategies

Mnemonic strategies "refers to systematic procedures specifically designed to improve one's memory" (Pressley, Levin, & McDaniel, 1987, p. 109) and may include keywords, pictures, or phrases to aid in definition recall. Stahl and Fairbanks (1986) reported mean effect sizes for the keyword method on contextual tasks and definitional tasks (0.66 and 0.57, respectively), but noted that effect sizes varied greatly, from 2.01 to -2.89. Jitendra et al. (2004) examined six studies that found the keyword approach, with a mean effect size of 1.93, with additional positive results for maintenance and transfer.

Additional reviews corroborate the positive impact of mnemonic strategies (Baker et al., 1998; Baumann et al., 2003b; Bryant et al., 2003; NRP, 2000).

Use of Technology

The use of technology was identified by NRP (2000) as strategy to enhance vocabulary learning. This conclusion is further supported by other vocabulary reviews (Baker et al., 1998; Jitendra et al., 2004). Studies included computer programs that presented drill and practice activities, programs that provided definitions and context, the use of digitized speech, computer-mediated texts (for example, students can link to definitions of unknown words), and the use of video to enhance vocabulary instruction.

Combined Strategies

Multiple or combined strategies are often used in vocabulary instruction. Combinations commonly found included graphic organizers in combination with semantic strategies, the combination of contextual and morphological analysis, or explicit teaching with any of the above mentioned strategies. Positive results for using explicit instruction in morphological analysis, contextual analysis, and a combination of these two were obtained by Baumann and colleagues in two studies (2003a; 2002). NRP (2000) concluded that no one method of vocabulary instruction is best. Multiple instructional methods are necessary to meet the varying age and ability levels of students. Others have echoed that because vocabulary knowledge involves increasing both depth and breadth of word meaning, a single strategy will not be sufficient (Beck et al., 2002; Bryant et al., 2003; Graves, 2006; Snow, 2002).

Knowledge of effective vocabulary strategies has continued to expand over the last 10 years. While much of this knowledge is based in narrative text, content area vocabulary strategies can be inferred. Recently, The National Literacy Institute report, *What Content-Area Teachers Should Know about Adolescent Literacy* (2007), outlined recommendations for content area vocabulary instruction including the importance of preteaching difficult words, using explicit instruction, incorporating prior knowledge, providing multiple practice opportunities, and using computer technology for additional practice, all of which have been discussed above.

There is converging empirical evidence identifying the positive impact of several vocabulary strategies, most of which have been applied to content area text. As the evidence base accrues to support vocabulary instruction and independent word learning strategies, it is important to understand the degree to which the knowledge base is being implemented in classrooms, particularly content area classrooms.

Observational Studies of Vocabulary Instruction in Content Area Classrooms

Given the difficulty of content area text and what is known effectively helps students navigate conceptually dense vocabulary, the logical next question becomes whether and to what degree teachers use validated vocabulary strategies in social studies instruction? This section reviews observational studies in social studies classrooms that included a vocabulary focus. A search for vocabulary observational studies in social studies classrooms was conducted using Academic Search Premier, Education Full Text (Wilson), Linguistics + Language, PsycINFO 1872-current, Social Sciences Full Text,

Education: A SAGE Full-Text Collection, ISI Web of Knowledge, and Google Scholar, resulting in three studies (Bailey, Shaw, & Hollifield, 2006; Durkin, 1978-1979; Scott, Jamieson-Noel, & Asselin, 2003). Each of these studies varied in their primary focus. Durkin (1978-79) and Bailey et al. (2006) observed only periods of social studies instruction, whereas Scott and colleagues (2003) observed entire school days. Scott and colleagues (2003) focused only on vocabulary instruction, while Durkin (1978-79) and Bailey et al. (2006) observed vocabulary only as a portion of their overall observation. The following section details each study's sample, methods, and results. In addition, Table III.1 provides a brief overview of each of the social studies vocabulary observational studies.

In 1978-79, Durkin conducted one of the first observational studies examining reading instructional practices in fourth grade classrooms. As part of a more comprehensive study, she observed 24 fourth grade teachers for a total of 2,775 minutes of social studies instruction. Although Durkin's primary focus was comprehension instruction, vocabulary instruction was embedded in the comprehension instruction and coded as preparation for reading. Durkin also included providing background knowledge in this code; therefore, it is impossible to disaggregate specific vocabulary instruction from background knowledge development. Durkin reported that 1.73% of the total observed social studies time was devoted to preparation for reading. Unfortunately, the exact amount of time devoted to vocabulary is not able to be determined. In a sub-study, Durkin observed 12 teachers in third, fourth, and sixth grades for a total of 975 minutes. Using the same codes, no instances of preparation for reading were observed, which one

Table III. 1.
Summary of Previous Vocabulary Observational Studies in Social Studies Classrooms

Study	Grade Level	Subject	Sample Size	Amount of Observation	Major Finding
Durkin 1 (1978-79)	4	Reading and Social Studies	24 teachers	3 consecutive days 2,775 minutes in Social Studies	*practically no time is allocated to comprehension instruction in reading or social studies
Bailey et al. (2006)	K-5	Social Studies	39 classrooms	5 days per week for 17 weeks	*most prominent activities during social studies were reading the book and answering questions, whole class discussions, and individual activities
Scott et al. (2003)	5-7	Vocabulary	23 classrooms	3 consecutive days 18,503 minutes	*6% of school time devoted to vocabulary development and only 1.4% devoted to academic vocabulary

can infer to suggest that no preteaching or development of vocabulary knowledge was observed.

Bailey et al. (2006) focused on all aspects of social studies instruction in their observational study. The purposes of this study were to describe social studies instruction in terms of actual total time, types of instructional strategies used, and the use of technology. The authors were motivated by the lack of curricular standards for social studies in No Child Left Behind (NCLB) ("No Child Left Behind Act of 2001,") and the resulting impact on instruction. Observations of 39 social studies teachers' instruction in Kindergarten to fifth grade were conducted by elementary education pre-service teachers during the social studies methods block for 13 weeks during the spring semester and 14 weeks during the fall semester. A Social Studies Teaching Log recording the total minutes of social studies instruction and the instructional strategies used by the mentor teacher was completed each week by the preservice teacher. Strategies coded included (a) read the book and answer questions, (b) define vocabulary words, (c) hands-on activity from book, (d) hands-on activity not from book, (e) inquiry lesson, (f) whole group discussion/activity, (g) cooperative learning group activity, and (h) individualized activity. Results were reported by grade level. Three fourth grade classrooms were included in the spring sample, resulting in an average of 15.9 minutes of social studies instruction a day (Range: 13.2 to 20.9 min). In the fall, daily social studies instruction of two teachers ranged from 21.9 to 27.9 minutes, with an average of 24.7 minutes. Four teachers skipped entire weeks of instruction (a combined total of 13 weeks), with one teacher not teaching social studies for seven weeks. Results for third (n = 5) and fifth (n =

10) grade mirror those of fourth. Daily social studies instruction averaged 24.2 minutes in third grade and 14.9 minutes in fifth grade. Again, teachers often skipped social studies instruction, with only one teacher from each grade level teaching social studies weekly.

When analyzing classroom instruction in fourth grade, the most prominent activities observed were (a) reading a book and answering questions (47 instances), (b) whole group discussions/activities (41 instances), and (c) individualized activities (30 instances). Nineteen instances of defining vocabulary words were coded (one in the spring and 18 in the fall). Third and fifth grade saw many more instances of vocabulary activities, with a total of 98 instances in third grade and 39 instances in fifth grade. The authors reported that the overall findings were dismal in not only the amount of time devoted to social studies instruction, but the lack of consistency, use of unchallenging instructional strategies, and lack of technology inclusion.

Observing 23 fifth, sixth, and seventh grade classrooms in Canada, Scott, Jamieson-Noel, and Asselin (2003) observed "when, where, how often, and how effectively" (p. 269) vocabulary was taught throughout the school day. Observations were conducted on three consecutive days over a four-month period for a total of 608 hours of observation. According to their observations, academic vocabulary (i.e., vocabulary for math, science, social studies, or art) was observed for 266 minutes. Only two teachers included four instances of vocabulary in their social studies instruction for a total of 92 minutes distributed over three days. Overall, vocabulary was the focus of instruction for 1.4% of the total time, with social studies accounting for 0.5% of the total time. The most common vocabulary instruction was whole group, followed by individual

work. Scott and colleagues (2003) also examined the types of vocabulary instruction observed, although these results were not broken down by content area. Vocabulary activities were divided into definitional, contextual, semantic knowledge, structural roots and affixes, and mnemonic association. The researchers reported that there was much mentioning and assigning, yet little actual instruction.

Almost 30 years have passed since the first observational study of instructional practices in social studies. Despite limited observational studies, a consistent finding across these investigations is the little time devoted to vocabulary instruction in social studies. In addition, studies have largely focused on documenting whether or not vocabulary instruction occurred, not on the type of vocabulary instruction and practice. The purpose of this strategy use study is to evaluate whether and to what degree evidence-based vocabulary strategies are used in fourth grade social studies instruction. Several important questions regarding vocabulary instruction in social studies remain unanswered. First, as the body of evidence supporting vocabulary instruction accrues, it is important to know what the current vocabulary practices of social studies teachers are. With only three observational studies situated in social studies, knowledge of actual classroom vocabulary practice is limited. As a result of the recent research and subsequent recommendations for effective vocabulary instruction, this study aims to focus observations on specific vocabulary strategies to determine whether research knowledge has translated into classroom practice. While there are recognized strategies to guide vocabulary instruction, there is limited evidence that these strategies are being applied in practice.

Purpose of the Study

Because of the importance of the teacher in helping students transition to vocabulary-dense content area text, a unique feature of the present study is to determine the impact of professional development on teachers' vocabulary strategy use in social studies instruction. Additionally, this study sought to determine whether professional development in evidence-based strategies increases the quantity and quality of teachers' use of these strategies. Descriptions of instructional and classroom practices were coded from digital audio recordings; therefore in this study, the word *audio documentation* will be used to refer to the process used to gather information on strategies teachers used during instruction. Research questions to be answered are:

1. What evidence-based vocabulary strategies do fourth grade teachers use during social studies instruction? To what extent are these strategies used during social studies instruction?
2. Does professional development in evidence-based vocabulary strategies result in differential use in the quantity and/or quality of strategies by teachers during social studies instruction?

Method

Participants

Participants were 26 fourth grade social studies teachers in nine ethnically and economically diverse schools in two public school districts in central Texas. District A was a large district (14,618 students) with 17 elementary schools, five of which participated in the study. District B was also a large district (8,666 students) with four of

the eight elementary schools participating. The sample in this study was part of the *Enhancing the Quality of Expository Text Instruction and Comprehension through Content and Case-Situated Professional Development* grant funded by the Institute for Education Science (Simmons, Rupley, & Vaughn, 2005) (IES grant contract number R305M050121A).

