I, Danielle J Hayes, hereby submit this original work as part of the requirements for the degree of Doctor of Philosophy in Communication Sciences and Disorders.

It is entitled:
Assessing Vocabulary in Context Using Graduated Prompting

Student's name: Danielle J Hayes

This work and its defense approved by:

Committee chair: Jo-Anne Prendeville, EdD
Committee member: Beth O'Brien, PhD
Committee member: Nancy Creaghead, PhD
Committee member: Lesley Raisor, PhD
Assessing Vocabulary in Context Using Graduated Prompting

A dissertation submitted to the

Division of Research and Advanced Studies
of the University of Cincinnati

In partial fulfillment of the
requirements for the degree of

**Doctor of Philosophy**

In the Department of Communication Sciences and Disorders
in the College of Allied Health Sciences

By

Danielle Janine Hayes

B.S. Tennessee State University, 2005
M.A. University of Cincinnati, 2007

Committee Chair: Jo-Anne Prendeville, Ed.D

Committee: Nancy Creaghead, Ph.D
Lesley Raisor-Becker, Ph.D
Beth O'Brien, Ph.D
ABSTRACT

Graduated prompting is a method of dynamic assessment that is used to make predictions about a child’s response to intervention (Gutierrez-Clellan & Pena, 2001). The purpose of this study was to use graduated prompting to assess the word learning ability of low-income fourth grade students when given context clues during a reading task.

Twenty-eight fourth grade students were administered the Expressive One Word Picture Vocabulary Test (EOWPVT; Brownell, 2000) and were placed into Pass (85 and higher) and Non-Pass (below 85) groups based on standard EOWPVT scores. The participants were administered the three phases of the Dynamic Assessment of Vocabulary in Context (DAVIC): 1) Pretest, 2) Graduated Prompting, and 3) Transfer task. Scores on the DAVIC phases were analyzed by group and by EOWPVT score. The frequency of responses at each prompt level was compared to EOWPVT score. DAVIC pretest and DAVIC transfer task scores were compared.

There was no significant difference between the scores of the Pass group ($M=67.00, SD=14.502$) and scores of the Non Pass group ($M=64.57, SD=9.525$) on the DAVIC graduated prompting phase $t(26)=.524, p=.605, d=.19$. Scores on EOWPVT and DAVIC pretest scores were significantly correlated, $r=.599, p=0.01$. EOWPVT scores were significantly correlated with Prompt 1, $r=.429, (p=0.05)$ and Prompt 3, $r=-.376, (p=0.05)$. There was not a significant correlation $r=.184 p=.184$, between scores on the EOWPVT and DAVIC transfer task scores. The participants scored significantly higher on the transfer task ($M=62.32, SD=10.61$) than on the pretest ($M=37.96, SD=5.87$); $t=-10.73, p=<.001, d=2.84$. Results of this study suggested that when assessing a child’s vocabulary skills, Graduated Prompting may be more effective in determining a child’s abilities than a standardized assessment alone.
ACKNOWLEDGEMENTS

I will forever be grateful for the love and support that I have received throughout this process. I would first like to acknowledge the members of my committee. I would like to thank Dr. Jo-Anne Prendeville, who has been my advisor and mentor since my Junior year at Tennessee State University. With her ongoing advice and support, she has played an integral part in my development as a researcher and as a person. I am grateful for Dr. Nancy Creaghead, who has provided me with so many opportunities and life lessons. I will always be thankful to have been her student. Dr. Lesley Raisor-Becker has such a giving spirit and is a wonderful person. I have enjoyed working with her on this and other projects and consider her to be a role model. I am very thankful to Dr. Beth O’Brien who sparked my initial interest in cognition. It is her guidance that planted the seed for my dissertation research.

I would like to thank the faculty and staff in the department of Communication Sciences and Disorders who have supported me since I began my Master’s degree in 2005. Specifically, I would like to thank Dr. Laura and Dr. Richard Kretschmer for being exactly who they are. They may never understand the positive impact that they have had on me. Additionally, I also thank my wonderful classmates at U.C. I am very grateful for the teamwork atmosphere that was created. I would like dedicate this work to the memory of Nathan Sanders who went above and beyond to assist me throughout my Master’s degree.

I am grateful to the individuals who have served as my professional and educational mentors over the years: Dr. Iris Johnson-Arnold, Dr. Dawn Harris, Dr. Nancy Miller and Dr. Carney Sotto. I could never repay them for the seeds that they have planted in my life.