Twenty-five of the teachers were female and one was male; all were elementary certified. Teaching experience ranged from first year teachers to a veteran teacher with 31 years of experience. The average participant had 7.36 years of teaching experience (SD = 8.4). Group demographic differences were explored using X^2 analysis. Teachers did not differ by degree, $X^2(1, N = 26) = .08$, $p > .05$; ethnicity, $X^2(1, N = 26) = .16$, $p > .05$; years teaching, $X^2(12, N = 26) = 14.32$, $p > .05$; gender, $X^2(1, N = 26) = 1.96$, $p > .05$; or additional certifications, $X^2(1, N = 26) = .41$, $p > .05$.

Table III.2 provides detailed demographics of study participants. One teacher was unable to complete the study because she changed grade levels during the course of the study.

Study Design

Schools were matched according to ethnicity and achievement, and then randomly assigned at the school level, resulting in nine teachers in the typical practice condition and 17 teachers in the professional development condition. It is important to note that District A was in the first year of an academic vocabulary program (Marzano & Pickering, 2005) required of both typical practice and professional development teachers. The typical practice condition did not receive any teacher or student materials or

Table III.2.
Teacher & Classroom Demographics

Characteristic	***Professional Development n (%)***	***Typical Practice n (%)***
Race	17	9
White	12 (70.6)	7 (77.8)
Hispanic	5 (29.4)	2 (22.2)
African-American	0 (0)	0 (0)
Gender		
Males	0 (0)	1 (11.1)
Females	17(100)	8 (88.9)
Teaching experience		
Average years teaching	5.35	11.4
Average years teaching grade 04	3.59	6.29
Degrees and certifications		
Advanced degrees	3 (17.6)	2 (22.2)
Reading specialist certification	1 (5.9)	1 (12.5)
Special education certification	2 (11.8)	1 (12.5)
ESL/Bilingual certification	10 (58.8)	5 (55.6)
Classroom organization		
Self-contained	12 (70.6)	5 (55.6)
Departmentalized	5 (29.4)	4 (44.4)

professional development training. This group was instructed to teach social studies as they normally would.

Teachers assigned to the professional development condition received both materials and professional development for the duration of the 18-week study. Teachers attended approximately 15 hours of professional development on the use of vocabulary

strategies situated within social studies instruction conducted by university personnel. Professional development sessions were distributed among three primary sessions and three teacher study team meetings over the course of the study. Strategies taught in the professional development sessions included the use of vocabulary maps to explicitly teach vocabulary words and expand their meaning, semantic features analysis to connect chapter content to new vocabulary, context clues for independent word learning, and vocabulary practice activities.

As part of the professional development program, materials aligned with state and district standards were provided for teachers to encourage their use of the vocabulary strategies in social studies instruction. Teachers received weekly lesson plans and PowerPoint files or transparencies to support their classroom instruction. Teachers were asked to implement 30-minute lessons three times per week, with the remainder of social studies instruction time planned by the teacher. Materials provided for students' use included folders containing vocabulary maps of target words, chapter overviews, and vocabulary cards.

Data Collection

Teacher Measures. Data collected from classroom teachers included surveys for demographics, teacher familiarity and use of reading and vocabulary strategies in social studies, and audio recordings of social studies instruction.

Demographics surveys requesting information on the degree or degrees, additional certifications, total years teaching, years teaching fourth grade, and classroom

organization were completed by each teacher. Teachers completed this survey at their convenience prior to the beginning of the study.

Social studies familiarity and use surveys were administered prior to the beginning of the study to determine teacher familiarity of reading strategies and use of those strategies in social studies instruction. Teachers rated their familiarity with five vocabulary strategies (practice activities, semantic maps, semantic features analysis, context clues, and explicit instruction) on a three-point scale (1 = *unfamiliar*, 2 = *somewhat familiar*, 3 = *very familiar*). Responses to these items were summed to create a Vocabulary Knowledge Score (maximum score = 15).

Classroom audio recordings were collected to document instructional strategy use and distribution of time during social studies instruction. Teachers digitally audio-recorded three social studies lessons over an 18-week period, one per six-week unit. Recorded lessons were either uploaded to the secure website by the teacher or collected from the teacher.

Strategy Documentation Instrument. To document teachers' strategy use and overall quality of instruction, the Instructional Content Emphasis-Revised (ICE-R) (Edmonds & Briggs, 2003) was modified. The ICE-R was created to record multiple dimensions of classroom reading instruction from the main category of reading instruction (e.g., comprehension or text reading) to more specific strategies (e.g., comprehension monitoring, silent reading). With training, reported inter-rater reliability on the ICE-R was 91%.

The ICE-R was modified for use in this study to more fully specify vocabulary strategies. In the following paragraphs, each additional dimension of the modified strategy use instrument will be discussed along with any changes. While ICE-R utilizes three dimensions of classroom activities, because the focus of this study was on vocabulary strategy use, I have restricted the documentation of strategy use to **Dimensions B** (instructional activities), and **C** (the focus of instruction). Dimension B, instructional activities, focused on specific activities within a larger instructional category. For example, Dimension B would include specific comprehension or vocabulary strategies (e.g., comprehension monitoring, definitional). This code will be elaborated on in the following section. The focus of instruction (Dimension C) was coded for each activity and included: instruction, review, application, assessment, or social studies content focus. In order to determine which code was most appropriate, the intention of the activity was analyzed. For example, initial introduction of a strategy would be coded as instruction, and questions or activities focused on determining student understanding of previously taught materials would be coded as assessment.

The most significant modifications were made to the comprehension dimension. In ICE-R, vocabulary was coded as a subcategory of comprehension instruction; however, the present study's focus on specific vocabulary strategies necessitated a more detailed documentation of vocabulary strategies. Specific instructional activities of vocabulary included definitional, explicit instruction of word meanings, contextual analysis, structural or morphological analysis, organizational or semantic strategies,

practice activities, keyword or mnemonic strategies, use of multiple strategies, and other strategies. Table III.3 summarizes the codes of the strategy use instrument.

Table III.3.
Three Instructional Dimensions of Social Studies Vocabulary Instruction in the Modified Strategy Use Instrument

Dimension A	**Dimension B**	**Dimension C (consistent across all dimensions)**
Oral language development	Teacher-initiated discussion Student-initiated discussion Teacher lecture	Instruction Review Application Assessment Social Studies Content
Text reading	Supported oral reading Choral reading Independent silent reading Independent oral reading Teacher reads aloud Teacher reads aloud while students read along Paired reading Other	
Comprehension	Prior knowledge/predicting Reading comprehension monitoring Listening comprehension monitoring Comprehension strategy instruction Other	

Table III.3 cont.

Dimension A	Dimension B	Dimension C (consistent across all dimensions)
Vocabulary	Definitional Explicit instruction Contextual analysis Morphological analysis Semantic Practice activities Keyword/mnemonics Multiple strategies Other	
Study Skills	Skimming Paraphrasing Use of guide words Outlining SQ3R Reading rate Use of references	
Logistical/ Instructional Tasks	Checking assignments Giving assignments Helping with assignments Collecting materials	
Logistical/ Noninstructional Tasks	Transitions Student redirection/correction Teacher correcting papers Other (fire drill, announcements, classroom visitors)	

Teachers' overall instructional quality was also rated on a scale from one to seven. One was considered instruction that was ineffective while a rating of seven signaled instruction that was considered highly effective. To determine the overall rating, discrete factors of instruction were indentified: teacher modeling, student practice, quality

feedback, adjusted instruction to meet the needs of students, progress monitoring, student participation, and appropriate pacing.

Piloting of the modified strategy use instrument was conducted with other recordings collected as part of the *Enhancing the Quality of Expository Text Instruction and Comprehension through Content and Case-Situated Professional Development grant funded by the Institute for Education Science* (Simmons et al., 2005) (IES grant contract number R305M050121A).

Coding of the audios was accomplished with three coders, all certified teachers with an average of 8.7 years of teaching experience. Inter-rater agreement was determined on several levels. First, time samples were considered consistent when beginning and ending times fit within a five-second window. Coders were .97 reliable across time codings. Inter-rater reliability was also determined for the s of instruction, with codings for Dimensions C highly reliable (.99). High reliability can be attributed to the use of recorded audios that allowed coders to replay portions or slow the pace of playback. Inter-rater reliability for Dimension B ranged from .83 to 1.00. Overall, inter-rater agreement on the coding of a pertinent sample of classroom audios including time and both dimensions was .91.

Data Analysis

Data analyses were conducted on several levels. Audio-recorded lessons were analyzed using the modified ICE-R strategy use instrument. The first level of analysis was on the specific strategies used during social studies instruction (Dimension B), and then the focus of instruction (Dimension C).

Because teachers submitted multiple recordings, individual results of strategy use were averaged by teacher to standardize each teacher's contribution. All data reported used averaged teacher results. Since differences in the two conditions (professional development and typical practice) are of interest to this study, average results were determined for each condition.

Further analyses were conducted on three dependent measures: time spent on vocabulary strategies, number of unique vocabulary strategies used, and the overall quality of instruction to determine how teachers in each condition were implementing evidence-based vocabulary strategies. The independent variable was the condition controlling for teacher preknowledge of vocabulary strategies using the Vocabulary Knowledge Score.

Results

Results of this study are organized according to the following framework. Descriptions of typical practice classroom instruction and professional development classroom instruction will first be addressed separately, and then compared. First the total time observed for social studies will be reported, followed by the most prominent vocabulary activities, in time and frequency (Dimension B). Last, the specific focus (Dimension C) of vocabulary instruction that occurred during the audio documentation will be reported by condition.

Description of Typical Social Studies Instruction

To determine what evidence-based strategies and to what extent typical practice fourth grade social studies use during social studies instruction, documentation of the

nine fourth grade social studies teachers who received no professional development were analyzed by Dimensions B and C. Teachers in the typical practice condition (n = 9) averaged 27.7 minutes (SD = 12.5) per lesson in social studies (Range: 13.0 – 48.7 min.). The total averaged observation time per lesson was 27.5 minutes for this condition.

Distribution of Time by Instructional Activity (Dimension B). When analyzing the data by instructional activity only (Dimension B), the most common classroom activities documented in typical practice classrooms were teacher questioning of content (24%), teacher lecturing on social studies content (19%), text reading (10%), and checking assignments (8%). Very little time was spent in student-initiated discussions (1%) or reading comprehension monitoring (2%) and no instances of listening comprehension monitoring or comprehension strategy instruction or use were observed. In actual classroom time, these percentages translate to approximately 17 seconds of student-initiated discussion and 34 seconds of reading comprehension monitoring.

Limited to documented vocabulary strategies, Table III.4 lists the mean documented frequency, mean time in seconds, and the total percentage of time. Time was reported in seconds due to the small amount of time observed. Focusing on vocabulary instruction only, 11% (approximately three minutes) of social studies was spent in vocabulary activities. The most common vocabulary activity documented in typical practice classrooms was the teacher or student providing the definition of a word, with an average documented frequency of 1.5 occurrences per lesson. Practice activities were the next most observed strategy, with an average documented frequency of 0.7 occurrences per lesson. No instances of semantic strategies or mnemonics were documented.

Table III.4.

Observed Use of Vocabulary Strategies by Frequency and Time in Seconds and by Condition (Dimension B)

Strategy	**Observed Frequency (M/SD)**		**Average time in seconds (M/SD)**		**Percentage of Total Time**	
	Typical practice	Professional Development	Typical Practice	Professional Development	Typical Practice	Professional Development
Definitional	1.46 (0.95)	1.01 (0.83)	62.10 (60.2)	42.3 (45.6)	3.77	1.79
Explicit instruction	0.28 (0.67)	1.31 (1.73)	36.50 (94.0)	150.9 (208.6)	2.21	6.41
Contextual analysis	0.09 (0.19)	0.83 (1.10)	7.10 (15.7)	48.8 (79.1)	0.43	2.07
Morphological analysis	0.09 (0.19)	0.08 (0.25)	2.07 (6.30)	10.31 (39.02)	0.13	0.44
Multiple exposures	0.65 (1.82)	1.22 (1.71)	16.30 (46.30)	210.66 (257.70)	.99	8.94
Semantic organizers	0.00 (0.00)	0.12 (0.19)	0.00 (0.00)	13.54 (31.02)	0.00	0.57
Mnemonic strategies	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00	0.00
Combined strategies	0.50 (1.50)	1.18 (1.11)	0.00 (0.00)	161.57 (201.23)	0.00	6.86
Other	0.50 (1.00)	0.31 (0.43)	50.92 (117.23)	10.96 (13.70)	3.09	0.47

With regards to time, the greatest amount of time was spent in definition instruction (62.1 sec), accounting for 4% of an average lesson. Other vocabulary activities (e.g., writing words, echoing words) accounted for 3% of an average lesson (50.9 sec). The third most common vocabulary activity recorded was explicit instruction (2%, 36.5 sec). Approximately 1% of the average lesson was devoted to vocabulary practice activities.

Distribution of Time by Instructional Focus (Dimension C). In addition to which vocabulary strategies were used, how these strategies were used was coded (Dimension C). The majority of all vocabulary activities observed in typical practice classrooms were instructional (teacher-led) or application, with a little over a minute each per lesson. Very little time was spent in review, practice, assessment, or content-based vocabulary activities. See Table III.5 for means and standard deviations for total time spent for each specific vocabulary strategy (Dimension B) and the instructional focus (Dimension C).

Description of Teachers with Professional Development

As a comparison to the typical practice teachers, strategy use data from 17 fourth grade social studies teachers who received professional development was also analyzed by Dimensions B and C. Teachers (n = 17) in the professional development condition averaged 40.1 minutes (SD = 9.3) per social studies lesson (Range: 21.0 – 51.6 min.).

Distribution of Time by Instructional Activity (Dimension B). The most common classroom activities (Dimension B) documented were the use of multiple strategies to learn vocabulary (17%), checking assignments (10%), and vocabulary practice activities (9%).

Table III.5.

Means and Standard Deviations for Time in Seconds for Vocabulary Strategies by Instructional Focus

Strategy	Instruction		Review		Application/ practice		Assessment		Content	
	TP M/SD	PD M/SD	TP M/SD	PD M/SD	TP M/SD	PD M/SD	TP M/SD	PD M/SD	TP M/SD	PD M/SD
Definitional	27.43/ 45.58	14.20/ 18.66	8.39/ 11.05	7.39/ 10.66	11.73/ 27.85	2.79/ 10.11	7.49/ 18.81	13.32/ 42.42	9.48/ 15.85	4.56/ 8.48
Explicit instruction	36.47/ 93.99	43.67/ 99.23	0.00/ 0.00	66.73/ 165.21	0.00/ 0.00	40.55/ 127.73	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00
Contextual analysis	0.00/ 0.00	0.44/ 1.89	0.00/ 0.00	16.06/ 29.18	4.87/ 14.81	30.79/ 62.45	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	1.50/ 6.36
Morphological analysis	0.00/ 0.00	1.11/ 4.71	0.00/ 0.00	0.00/ 0.00	2.07/ 6.30	9.19/ 39.01	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00
Multiple exposures	16.30/ 46.30	7.86/ 24.46	0.00/ 0.00	10.78/ 22.85	0.00/ 0.00	131.64/ 232.10	0.00/ 0.00	19.58/ 64.37	0.00/ 0.00	27.26/ 115.65
Semantic organizers	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	13.54/ 31.002	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00
Mnemonic strategies	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00
Combined strategies	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	161.57/ 201.23	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00
Other	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	50.92/ 117.23	10.96/ 13.70	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00	0.00/ 0.00

Note: All times reported in seconds.

Focusing on vocabulary instruction only, approximately 55% of social studies time was spent in vocabulary activities. Table III.4 lists the mean documented frequency, mean time in seconds, and the total percentage of time. The time was again reported in seconds due to the small amount of time each strategy or activity was observed. Explicit vocabulary instruction by the teacher was documented the most (1.3 occurrences per lesson) with an average time of 2.5 minutes per lesson. No instances of mnemonics were documented in the professional development classrooms. Vocabulary practice occurred on average of 3.5 minutes per lesson with a documented frequency of 1.2 occurrences per lesson. The next most common activities were multiple strategies and explicit instruction (7% and 6%, respectively). Lessons averaged 2.7 minutes of multiple strategies and 2.5 minutes of explicit instruction in the audio documented lessons.

Distribution of Time by Instructional Focus (Dimension C). In addition to which vocabulary strategies were used, how these strategies were used was also coded. The majority of all vocabulary activities documented were application activities, with an average of 3.5 minutes per documented lesson. The majority of the application focus was coded in multiple strategies (161.6 sec). The main activity documented during this time was students independently working with vocabulary maps that combined definition, context, and semantic connections. With an average of 131.6 seconds per lesson, application of practice activities was the next most documented activity. All vocabulary activities coded as other were application—for example, students using vocabulary words in a writing activity. While teachers in the professional development group used vocabulary activities across all instruction types, the smallest amounts of time were spent

with strategies that focus on assessment and content. Table III.5 provides means and standard deviations for each specific vocabulary strategy (Dimension B) and the instructional focus (Dimension C) documented per category.

The Impact of Professional Development on Documented Vocabulary Instruction

To determine if professional development in evidence-based vocabulary strategies results in differential use in the quantity and/or quality of strategies by teachers during social studies instruction, a between-groups analysis was conducted.

Overall instructional quality was determined for each teacher from the audio recordings. On a scale from 1 to 7, the average instructional quality of the audio documented professional development teachers was 4.9 (SD = 0.79), while the instructional quality score for the typical practice teachers was 3.6 (SD = 0.72).

A one-way between-groups multivariate analysis of covariance (MANCOVA) was performed to investigate differences between teachers in the typical practice and professional development conditions. Three dependent variables were used: instructional quality, time devoted to social studies instruction, and the unique number of strategies used. The independent variable was to which group the teacher was assigned: typical practice or professional development. Teachers' Vocabulary Knowledge Score was used as a covariate due to initial group differences, $X^2(8, N = 26) = 16.0$, $p < .05$. Preliminary assumption testing was conducted to check for normality, linearity, univariate and multivariate outliers, homogeneity of variance, and multicollinearity, with no serious violations noted. There was a statistically significant difference between typical practice and professional development teachers on the combined dependent variables: F(3, 21) =

7.32, p = .002; Wilks' Lambda = .49 ; η_p^2= .51. When the results for the dependent variables were considered separately, using a Bonferroni adjusted alpha level of .017, all reached statistical significance: for the unique number of strategies F(1, 24) = 10.43, p = .004, partial eta squared = .31; for instructional quality F(1, 24) = 15.15, p = .001, partial eta square = .40; and for average social studies instructional time F(1, 24) = 7.76, p = .01, partial eta squared = .25. R-square values for the three dependent variables were: instructional quality (R^2 = .416), the unique number of strategies (R^2 = .312), and average time (R^2 = .259). An inspection of the mean scores indicated that teachers who participated in professional development had higher mean scores on the three dependent variables. Results of the MANCOVA are reported in Tables III.6 and III.7. Table III.8 reports the mean scores for the three independent variables.

Discussion

The purpose of this strategy use study was to evaluate whether and to what degree evidence-based vocabulary strategies were used in fourth grade social studies classrooms and to determine whether professional development in evidence-based strategies increases the quantity and quality of teachers' use of these strategies. Results of the classroom documentation reveal little vocabulary instruction in fourth grade social studies classrooms, similar to previous findings of Durkin (1978-1979), Bailey et al. (2006), and Scott, et al. (2003). Unfortunately, the findings of these three studies are not relevant to the findings in the present study, particularly since only one study attempted to document particular vocabulary strategies. While the present study builds on these past inquiries, a direct comparison of results is quite difficult.

Table III.6

Multivariate Test Results of a MANCOVA of Group using Preknowledge as a Covariate

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.907	68.594 (a)	3.000	21.000	.000	.907
	Wilks' Lambda	.093	68.594 (a)	3.000	21.000	.000	.907
	Hotelling's Trace	9.799	68.594 (a)	3.000	21.000	.000	.907
	Roy's Largest Root	9.799	68.594 (a)	3.000	21.000	.000	.907
Preknowledge score	Pillai's Trace	.038	.273(a)	3.000	21.000	.844	.038
	Wilks' Lambda	.962	.273(a)	3.000	21.000	.844	.038
	Hotelling's Trace	.039	.273(a)	3.000	21.000	.844	.038
	Roy's Largest Root	.039	.273(a)	3.000	21.000	.844	.038
Group	Pillai's Trace	.511	7.317(a)	3.000	21.000	.002	.511
	Wilks' Lambda	.489	7.317(a)	3.000	21.000	.002	.511
	Hotelling's Trace	1.045	7.317(a)	3.000	21.000	.002	.511
	Roy's Largest Root	1.045	7.317(a)	3.000	21.000	.002	.511

Table III.7

Tests of Between-Subject Effects for Three Independent Variables on Overall Instructional Quality

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Unique strategy	6.938(a)	2	3.469	5.216	.014	.312
	Quality	9.986(b)	2	4.993	8.188	.002	.416
	Time	910.000 (c)	2	455.000	4.025	.032	.259
Intercept	Unique strategy	23.317	1	23.317	35.060	.000	.604
	Quality	110.313	1	110.313	180.912	.000	.887
	Time	7416.695	1	7416.695	65.617	.000	.740
Preknowledge score	Unique strategy	.576	1	.576	.867	.362	.036
	Quality	.014	1	.014	.023	.882	.001
	Time	13.670	1	13.670	.121	.731	.005
Group	Unique strategy	6.935	1	6.935	10.427	.004	.312
	Quality	9.238	1	9.238	15.150	.001	.397
	Time	877.070	1	877.070	7.760	.011	.252
Error	Unique strategy	15.296	23	.665			
	Quality	14.025	23	.610			
	Time	2599.684	23	113.030			
Total	Unique strategy	112.706	26				
	Quality	539.142	26				
	Time	36810.847	26				
Corrected Total	Unique strategy	22.234	25				
	Quality	24.010	25				
	Time	3509.684	25				

a R Squared = .312 (Adjusted R Squared = .252)

b R Squared = .416 (Adjusted R Squared = .365)

c R Squared = .259 (Adjusted R Squared = .195)

Table III.8
Means and Standard Deviations of the Three Independent Variables in the MANCOVA

Variable	Mean	Standard Deviation
Unique number of strategies		
PD	2.23	.84
TP	1.19	.75
Instructional Quality		
PD	4.90	.79
TP	3.60	.72
Average Time		
PD	40.06	9.25
TP	27.72	12.46

Note: PD = Professional Development, TP = Typical Practice

There was a positive impact on the teachers' use of vocabulary strategies, both in time and variety, and overall instructional quality. Differences emerged as a result of the activities documented in the typical practice and professional development classrooms. Typical practice classrooms spent little time reading social studies text and more time with teacher lecturing. This finding may be a result of the difficulties students encounter with content area text referred to earlier. In order to compensate for students' difficulties, teachers chose to read to students or simply explained the concepts orally.