There are many organizations, groups and individuals throughout Cincinnati who have supported me on a number of levels. Specifically, I would like to thank Rev. Dr. Alphonse Allen and the entire Allen Temple A.M.E. Church family. Upon moving to Cincinnati for school, Allen Temple took me in and gave me a family whom I have grown to love very much. I have made many special friends at Allen Temple and throughout Cincinnati who are too many to name. However, I would like to particularly thank the Cannon family for giving me a home away from home.
I must thank Tennessee State University for providing me with the solid foundation to move forward and strive for excellence. My blood still runs blue!

I am very grateful to the members of Unity and my entire extended family in Nashville, TN. You all are very special to me and I thank you for always cheering for me.

To my friends, otherwise known as the “Cincinnati girls” (in no particular order): Jamie, Alayna, Kellie, Erin and Genise- this hasn’t been the first milestone (and won’t be the last) that you have supported me through. I love you guys.

To Kwesi… “thank you” will never be enough. I love you very much and I’m grateful that God saw fit to show me you.

To my dear family, I could write thanks for many pages. To my grandparents (late and living), aunts, uncles, cousins, and my goddaughter- I love you all very much and I thank you for supporting and loving me. To my father-it is because of you that I am who I am today. Success was never an option…it was a commandment and I thank you for that. To my mother- Who knew that after my horrible teenage years that we would end up being best friends? You are TRULY my rock. I know that it was your prayers that got me through this. I’m so happy that God chose me to be your daughter. To Danny-You are my little brother, but I thank you for serving the role of my big brother and simply being wonderful. To DeeDee-In you I have my own personal super cheerleader …who could ask for more? To both Danny and DeeDee, you give me my fire. I run because I know you’re chasing me. Never stop! To Mama, Dad, Danny and DeeDee, I love you all very much and I would have never reached this point without you.

Finally, I would like to thank the One who made everything possible. Thank You and I love You.
## Table of Contents

Abstract........................................................................................................... iii  
Acknowledgements............................................................................................ vi  
List of Tables and Figures.................................................................................... x  
Chapter I: Introduction....................................................................................... 10  
  Purpose of Study............................................................................................... 15  
Chapter II: Literature Review........................................................................... 17  
  Vocabulary and Reading.................................................................................. 17  
  Influence of the home environment................................................................. 18  
  Word knowledge ............................................................................................. 20  
  Fourth grade slump......................................................................................... 25  
  Standardized assessment................................................................................ 28  
  Response to intervention and Dynamic assessment......................................... 32  
  Word Learning and Context Clues................................................................... 40  
  Summary......................................................................................................... 43  
  Research Questions......................................................................................... 44  
Chapter III: Methods....................................................................................... 45  
  Setting............................................................................................................. 45  
  Participants..................................................................................................... 45  
  Procedures..................................................................................................... 47  
  Data Analysis................................................................................................ 54  
Chapter IV: Results......................................................................................... 60  
Chapter V: Discussion...................................................................................... 66
Vocabulary knowledge and word learning ability………………………….67
Unknown, acquainted, established………………………………………….73
Congruence with prior research……………………………………………..76
Teaching a word learning strategy…………………………………………..78
Future research…………………………………………………………………..82
Limitations…………………………………………………………………………83
Implications and Conclusion…………………………………………………..84
References…………………………………………………………………………87
Appendix A………………………………………………………………………..97
Appendix B………………………………………………………………………..98
Appendix C………………………………………………………………………..100
Appendix D………………………………………………………………………..101
Appendix E………………………………………………………………………..103
Appendix F………………………………………………………………………..105
Appendix G………………………………………………………………………..106
List of Tables and Figures

Table 1 ..................................................................................................................62
Table 2 ..................................................................................................................64
Table 3 ..................................................................................................................75
Table 4 ..................................................................................................................82

Figure 1 ..............................................................................................................70
Figure 2 ..............................................................................................................71
Figure 3 ..............................................................................................................72
Chapter I

Introduction

Literacy is a fundamental skill that is utilized in practically every aspect of formal education. As a result, having the ability to read is necessary for educational achievement. However, many children have trouble acquiring this important skill. This has been particularly noted among fourth grade students who are considered ‘at risk’ (Chall, Jacobs & Baldwin, 1990) for reading difficulties. Standardized reading scores from the National Assessment of Educational Progress (NAEP) show that in 2009, only 33% of fourth grade students were at or above a proficient reading level. The Committee on the Prevention of Reading Difficulties in Young Children reports that children from low income households are among those at risk of having a weakness in reading (Snow, Burns & Griffin, 1998). The NAEP reveals that only 17% of low income children are at or above a proficient level of reading (U.S. Department of Education, 2009). Fourth grade students who had a higher income status performed considerably better on the same assessment, with 45% of students performing at or above a proficient level of reading (U.S. Department of Education, 2009).