Instances of vocabulary instruction were rare, corroborating the findings of other studies (Bailey et al., 2006; Durkin, 1978-1979; Scott et al., 2003). When vocabulary instruction was observed, it was primarily definitional in nature with the teacher providing the definition or students looking up words in a glossary. Little time was devoted to the instruction of important social studies content words, and even less time

was given to vocabulary practice or planned multiple exposures. These findings are counter to the most recent vocabulary instruction recommendations (National Literacy Institute, 2007).

Unique to this study was the comparison of teachers who had received professional development in vocabulary strategies when compared to typical practice teachers. Teachers who received professional development in incorporating vocabulary instruction in social studies content differed from the typical practice teachers on several levels. Not only did they devote more time to social studies instruction itself, but more vocabulary instruction was evident. Teachers also showed a greater repertoire of evidence-based vocabulary instructional strategies. Not only did they provide more vocabulary instruction, but they used strategies to meet students' needs. The documented instructional quality of these teachers was also significantly greater. One unexpected finding was the large percentage of time spent in logistical/instructional tasks. One explanation may be the time spent checking vocabulary activities. Another explanation may be the need for teachers to provide more directions for students because of the new activities teachers introduced.

Professional development situated within content instruction may be the bridge to link vocabulary instructional strategies with content instruction. It is evident that the advances in vocabulary research strategies in the past 30 years are not systematically transferring into classroom practice, as highlighted by the typical practice classroom results. However, when professional development on vocabulary strategies was situated

within the content teachers were required to teach, they implemented a wider variety of strategies more regularly and with higher quality.

The MANCOVA resulted in statistically significant results for quality, use, and variety, accounting for preknowledge of the strategies. Even when accounting for existing knowledge of evidence-based vocabulary strategies, teachers in the professional development group were rated higher in overall instructional quality, and used more vocabulary strategies more often. It is interesting to note that quality was the variable most impacted by professional development. The professional development provided may have provided teachers with routines and classroom strategies that aligned with other factors of overall instructional quality, such as general pacing, type of feedback, and student engagement.

There are several limitations of this study that may limit its generalizability. The findings of vocabulary instruction in the professional development group may be exaggerated due to the nature of the professional development and the provision of classroom materials. Regardless, the impact of training on teacher practice is still relevant. Any findings must be tempered with the knowledge of the small sample. Because the instructional activities were documented using audio tapes, information on student participation is unavailable, which limits findings regarding student engagement. Additionally, teacher actions that were nonverbal were unable to be recorded.

Conclusion

A primary purpose of this study was to document vocabulary instructional practices in fourth grade social studies classrooms and the impact of professional

development of evidence-based vocabulary strategies. It is evident in this study that professional development increased teachers' use of evidence-based vocabulary strategies in social studies classrooms. Findings of this study show some translation of the growing knowledge base of vocabulary instruction specifically in social studies. It is possible the differences found may be a result of the professional development program and therefore not indicative of wide-spread change. Certainly, this study opens the possibility of what can be done. Without past results with which to compare, these results can be viewed as exploratory and in need of replication.

The right balance between content coverage and vocabulary strategy instruction is currently unknown. Students perform better on content task when they have stronger knowledge of the content's vocabulary; therefore, one unique aspect of vocabulary instruction in the content areas to consider is the optimal time to devote to vocabulary instruction, which reaps the greatest benefits to content understanding without taking away from content instruction.

Four broad areas for future research are recommended: content area vocabulary strategies, professional development, documentation of classroom practice, and impact on student achievement. First, research is needed to determine which vocabulary strategies currently identified in the research are most appropriate and beneficial in content area instruction. Just as the nature of the words in content area text is different from narrative text, so should the instructional strategies vary. Second, models of research dissemination and professional development need to be developed and studied to identify methods to narrow the gap between what is understood to be effective and classroom

implementation. While much research has focused on identifying what works, little has focused on what's used. Observational studies are costly in terms of time and funding, yet knowledge of typical classroom practice may provide an avenue to the development of robust models of research dissemination and professional development. These studies provide insight into the multiple factors that shape classroom practice. The ultimate goal of vocabulary instruction is students' increased vocabulary knowledge and subsequent content area understanding. Therefore, studies linking the instruction of vocabulary strategies to students' gains in vocabulary and subject matter knowledge are important.

CHAPTER IV

ADDITIONAL ANALYSES

Durkin's (1978-1979) seminal observational study of comprehension instruction in fourth grade reading and social studies classrooms identified the lack of reading comprehension instruction not only in reading classrooms, but social studies classrooms as well. This set of studies published 30 years ago documented the marked absence of comprehension instruction during general and content-area instruction. In the first phase of the study, observations of 24 fourth grade classrooms revealed that only 0.63% of the total time in the reading period was dedicated to comprehension instruction, and 17.65% of the time was spent in comprehension assessment. Similar results were found in the observation of 24 teachers for 2,775 minutes during social studies instruction. Comprehension assessment, defined as questions to assess the content, was the most typical comprehension activity (8.25% of the total time) during social studies. After almost 3,000 minutes of observation, Durkin summarized that "practically no comprehension instruction was seen" and social studies was not "a time to improve children's comprehension abilities" (pp. 520-521).

Since Durkin, researchers have attempted to replicate and extend the understanding of instructional practices during reading and content area instruction.Kurth and Greenlaw (1980) studied 16 teachers in grades Kindergarten, second, fourth, and sixth grades, and found that 32% of reading instruction time was dedicated to comprehension, with the amount of time increasing at each grade level.

Wendler, Samuels, and Moore (1989) observed 36 third to sixth grade teachers during reading instruction based on Durkin's (1978-1979) observational categories. Comparing the comprehension instructional practices of award-winning teachers, teachers with master's degrees, and a control group, they found few differences between the three groups in regards to the time allocated to comprehension instruction, with approximately 8% of the observed time coded as any type of comprehension activity. They hypothesized that the cause may be the reliance on basal textbooks. The authors concluded, "…teachers apparently are allocating their time in much the same way they did when Durkin (1978-1979) conducted her study" (Wendler et al., 1989, p. 396). Both the Kurth et al. (1980) and Wendler et al. (1989) studies documented the minimal amounts of reading comprehension instruction during time designated for reading instruction.

Almost two decades have passed since the most recent of these studies. During this time, the knowledge base in comprehension instruction has expanded. A specific purpose of this study was to examine whether there is evidence that teachers are using the accruing knowledge base of comprehension instruction when they teach subject matter and use content area texts. Specifically, determining whether teachers' use of practices in comprehension and vocabulary have changed since Durkin's seminal observation study was of interest. Toward that end, this study compared findings from a study conducted in 2007 to findings from the Durkin study. The primary question to be answered is how do the findings in the present study of fourth grade social studies

teachers compare to Durkin's findings 30 years ago regarding the proportion of time dedicated to comprehension instruction.

To accomplish this comparison, the categories coded by Durkin (1978-79) had to correspond to the activity codes in the present study. The present study used a modified version of the ICE-R (Edmonds & Brigg, 2007), which was designed to document specific reading instructional practices. For specifics on the coding instrument, refer to Chapter III for modification. The coding instrument can be found in Appendix B. While there is currently not a one-to-one correspondence of the Durkin codes to the modified strategy documentation instrument, this section was completed to satisfy agreed upon outcomes from the proposal. Future analysis of the raw data will allow for this one-to-one correspondence and subsequent in-depth analysis.

Nine 4th grade social studies teachers who were participating in *Enhancing the Quality of Expository Text Instruction and Comprehension through Content and Case-Situated Professional Development funded by the Institute for Education Science* (Simmons et al., 2005) (IES grant contract number R305M050121A) comprised the study sample. This study is termed a strategy use study using audio documentation. Codings were obtained from listening to digitally audio-recorded lessons, not in person observations. Audio recordings of classroom instruction were collected at three time points in the 18-week vocabulary intervention. A total of 25 audio recordings comprising a total of 604 minutes were collected. For additional information on the demographics of the teachers refer to Chapter III. A cross reference of the codes used by Durkin (1978-79) and those of the present study are provided in Table IV. 1.

Table IV.1.
Cross Reference of Strategy and Classroom Activity Codes from Durkin (1978-1979) and Hairrell (2008)

Durkin Category-Social Studies	Modified Strategy Use Instrument Category
Assignment: checks (teacher checks answers connected with an assignment)	13a Logistical/ Instructional Tasks: Checking assignments
Assignment: gives (all assignments except comprehension and study skills)	13b Logistical/ Instructional Tasks: Giving assignments
Collects materials (collecting materials)	13d Logistical/ Instructional Tasks: Collection materials
Comprehension: review of instruction (review or repeat previous comprehension instruction)	9e/b Comprehension strategy instruction/ use
Comprehension: application (comprehension instruction enables children to understand connected text)	9e/c Comprehension: strategy instruction/ use
Comprehension: assessment (like assignment: checks but related to comprehension)	9c/d Comprehension: monitoring (reading and listening)
Comprehension: assignment (assignment that requires comprehension of connected text)	9e/c Comprehension: strategy instruction/ use
Comprehension: help with assignment (teacher helps with given assignment)	9e/b Comprehension: strategy instruction/use
Comprehension: instruction (helps children understand or work out the meaning of more than a single word)	9e/a Comprehension: strategy instruction/use
Comprehension: prediction (If a teacher says something like, "Now that you've read the first part of the story, what do you think is likely to happen in the next part?" the behavior goes here.)	9b Comprehension: prior knowledge/ predicting

Comparison to Durkin

One goal of this study was to determine if comprehension instruction in the context of social studies had changed since Durkin's observational study in the late 1970s. The ultimate goal of reading is comprehension. In social studies, the expectation is not just basic comprehension of text, but the expectation that students will gain content concepts from their independent reading. Vocabulary plays a significant role in comprehension, although it is difficult to disaggregate its effects in reading comprehension (NRP 2000).

Similarities were numerous between Durkin's findings and findings in the present study. Both studies were found to devote none of their social studies time to comprehension instruction, review, application, help, or prediction. Percentages of time were also similar for checking assignments (D = 3.4%, TP = 3.0%), non-instructional time (D = 7.7%, TP = 7.6%), and study skills review (D = .5%, TP = .1%). Figure IV.1 provides a graphic representation of the similarities and differences between the two studies.

Several differences in the two data sets were also evident. While helping with assignments accounted for 11.5% of Durkin's observations, it only accounted for .07% of the typical practice classroom practice. Although no time was observed for oral review in typical practice classrooms, this activity was coded at 11.2% of Durkin's observed time. Almost double the time was spent in typical classrooms listening (D = 11.0%, TP = 20.7%). A large discrepancy was evident in comprehension, preparation for reading (D = 1.7%, TP = 10.5%).

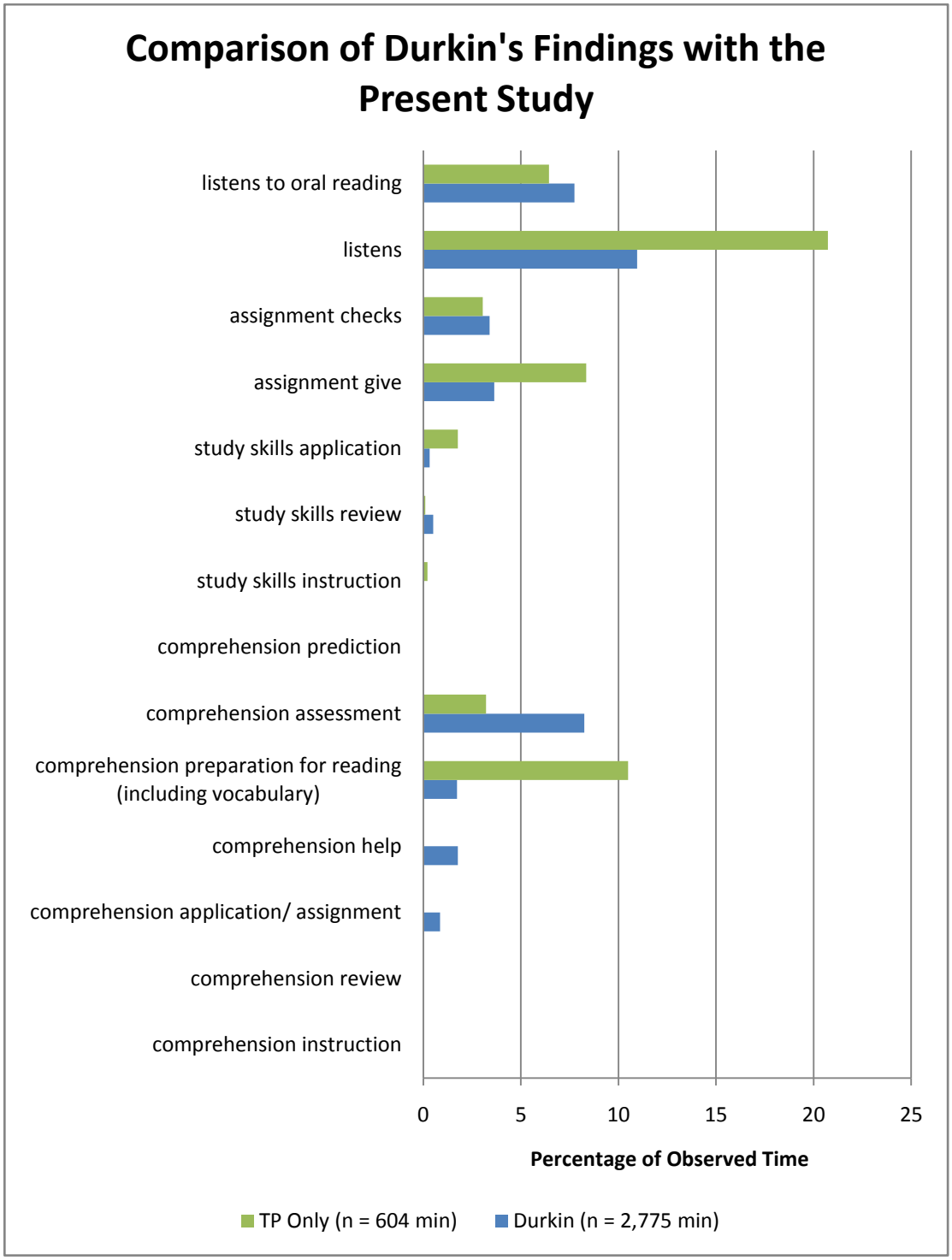

Figure IV.1.

Comparison of Durkin (1978-79) to Hairrell (2008)

The three most common activities reported by Durkin were help with assignment (11.5%), transition (11.2%), and listening (11.0%). Listening (20.7%), preparation for reading (10.5%), and giving assignments (8.4%) were the most observed in the typical classrooms only.

Conclusion

One limitation of this analysis is that while Durkin gathered her data through classroom data gathering in the present study was done by digital audio recording. As a result, the coding in the present study may have failed to capture instances of the teacher helping students that were non-verbal or out of range of the audio recorder. This could account for the large disparity in the category of helping with assignments observed between the groups.

Findings of this study show that little progress has been made in the past 30 years regarding comprehension and the inextricable component of vocabulary in social studies instruction. Unfortunately, as our research knowledge base grows, there is not widespread transfer and implementation of this knowledge into the classrooms observed in this study.

CHAPTER V

CONCLUSION

The purpose of this two-part dissertation was to (a) document the quantity and methodological characteristics of vocabulary research from 1998 to 2007 and (b) conduct an strategy use study of the implementation of evidence-based vocabulary strategies in classrooms where reading of informational text is a principal source of learning. This research was designed to examine the quantity and methodological characteristics of vocabulary instruction research and to document to what extent such practices are being implemented in classroom instruction. I have presented two stand alone journal articles: (a) a systematic literature review of vocabulary interventions since 1998, and (b) a strategy use study of the use of evidence-based strategies in fourth grade social studies classrooms. Chapter IV provides additional statistical analyses from the data collected as part of the strategy use study.

Based on the systematic literature review presented in Chapter II, vocabulary research is experiencing increasing interest, corresponding to the need for vocabulary to be an important research topic (Cassidy, Garrett, & Barrera, 2006). This study contributes to the vocabulary literature in two ways. First, it provides an update of the research in vocabulary instructional strategies in second through eighth grades since the publication of the NRP. Second, it describes the methodological characteristics of the body of vocabulary research in print according to the standards of research recently published by several researchers and research entities.

The findings of this review align with the findings of past meta-analyses and reviews; there are several vocabulary strategies effective at increasing vocabulary knowledge. Additionally, the variety of intervention strategies in the studies heightens our understanding of the multi-dimensional nature of vocabulary and the many factors that influence the acquisition and maintenance of vocabulary knowledge.

The methodological characteristics of studies published from 1999 to 2007 indicate that many vocabulary instructional studies use experimental and quasi-experimental designs with treatment and control groups and adequately describe the interventions used. Statistical analyses are generally limited to t-tests and ANOVAS, with effect sizes reported in the majority of published studies. There is a need for complete sample descriptions to allow for replication studies and the reporting of fidelity and fidelity quality. A strong dependence on researcher-developed measures may be indicative of the need for accurate measures of vocabulary knowledge. In sum, while the body of research is generally approaching the rigorous research standards in the areas of design and intervention characteristics, the areas of sample and statistical characteristics have room for improvement.

Based on my findings in Chapter II, I believe several implications for future study should be noted. Vocabulary instruction can increase vocabulary knowledge; however, the best program of instruction to improve students' vocabulary, including the type and dosage of vocabulary instruction, is still elusive. Several studies have also shown differential effects for different populations, especially those with varying initial vocabulary levels. Currently, most studies are conducted with small to medium sample

sizes, lacking the robustness needed to have respectable external validity. Furthermore, studies designed to determine the most appropriate amount of time to devote to vocabulary instruction, whether daily or weekly, are needed.

A component of vocabulary that is often ignored is the retention of word knowledge over time. Eight studies utilized delayed post-testing, thus our knowledge of which vocabulary instructional strategies support long-term learning of words is mostly unknown. Therefore; measures must be sensitive to the breadth, as well as depth of word knowledge, while also determining the difference in short-term learning and long-term maintenance.

The next generation of vocabulary research must commit to advancing the field through the use of best practices regarding research design, analysis, and reporting. Improvement in the quality of vocabulary research necessitates that researchers be conscientious in their reporting and provide adequate descriptions of the participants, interventionists, and intervention. Establishment of a line of testable findings cannot be achieved through replication inquiries unless all characteristic features of studies are reported. Researchers must strive to eliminate a primary source of confusion when comparing studies, which is the lack of a clear set of terminology for instructional practices and research methodological characteristics. The multi-dimensional nature of vocabulary calls for the use of high levels of statistical analysis and the use of statistical models to determine which factors contribute most to vocabulary learning. Through the use of hierarchical linear modeling and structural equation modeling, the many factors that contribute to vocabulary learning may be discovered.

The study presented in Chapter III, contributes to the literature in two primary ways. First, the documentation of strategy use provides a portrait of the amount and type of vocabulary instruction used in fourth grade social studies classrooms. Second, this study documented the impact of professional development on teachers' use of vocabulary instruction in social studies.

Findings reveal that vocabulary instruction in typical practice classrooms was negligible. Time that was spent in vocabulary instruction was generally definitional in nature. In comparison, teachers in the professional development group showed not only more time devoted to vocabulary instruction, but a larger repertoire of strategies.

Based on the findings in Chapter III, implications for vocabulary instruction and research are clear. First, the research-to-practice gap is evident. Even with the converging knowledge base of effective vocabulary instruction, very little use of those strategies is evident in typical practice classrooms. Second, this study documented the positive impact of professional development situated within the content and context of the classroom on teachers' use of evidence-based strategies. When teachers were presented with vocabulary strategies embedded within the social studies content, they were able to use these strategies during content instruction.

Based on the findings in Chapter III, I recommend three broad areas for future research(a) the application of vocabulary strategies in content areas, (b) how to provide effective professional development and (c) the impact on student achievement. First, research is needed to determine which vocabulary strategies currently identified in the research are most appropriate and beneficial in content-area instruction. Just as the

nature of the words in content area text is different from narrative text, so should the instructional strategies vary. Second, models of research dissemination and professional development need to be developed and studied to identify methods to narrow the gap between what we understand to be effective and classroom implementation. While much research has focused on identifying what works, little has focused on what's used. Observational studies are costly in terms of time and funding, yet knowledge of typical classroom practice may provide an avenue to the development of robust models of research dissemination and professional development. These studies provide insight into the multiple factors that shape classroom practice. Third, the ultimate goal of vocabulary instruction is students' increased vocabulary knowledge and subsequent content area understanding. Therefore, studies linking the instruction of vocabulary strategies to students' gains in vocabulary and subject matter knowledge are important.

Chapter IV provided a comparison of the findings of the strategy use study with the findings of Durkin's (1978-1979) seminal observational study. Findings revealed few changes over the past 30 years in regards to comprehension instruction. Even with the exponential growth in the knowledge of effective comprehension instruction during this time period, changes in practice are minimal. Similar to the recommendations from Chapter III, these findings reveal the necessity to identify strategies to disseminate research findings and narrow the research-to-practice gap.