Children in general who are from low income backgrounds are at risk for reading difficulties, but statistics suggest that these difficulties may be amplified among minority populations. For example, African American students represent a large percentage of the low income population, making up 16.9% of all students who are eligible for free or reduced lunch (U.S. Department of Education, 2006). Also, according to data from the NAEP, in 2009, 16% of African American fourth grade students read at or above a proficient level of reading. The majority of African American fourth grade students (52%) performed below a basic level of
reading on the NAEP. In addition to having the most students (among the represented ethnic groups) below a basic level of reading, statistics from the U.S. Department of Education suggest that African Americans represent a disproportionate number of students served under the Individuals with Disabilities Education Improvement Act (IDEIA) (Blackorby et al., 2010; Synder & Dillow, 2010). Although African American students make up only 17% of the total population in American schools (Snyder & Dillow, 2010), these children make up 16.67% of school aged students served under IDEA (Blackorby, et al., 2010). This statistic can be compared to White American students, who make up 55.8% of the total population in American schools (Snyder & Dillow, 2010) and 14.05% of all students served under IDEA (Blackorby, et al, 2010).

Reading is the process of making meaning from text (Rosenblatt, 1994). This meaning is based on one’s schema, or the knowledge that a reader has prior to encountering that particular text (Weaver, 2002). Consequently, one’s prior experiences will greatly influence how a particular text is comprehended. The reading weakness experienced by many low income children may be the result of differences in background experiences (Hart & Risley, 1995). If a child does not have the prior experiences required to comprehend the text that they are encountering, then it can be assumed that they would have difficulty understanding it. Another factor that can influence a child’s ability to comprehend a text is their background experiences with vocabulary. Vocabulary knowledge is a strong predictor of a child’s ability to comprehend what they read (Davis, 1944, 1972). There are multiple relationships between vocabulary skills and reading comprehension. First, vocabulary abilities can predict reading comprehension (Zinar, 2000). Next, readers can learn word meanings during the normal reading process (Swanborn & Glopper, 1999). Also, instruction in vocabulary can improve reading comprehension (McKeown, Beck, Omanson & Perfetti, 1983).
Because of the disproportionate number of low income and African American students in special education, and the influence of background experiences on learning, prevention is possibly the most useful course of action in literacy learning (Snow, Burns & Griffin, 1998). Therefore schools need a preventative model to best serve their students who are at risk for reading difficulties. Current legislation seeks to provide such a model in the Individuals with Disabilities Education Improvement Act (IDEIA, 2004) and the implementation of Response to Intervention (RtI). Response to Intervention is a preventative model (Justice, 2006) that allows educators to target student areas of weakness without or before labeling them with a specific learning disability. It is important to accurately identify children who are in need so they can receive the necessary assistance. However, children are often evaluated using some form of standardized assessment which, because of content and linguistic bias (Laing & Kamhi, 2003), ultimately leads to them being labeled with a disability (Washington & Craig, 1992). In an RtI model, assessment is paired with intervention in order to determine and target a student’s area of weakness and to provide the necessary assistance to avoid the over identification of children for special education services.

Dynamic Assessment, which is traditionally an assessment/intervention method, may be a useful tool for identifying and examining student weaknesses within an RtI model. Although there are multiple dynamic assessment models, the graduated prompting method is one successful model that is used to make predictions about a child’s response to intervention (Gutierrez-Clellan & Pena, 2001). The Graduated Prompting method has been shown to be useful when observing and assessing the language skills of young children (Gutierrez-Clellan & Pena, 2001). Researchers have also used Graduated Prompting to study the ability of older children to use morphological analysis to determine word meanings (Larsen & Nippold, 2007).
However, there are other vocabulary strategies, besides morphological analysis, that have not been examined using the graduated prompting method. One such strategy is the use of context clues. All words may not provide morphological cues, and many times readers have only the reading passage, context clues and prior knowledge on which to depend. As a result, researchers need to determine if graduated prompting can be used effectively to assess the vocabulary abilities of fourth grade students when they are presented with a reading task.