As a result of this two-part dissertation, I have examined the efficacy of vocabulary research conducted since 2002 and documented the use of evidenced-based vocabulary strategies in fourth grade social studies classrooms through a documentation

study. In a journal article format, I have presented two stand alone scholarly journal articles and one supplementary chapter containing additional statistical analyses. Chapter II presented the systematic literature review of vocabulary interventions since 1998, Chapter III presented the results of a strategy use study of the use of evidence-based strategies in fourth grade social studies classrooms, and Chapter IV presented additional analysis comparing the findings of the strategy use study to a previous study (Durkin, 1978-1979).

REFERENCES

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*, 364-373.

American Education Research Association Task Force on Reporting of Research Methods in AERA Publications. (2006). Standards for reporting on empirical social science research in AERA publications. *Educational Researcher, 35*(6), 33-40.

American Psychological Association. (2002). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.

Apthorp, H. S. (2006). Effects of a supplemental vocabulary program in third-grade reading/language arts. *The Journal of Educational Research, 100*, 67-79.

Armbruster, B. B., & Anderson, T. H. (1988). On selecting "considerate" content area textbooks. *Remedial and Special Education, 9*, 47-52.

Armbruster, B. B., & Gudbrandsen, B. (1986). Reading comprehension in social studies programs. *Reading Research Quarterly, 21*, 36-48.

Bailey, G., Shaw, E. L., Jr., & Hollifield, D. (2006). The devaluation of social studies in the elementary grades. *Journal of Social Studies Research, 30*(2), 18-29.

Baker, S. K., Simmons, D. C., & Kame'enui, E. J. (1998). Vocabulary acquisition: Synthesis of the research, Tech. Rep. No. 13. Retrieved September 3, 2005 from http://idea.uoregon.edu/~ncite/documents/techrep/tech13.html

Baumann, J. F., Edwards, E. C., Boland, E. M., Olejnik, S., & Kame'enui, E. J. (2003a). Vocabulary tricks: Effects of instruction in morphology and context on fifth-grade students' ability to derive and infer word meanings. *American Educational Research Journal, 40*, 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*, 150-176.

Baumann, J. F., Font, G., Edwards, E., & Boland, E. M. (2005). Strategies for teaching middle-grade students to use word-part and context clues to expand reading vocabulary. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary* (pp. 179-205). Mahwah, NJ: Erlbaum.

Baumann, J. F., & Kame'enui, E. J. (2003). *Vocabulary instruction: Research to practice*. New York: The Guilford Press.

Baumann, J. F., Kame'enui, E. J., & Ash, G. E. (2003b). Research on vocabulary instruction: Voltaire redux. In J. Flood, D. Lapp, J. R. Squire & J. M. Jensen (Eds.), *Handbook of research on teaching the English language arts* (pp. 752-785). Mahwah, NJ: Erlbaum.

Baxter, S., & Reddy, L. (2007). *What content-area teachers should know about adolescent literacy*. Jessup, MD: National Institute for Literacy.

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

Best, R. M., Dockrell, J. E., & Braisby, N. (2006). Lexical acquisition in elementary science classes. *Journal of Educational Psychology, 98*, 824-838.

Biancarosa, C., & Snow, C. E. (2006). *Reading next: A vision for action and research in middle and high school literacy: A report to the Carnegie Corporation of New York* (2nd ed.). Washington, DC: Alliance for Excellent Education.

Biemiller, A. (2003a). Oral comprehension sets the ceiling on reading comprehension. *American Educator, 27*, 23, 44.

Biemiller, A. (2003b). Vocabulary: Needed if more children are to read well. *Reading Psychology, 24*, 323-335.

Blachowicz, C., & Fisher, P. (2000). Vocabulary instruction. In M. L. Kamil, P. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. III, pp. 503-523). Mahwah, NJ: Erlbaum.

Bos, C. S., & Anders, P. L. (1990). Effects of interactive vocabulary instruction on the vocabulary learning and reading comprehension of junior-high learning disabled students. *Learning Disability Quarterly, 13*, 31-42.

Boulware-Gooden, R., Carreker, S., Thornhill, A., & Joshi, R. M. (2007). Instruction of metacognitive strategies enhances reading comprehension and vocabulary achievement of third-grade students. *The Reading Teacher, 61*, 70-77.

Bryant, D. P., Goodwin, M., Bryant, B. R., & Higgins, K. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disability Quarterly, 26*, 117-128.

Cain, K. (2007). Deriving word meanings from context: Does explanation facilitate contextual analysis? *Journal of Research in Reading, 30*(2), 1-13.

Cain, K., Oakhill, J., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language, 30*, 681-694.

Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*, 671-681.

Carlisle, J. F., Fleming, J. E., & Gudbrandsen, B. (2000). Incidental word learning in science classes. *Contemporary Educational Psychology, 25*, 184-211.

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

Carnine, D. (1997). Bridging the research-to-practice gap. *Exceptional Children, 63*, 513-521.

Carnine, D., Crawford, D. B., Harniss, M. K., Hollenbeck, K. L., & Miller, S. K. (2002). Effective strategies for teaching social studies. In E. J. Kame'enui, D. W. Carnine, R. C. Dixon, D. C. Simmons & M. D. Coyne (Eds.), *Effective teaching*

strategies that accommodate diverse learners (Vol. 2). Upper Saddle River, NJ: Pearson Education, Inc.

Cassidy, J., Garrett, S. D., & Barrera, E. S. (2006). What's hot in adolescent literacy 1997-2006. *Journal of Adolescent & Adult Literacy, 50*(1), 30-36.

Chall, J. S. (1996). *Stages of reading development.* 2nd ed. United States: Harcourt Brace College.

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

Cooper, H., & Hedges, L. V. (Eds.). (1994). *Handbook of research synthesis*. New York: Russell Sage Foundation.

Cromley, J. G., & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology, 99*(2), 311-325.

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

Curtis, M. E., & Longo, A. M. (2001). Teaching vocabulary to adolescents to improve comprehension [Electronic Version]. *Reading Online*, *5*, from http://www.readingonline.org/articles/art_index.asp?HREF=curtis/index.html

de Leon, A. G. (2002). Moving beyond storybooks: Teaching our children to read to learn [Electronic Version]. *Carnegie Reporter, 2*. Retrieved February 1, 2007, from www.carnegie.org/reporter/05/learning/index.html

Duke, N. K. (2000). 3.6 minutes per day: The scarcity of informational texts in first grade. *Reading Research Quarterly, 35*, 202-224.

Durkin, D. (1978-1979). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*, 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.

Edmonds, M., & Briggs, K. L. (2003). Instructional content emphasis instrument. In S. R. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for observing teaching and learning*. Baltimore: Brookes.

Edwards, E., Font, G., Baumann, J. F., & Boland, E. (2003). Vocabulary instruction for students with learning disabilities: A review of the research. *Learning Disabilities Quarterly, 26*, 117-128.

Edwards, E., Font, G., Baumann, J. F., & Boland, E. (2004). Unlocking word meanings: Strategies and guidelines for teaching morphemic and contextual analysis. In J. F. Baumann & E. J. Kame'enui (Eds.), *Vocabulary instruction: Research to practice*. New York: Guildford Press.

Fore III, C., Boon, R. T., & Lowrie, K. (2007). Vocabulary instruction for middle school students with learning disabilities: A comparison of two instructional models. *Learning Disabilities — A Contemporary Journal, 5*(2), 49-63.

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

Gardner, D. (2004). Vocabulary input through extensive reading: A comparison of words found in children's narrative and expository reading materials. *Applied Linguistics, 25*(1), 1-37.

Gardner, D. (2007). Children's immediate understanding of vocabulary: Contexts and dictionary definitions. *Reading Psychology, 28,* 331-373.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M. D., Greenwood, C. R., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71,* 149-164.

Graves, M. F. (2006). Building a comprehensive vocabulary program. *The New England Reading Association Journal, 42*(2), 1-7.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education, 24*(4), 276-289.

Gresham, E. M., MacMillan, D. L., Beebe-Frankenberger, M. E., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice, 15,* 198-205.

Harmon, J. M., Hedrick, W. B., & Wood, K. D. (2005). Research on vocabulary instruction in the content areas: Implications for struggling readers. *Reading & Writing Quarterly, 21,* 261-280.

Hart, B., & Risley, T. R. (2003). The early catastrophe: The 30 million word gap by age 3 [Electronic Version]. *American Educator, 27*. Retrieved March 3, 2006, from http://www.aft.org/pubs-reports/american_educator/spring2003/catastrophe.html.

Heimlich, J. E., & Pittelman, S. D. (1986). *Semantic mapping: Classroom applications*. Newark, DE: International Reading Association.

Heller, R., & Greenleaf, C. (2007). *Literacy instruction in the content areas: Getting to the core of middle and high school improvement.* Washington, DC: Alliance for Excellent Education.

Hirsch, E. D. (2003). Reading comprehension requires knowledge--of words and the world: Scientific insights into the fourth-grade slump and the nation's stagnant comprehension scores. *American Educator, 27*, 10-13, 16-22, 28-29, 48.

Hsieh, P., Acee, T., Chung, W., Hsieh, Y., Kim, H., Thomas, G. D., You, J., Levin, J. R., & Robinson, D. H. (2005). Is educational intervention research on the decline? *Journal of Educational Psychology, 97*, 523-529.

Institute of Education Sciences. (2006). Procedures for peer review of grant applications [Electronic Version]. Retrieved July 18, 2007 from http://www.ed.gov/about/pubs/intro/index.html?src=gu

Jitendra, A. K., Edwards, L. L., & Sacks, G. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*, 299-322.

Jones, M. S., Levin, M. E., Levin, J. R., & Beitzel, B. D. (2000). Can vocabulary-learning strategies and pair-learning formats be profitably combined? *Journal of Educational Psychology, 92*, 256-262.

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

Kauffman, J. M. (1996). Research to practice issues. *Behavioral Disorders, 22*, 55-60.

Kennedy, M. M. (1997). The connection between research and practice. *Educational Researcher, 26*, 4-12.

Kuhn, M. R., & Stahl, S. A. (1998). Teaching children to learn word meanings from contexts: A synthesis and some questions. *Journal of Literacy Research, 30*, 119-138.

Kurth, R. J., & Greenlaw, M. J. (1980). *Research and practices in comprehension instruction in elementary classrooms*. Paper presented at the American Reading Conference, Sarasota, FL.

Lee, J., Grigg, W., & Donahue, P. (2007). *The Nation's Report Card: Reading 2007 (NCES 2007-496)*. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.

Lubliner, 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*, 163-200.

Martin-Chang, S. L., Levy, B. A., & 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.

Marzano, R. J., & Pickering, D. J. (2005). *Building academic vocabulary: Teacher's manual*. Alexandria, VA: Association for Supervision and Curriculum Development.

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

Mezynski, K. (1983). Issues concerning the acquisition of knowledge: Effects of vocabulary training on reading comprehension. *Review of Educational Research, 53*, 253-279.

Moats, L. C. (1999). *Teaching reading is rocket science: What expert teachers of reading should know and be able to do*. Washington, D. C.: American Federation of Teachers.

Moats, L. C. (2004). Efficacy of a structured, systematic language curriculum for adolescent poor readers. *Reading & Writing Quarterly, 20*, 145.

Moss, B., & Newton, E. (2002). An examination of the informational text genre in basal readers. *Reading Psychology, 23*, 1-13.

Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second readers and at-risk fourth-grade writers. *Journal of Educational Psychology, 95*, 730-742.

Nagy, W., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D. Pearson & R. Barr (Eds.), *Handbook of reading research* (Vol. 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 & Communication Disorders, 41*, 335-354.

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

Nelson, J. R., & Stage, S. A. (2007). Fostering the development of vocabulary knowledge and reading comprehension through contextually-based multiple meaning vocabulary instruction. *Education and Treatment of Children, 30*(1), 1-22.

No Child Left Behind Act of 2001, Congress of the United States of America, 107th Sess. 1425-2094.

Pittelman, S. D., Heimlich, J. E., Berglund, R. L., & French, M. P. (1991). *Semantic feature analysis: Classroom applications*. Newark, DE: International Reading Association.

Pressley, M., Disney, L., & Anderson, K. (2007). Landmark vocabulary instructional research and the vocabulary instructional research that makes sense now. In R. K. Wagner, A. E. Muse & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 205-232). New York: Guilford Press.

Pressley, M., & Harris, K. R. (1994). Increasing the quality of educational intervention research. *Educational Psychology Review*, 6, 191-208.

Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 107-128*)*. Hillsdale, NJ: Erlbaum.

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics, 24*, 146-161.

Robinson, V. M. J. (1998). Methodology and the research-practice gap. *Educational Researcher, 27*(1), 17-26.

Ross, A. H., & Stevens, K. B. (2003). Teaching spelling of social studies content vocabulary prior to using the vocabulary in inclusive learning environments: An examination of constant time delay, observational learning, and instructive feedback. *Journal of Behavioral Education, 12*, 287-309.

Scammacca, N., Roberts, G., Vaughn, S., Edmonds, M., Wexler, J., Reutebuch, C. K., et al. (2007). *Interventions for adolescent struggling readers: A meta-analysis with implications for practice*. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Schatschneider, C., Buck, J., Torgesen, J. K., Wagner, R. K., Hassler, L., Hecht, S., et al. (2004). *A multivariate study of factors that contribute to individual differences in performance on the Florida Comprehensive Reading Assessment Test. (Technical Report No. 5)*. Tallahassee, FL: Florida Center for Reading Research.

Scott, J. A., Jamieson-Noel, D., & Asselin, M. (2003). Vocabulary instruction throughout the day in twenty-three Canadian upper-elementary classrooms. *The Elementary School Journal, 103*, 269-286.

Shavelson, R. J., & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Simmons, D. C., Rupley, W. H., & Vaughn, S. (2005). Enhancing the Quality of Expository Text Instruction and Comprehension through Content and Case-Situated Professional Development: Institute of Education Sciences grant contract number R305M050121A.

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Society for Prevention Research. (2004). Standards of evidence: Criteria for efficacy, effectiveness and dissemination [Electronic Version]. Retrieved July 29, 2007, from http://www.preventionresearch.org

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

Stevens, R. J. (2003). Student team reading and writing: A cooperative learning approach to middle school literacy instruction. *Educational Research & Evaluation, 9*, 137-160.

Swanborn, M. S. L., & de Glopper, K. (1999). Incidental word learning while reading: A meta-analysis. *Review of Educational Research, 69*, 261-285.

Swanborn, M. S. L., & de Glopper, K. (2002). Impact of reading purpose on incidental word learning from context. *Language Learning, 52*(1), 95-117.

Twyman, T., McCleery, J., & Tindal, G. (2006). Using concepts to frame history content. *Journal of Experimental Education, 74*, 331-349.

U.S. Department of Education, Institute of Education Sciences, & National Center for Evaluation and Regional Assistance. (2003). *Identifying and implementing educational practices supported by rigorous evidence: A user friendly guide*. Washington, D.C.: Coalition for Evidence-Based Policy.

Uberti, H. Z., Scruggs, T. E., & Mastropieri, M. A. (2003). Keywords make the difference! *Teaching Exceptional Children, 35*, 56-61.

Veerkamp, M. B., Kamps, D. M., & Cooper, L. (2007). The effects of class-wide peer tutoring on the reading achievement of urban middle school students. *Education and Treatment of Children, 30*(2), 21-51.

Wendler, D., Samuels, S. J., & Moore, V. K. (1989). Comprehension instruction of award-winning teachers, teachers with master's degrees, and other teachers. *Reading Research Quarterly, 24*, 382-401.

What Works Clearinghouse. (2006). What Works Clearinghouse evidence standards for reviewing studies [Electronic Version]. Retrieved July 15, 2007, from http://www.whatworks.ed.gov/reviewprocess/study_standards_final.pdf

White, T., Sowell, J., & Yanagihara, A. (1989). Teaching elementary students to use word-part clues. *The Reading Teacher, 42*, 302-308.

Xin, J. F., & Rieth, H. (2001). Video-assisted vocabulary instruction for elementary school students with learning disabilities. *Information Technology in Childhood Education Annual*, 87-103.

APPENDIX A

Vocabulary Instruction: The State of the Research

Hairrell Dissertation
Texas A&M University
Abstraction Sheet

Study ID # _____

Date of Abstraction: _____

Authors:

Year of Publication:

Title:

Journal:	Volume:	Issue:	Pages:

ISI Journal Impact Factor: ❍ Not reported ❍ Low (<.5) ❍ Medium (.5-.99) ❍ High (>1.0)

Research Question(s):

Dependent Variable(s):

Theoretical Framework/Empirical Justification of the study:

❍ No Theoretical Framework/Empirical Justification Presented

Research Paradigm:

❍ Quantitative ❍ Mixed

Design:

❍ Correlational ❍ Quasi-experimental ❍ Experimental ❍ Longitudinal

Study includes a control group: ❍ Yes ❍ No

Study includes a comparison group: ❍ Yes ❍ No

Method for assigning participants to the study conditions is through: ❍ random assignment ❍ stratified random assignment

Sample Characteristics

N (total) =	Ethnicity(ies):	Ages/Grade Level:
❍ Not Reported ❍ < 100 = Small ❍ >100 and < 300 = Medium ❍ > 300 = Large		

Sample Design:	Method of Randomization:	Comparable Across Conditions or incomparability addressed and reflected in the analysis:	Attrition:
❍ Non-random ❍ Random	❍ Reported ❍ Unreported	❍ Yes ❍ No	❍ Reported ❍ Unreported ❍ Less than 30% or explained

Was the sample described so completely that you can identify the population?

❍ Yes
❍ No

Intervention Characteristics

Description of interventionist/teachers:	Description of Intervention:
❍ Yes ❍ No	❍ Intervention and comparison described ❍ Intervention only described ❍ Comparison only described ❍ Neither described

Length of intervention:	Instructional Strategy(ies) incorporated:	
	❍ Context ❍ Morphology ❍ Semantics ❍ Graphic Organizers ❍ Multiple/combined strategies ❍ Other:	❍ Read aloud ❍ Incidental word learning ❍ Visual/Illustrations ❍ Repeat Practice/ Multiple exposures

Methods for Data Collection

Design:	Data Collection Instrument:
❍ Group administered assessment ❍ Individually administered assessment ❍ Interview ❍ Other: _____	❍ Standardized measures only: _____ ❍ Combination of standardized and unstandardized measures: _____ ❍ Unstandardized measures only: _____ ❍ Other: _____

Measure(s) appropriate for construct (valid):	Measured beyond immediate post-test:
❍ Yes ❍ No	❍ Yes ❍ No

Data Analysis (statistical techniques employed)		
❍ Univariate Statistics / Descriptive (frequencies and mean distributions; percentages) ❍ Bivariate Statistics/ ANOVA - Correlation or Crosstabulation (Chi-Square) ❍ Multiple / Logistic Regression ❍ Multivariate Statistics – Canonical correlation analysis; discriminant function analysis; path analysis, structural equation modeling		
Data Reported		
Fidelity reported: ❍ Yes ❍ No ❍ Quality	Effect Size(s) reported: ❍ Yes ❍ No How computed?_____	Level of analysis matches level of assignment: ❍ Yes ❍ No
Validity and Reliability		
Data were reported on reliability of Vocabulary scores: ❍ Yes ❍ No	Data were reported on validity and/or reliability of DV scores: ❍ Yes ❍ No	
Test-retest reliability reported: ❍ Yes ❍ No	Inter-rater reliability reported: ❍ Yes ❍ No	

Findings			
Test / Finding / Relationship	DV	Vocabulary Instructional Strategy	Direction of Finding (-1, 0 or +1)

Findings, cont.			

Does conclusion *appropriately* imply causality: ❍ Yes ❍ No

Quote(s):

Reporting of the results are clear and coherent: ❍ Yes❍ No

Comments / Notes / Potential Quotes:

APPENDIX B

Modified Instructional Content Emphasis Strategy Documentation Instrument

Recording Forms

Observation Recording Form

Recording Code: ____________________________ Total Time: ____________________
Coder: __________ Date of Coding: _________ Date of Entry: _________________

Time (begin/end)	Summary of Activity	Grouping	Materials

Observation Recording Form Cont. Page ____ of ____

Time (begin/end)	**Summary of Activity**	**Grouping**	**Materials**

Describe the content of the observation and each instructional activity using the Code Book.

Dimension A: Content/General category
Dimension B: Content/General subcategory
Dimension C: Focus of instruction
Dimension D: Materials

		Dimension						
Time	**Brief summary of activity**	**A**	**B**	**C**	**D**	**Content Emphasis**	**Quality Indicator**	**Text Reading**

Codebook

Content Codes: Short Form

Dim. A	6. Oral Language Development	8. Text Reading	9. Comprehension	11. Vocabulary*
Dimension B	a. Teacher initiated structured opportunities to talk with teachers, peers. b. Expansion of student initiated language (incidental language strategies) c. Other	a. Supported oral reading b. Choral reading c. Independent silent reading d. Independent oral reading e. Teacher reads aloud f. Teacher reads aloud while students read along g. Paired reading* h. Other	a. Vocabulary* *Coded in new category* b. Prior knowledge/ predicting c. Reading comprehension monitoring d. Listening comprehension monitoring e. Comprehension strategy instruction/use f. other	a. Definitional b. Explicit instruction of words 1. teacher directed 2. student directed c. Contextual analysis d. Structural/ Morphological e. Organizational/Semantic strategies f. Practice Activities g. Keyword/Mnemonics h. Multiple strategies (code strategies from above) i. Other

Dim. A	12. Study Skills*	13. Logistical/ Instructional Tasks*	14. Logistical/ Noninstructional Tasks*	
Dimension B	Examples of study skills are skimming, paraphrasing, use of guide words, outlining, SQ3R, varying reading rate	a. Checking assignments b. Giving assignments c. Helping with assignments d. Collecting materials	a. Transition b. Student redirection/ correction c. Correcting papers at desk d. Other	

Dimension C: Focus of instruction	Dimension D: Materials	Emphasis Coding Scale	Quality Indicators
a. Instruction b. Review c. Application d. Assessment e. Social Studies Content	a. Puzzles & games b. Manipulatives c. Word wall d. Text-basal e. Text-decodable f. Text-pattern g. Text-unknown h. Text-St or T made i. Big book or similar j. Pencil and paper k. Words out of context l. Computers m. Audio tapes n. Workbooks/ worksheets o. Oral language p. Chalkboard or equivalent q. Other r. Visuals-with print s. Visuals-without print	5. Maximum emphasis (91-100%) 4. High emphasis (71-90%) 3. High moderate emphasis (41-70%) 2. Low moderate emphasis (11-40%) 1. Minimal emphasis (10% or less)	4. Excellent 3. High Average 2. Low Average 1. Weak

Descriptors of Codes: Dimensions A & B

Dimension A	Dimension B	Descriptors
1. Oral Language Development		**Definition: Focus is on listening and speaking to communicate meaning.** • Discussion is academic and not logistic or disciplinary • Teacher and students engage in discussion about words or relevant topics
	a. Teacher initiated structured opportunities to talk with teachers, peers	• Teacher expands students' responses by elaboration on key concepts *Note: purpose is for students to talk, not just the teacher; this does not include students incidentally talking while working on another assigned activity*
	b. Expansion of student initiated language (incidental language strategies)	• Teacher expands student responses by elaborating on key concepts • Activities that develop students' understanding of words or concepts • Activities that extend students' understanding of the world by making connections between concepts and experiences, and not necessarily learning a specific

		• definition for a word • Teacher uses scaffolding to extend students' language • Teacher uses incidental language intervention strategies such as following students' leads to model expanded language, elaborate with vocabulary *Note: If discussion is focused on vocabulary, code below in the vocabulary section.*
	c. Social Studies content focused	• Focus of discussion was social studies content
	d. Other	• Incidental class discussions without another academic purpose (teacher does not expand on student language)
8. Text Reading	a. Supported oral reading	• Students engage in reading either with the class, small group, or one-on-one • Guidance provided by a teacher, peer, or parent • Shared reading during which the teacher and students share the reading task • Students may need the teacher's help to read aloud (teacher prompts) • Students are guided to use semantic (does it make sense) and syntactic (does it sound right) clues to read
	b. Choral reading	• class or group reads aloud as a group simultaneously
	c. Independent silent reading	• Student reads text independently • Students read silently on their own
	d. Independent oral reading	• Student reads text independently • Student reads text orally on their own (absence of partner)
	e. Teacher reads aloud	• Students listen to books read aloud by the teacher with minimal emphasis on instruction • Students do not have a copy of text
	f. Teacher reads aloud while students read along	• Students have a copy of text to read while the teacher is reading • Students are following along as the teacher reads *Note: also includes students reading their own text while listening to recorded text reading*
	g. Paired reading*	• Students read orally in pairs or groups of three
	h. Other	• Students listen to books read aloud on computer or tape with minimal emphasis on instruction • Singing or chanting a know pattern or song with text
9. Comprehension	a. Vocabulary* *Coded in new category*	
	b. Prior knowledge/ predicting	• students preview the material before reading • students predict outcomes based on prior knowledge • students participate in activities designed to measure their level of knowledge before reading

	c. Reading comprehension monitoring	**Definition: Monitoring may occur during or after reading. Students learn to be aware of their understanding of text. Tends to be discussion-oriented with little focus on a product or goal. Teacher and students summarize the story as the intent of discussion or activities.** *Note: students must have their own copies of text in order to code the even as reading comprehension. If not, see listening comprehension below.* • During or after reading, students answer questions generated by teacher or student • Teacher and students discuss or respond to reading • Students discuss elements not explicitly found in the text • Students retell a story (verbally or through acting out events) • Students summarize a story's main events • Students identify the main idea • Students put story events into a sequence (including picture sequencing)
	d. Listening comprehension monitoring	**Definition: Monitoring of comprehension occurs during or after reading done by the teacher or other students.** *Note: All indicators under "reading comprehension monitoring" apply with listening comprehension if the focus is comprehension of text read aloud by someone else when students do not have text copy.* • Students are listing to reading done by teacher or students and the focus is on listening comprehension • Student responses may be oral or written but are based on reading performed by others *Note: If all students have copy of text, the code as reading comprehension.*
	e. Comprehension strategy instruction/use	**Definition: Teacher and students analyze text with a specific goal in mind (e.g., character analysis). Teacher provides direct instruction in the strategy. Students may have a product or shared understanding once activity is completed.** *Note: students do not necessarily have to have their own copies of text.* • Students are taught specific comprehension strategies • Students practice using comprehension strategies • Students use graphic or semantic organizers to make representations of material and assist in comprehension

	f.	• Students learn to use story structure to facilitate comprehension and recall • Students categorize text • Students involved in the identification and understanding of story elements • Students instructed in text features • Students are taught to integrate ideas and make generalizations from text
	g. other	• other instruction involving getting meaning from the text
11. Vocabulary*		Definition: Students have the opportunity to develop print vocabulary in the context of reading, discussion, practice, or explicit instruction
	a. Definitional	• Definition of the word is provided by teacher, student or a reference aid (e.g., dictionary, glossary, thesaurus, word bank). The definition is not expanded through discussion.
	b. Explicit instruction of words	• Explicit instruction in word meaning is provided by the teacher of student. Definition of word may be provided, but the definition is expanded on through the use of examples, discussion, etc. *Note: Indicate if instruction was teacher or student directed.*
	c. Contextual analysis	• Instruction is provided in using context to determine word meanings • Students use context strategies to determine word meanings • May be done at the sentence, passage or picture level
	d. Structural/ Morphological	• Instruction is provided in using morphological analysis or structural analysis to determine word meanings • Students use morphological analysis or structural analysis to determine word meanings
	e. Organizational/Semantic strategy	• Instruction is provided in using organization or semantic strategies to determine word meanings • Students use morphological analysis or structural analysis to determine word meanings Examples of strategies for this category ○ semantic mapping ○ semantic feature analysis ○ categorization/ classification ○ analogies ○ examples/nonexamples ○ synonyms/antonyms ○ homophones
	f. Practice Activities	• Activities that provide multiple exposures to vocabulary Examples of strategies for this category ○ games ○ worksheets ○ word wall activities

	g. Keyword/Mnemonics	• Activities that use keywords or mnemonics to learn word meaning
	h. Multiple strategies (code strategies from above)	• More than one of the above listed strategies are used in combination to learn or determine word meaning *Note: Indicate all strategies employed using codes above.*
	i. Other	• Other instruction used to learn word meanings
12. Study Skills		**Definition: Instruction or practice of skills designed to increase student ability to work in the classroom (e.g., skimming, paraphrasing, use of guide words, outlining, SQ3R, varying reading rate).**
13. Logistical/ Instructional Tasks*	a. Checking assignments	• Teacher and students check assignment together
	b. Giving assignments	• Teacher explains assignment to entire class
	c. Helping with assignments	• Teacher provides assistance to one or more students with an assignment while others are working
	d. Collecting materials	• Time spent collecting materials
14. Logistical/ Noninstructional Tasks*	a. Transition	• Time spent moving from one area to another • Time spent moving from one activity to another Examples of strategies for this category ○ finding materials ○ taking out books ○ distributing materials ○ moving to another area
	b. Student redirection/ correction	• Teacher corrects or redirects student behavior
	c. Correcting papers at desk	• Teacher corrects papers at desk while students work on an assignment
	d. Other	• Other noninstructional activities

Dimension C: Focus of Instruction

a. Instruction	• Teacher provides instruction on a skill
b. Review	• Teacher repeats or reviews instruction that was previously taught
c. Application	• Students apply a skill that was previously taught
d. Assessment	• Teacher assessment of student learning (oral or written)

Quality Indicators and Descriptions

4 Excellent	3 High Average	2 Low Average	1 Weak
Uses language that is direct and explicit.	Inconsistently uses language that is direct and explicit.		Uses language that is indirect and implicit.
Models many examples.	Provides some examples.		Provides no models or demonstrations.
Provides sufficient and varied opportunities for practice.	Provides many opportunities for practice with little variation. Practice opportunities do not seem to be based on student need.		Provides insufficient opportunities for practice with no variation.
Provides immediate and corrective and descriptive feedback.	Provides inconsistent feedback.		Provides little feedback that is nonspecific or no feedback.
Adjusts time to meet student needs.	Uses time appropriately, but use does not seem based on student need, yet still seems adequate for given activity.		Demonstrates poor use of time that is not differentiated and unrelated to student need or task difficulty.
Constantly monitors student performance.	Monitors some students or monitors all students for some activities.		Demonstrates lack of monitoring or monitoring very few students.
Encourages high student engagement and time on task.	Encourages student engagement and time on task varies.		Does not encourage student engagement and time on task.
Scaffolds tasks and materials to meet student needs.	Uses scaffolding inconsistently and does not always tailor it to student needs.		Scaffolds inappropriately or insufficiently.
Uses appropriate pacing, including wait time.	Uses inconsistent pacing that varies between appropriate at times “too fast” or “too slow” and provides insufficient wait time.		Demonstrates poor pacing, either too slow to too fast with no wait time provided.

Notes: Teachers must meet most of the observable indicators to be coded in a particular category. For example, if a teacher is rated as excellent in three categories, and high average in one, the overall rating would be excellent. However, if the behavior that is rated as average is the most salient or frequently observed behavior for a particular lesson or activity, the overall rating for the category should be adjusted. Remember to base ratings only on observable behaviors related to lessons and activities.

Rules of Determining Quality Indicators

Use the following guidelines for assigning quality indicators for each instructional even or activity.

1. The majority determines the quality rating
- Rating should be based on observable behavior using professional judgment, not inference.
- The framework for thinking about teacher quality is based on assumption that a teacher who falls into the *Excellent* category is one who addresses the needs of a struggling reader.
- A rating of *High Average*, *Low Average* or *Weak* represents the degree to which a teacher deviates from this standard. For example, a teacher who is rated Low Average may be an effective teacher for most students, but is not addressing the needs of struggling readers.

2. Assignment of *Low Average* or *High Average*
    - *Low Average*: Some indicators under *Weak* are present, but the majority fall under Average.
    - *High Average*: Some indicators under Excellent are present, but the majority fall under Average.
    - Special consideration; If a teacher meets a majority (5) of indicators under *Weak* and all others under *Excellent* the teacher's rating would be *Low Average* for that event.

3. Assignment of *Weak* or *Excellent*
    - To clearly assign either of these extreme ratings, almost all (or supermajority) of indicators must fall within the *Excellent* or *Weak* range.
    - Distinguish between *Excellent* and *High Average* by considering how closely the teacher meets the needs of a struggling reader.

4. Situation: **All** indicators fall within *Average* column
    - Professional judgment should be used to determine whether to rate as *Low* or *High Average*.
    - Remember to keep the struggling reader in mind.
    - If the teacher has farther to go to meet the needs of the struggling reader, rate as *LowAverage*.

VITA

Name:	Angela Renee Sebesta Hairrell
Address:	Department of Educational Psychology Texas A&M University, MS 4225 College Station, TX 77845
Email Address:	ahairrell@tamu.edu
Education:	B.S., Interdisciplinary Studies, Texas A&M University, 1991 M.S., Curriculum and Instruction, Texas A&M University, 1994 Ph.D., Curriculum and Instruction, Texas A&M University, 2008