**Significance of the Problem**

Vocabulary has been shown to be problematic for some fourth grade students, particularly when the words are academic in nature (Chall, Jacobs & Baldwin, 1990). As stated previously, because only 33% of fourth grade students perform at or above a proficient level of reading, it is clear that many fourth grade students, despite ethnicity or SES, often have difficulties in reading (U.S Department of Education, 2009). The occurrence of these difficulties becomes more acute in populations described as “at risk” (Snow, Burns & Griffin, 1998) for reading difficulties. As a result, researchers have focused on the “fourth grade slump” which refers to the decline in reading skill that occurs (predominantly) among low income, fourth grade children (Chall, Jacobs, & Baldwin, 1990). There could be several causes of this, including: factors related to socio economic status (Hart & Risley, 1995; Noble, Wolmetz, Ochs, Farah, McCandliss, 2006), differences in home experiences (Scarborough, Dobrich & Hager, 1991; Hart & Risley, 1995), and lack of experience with expository text (Duke, Bennett-Armistead & Roberts, 2003). Another cause of the decline could be the increasing complexity of vocabulary that occurs in fourth grade. Children do not simply “know” or “not know” a word. Rather, word learning occurs in increments (Beck, McKeown, McCaslin & Burkes, 1979), and is the result of prior experiences. As a result, it is likely that many students, even those without reading
difficulties, will encounter unknown words in the academic environment. When this occurs, steps need to be taken to determine if the children’s vocabulary difficulty is the result of an actual weakness or simply a lack of exposure before labeling them with a learning disability. In order to determine this information, it is useful to determine a student’s zone of proximal development, that is, to determine how the student performs independently and then how they perform with assistance (Vygotsky, 1978). Teachers need to be aware of methods that can be used to determine this information. Speech Language Pathologists (SLPs) and other qualified professionals often use assessments such as the Peabody Picture Vocabulary Test (PPVT-III, Dunn & Dunn, 1997) and the Expressive One-Word Picture Vocabulary Test-Revised (EOWPVT; Gardner, 1990) to determine a child’s vocabulary abilities. Performance on standardized tests, however, can be influenced by income (Burton & Watkins, 2007, Noble, et al., 2006), and does not completely measure word learning ability among children from high risk populations (Burton & Watkins, 2007). Additionally, biases (construct, method and item) that are based on cultural differences need to be considered when administering standardized tests (Scholmerich, Leyendecker, Citlak, Caspar & Jakel, 2008). Further, a standardized assessment alone does not reveal a child’s zone of proximal development because scaffolding is usually not allowed beyond scripted test trials.

Unlike standardized tests, performance on Dynamic Assessment has shown to be uninfluenced by income (Burton & Watkins, 2007), and can be used to determine a child’s zone of proximal development by integrating intervention into the assessment process (Gutierrez-Clellen & Pena, 2001). Graduated prompting is a Dynamic Assessment method that can be used to assess a child’s use of vocabulary strategies (Larsen & Nippold, 2007), but there is little information about how this method can be used during a reading task. This method needs to be
examined in order to: demonstrate the usefulness of Graduated Prompting during a reading task; to reveal the level of prompting that fourth grade students need in order to determine word meanings; and to determine if using context clues as a strategy results in transfer of knowledge. Using the Graduated Prompting strategy, teachers may be able to use scaffolding to determine a child’s response to intervention in the area of vocabulary.

**Purpose of Study**

This is a quantitative investigation of the assessment of vocabulary abilities in low income fourth grade students. The purpose of this study was to use the Graduated Prompting method of dynamic assessment to assess the word learning ability of low income fourth grade students when given context clues during a natural reading task. Specifically, this research sought to determine if the information revealed during graduated prompting could be used to measure vocabulary knowledge and word learning potential. The purpose was also to determine if graduated prompting directed at context clues results in improved word knowledge and to measure students’ level of independence in deriving word meaning from context. Finally, the purpose of this study was to determine if there is a relationship between static word knowledge and word learning potential (as measured by a standardized assessment and graduated prompting, respectively).

This study used a model similar to a study conducted by Larsen and Nippold (2007) in which children were given a series of prompts related to morphological knowledge to assess word learning ability. The research questions of the current study were:

1. Is there a significant difference between the scores on the Dynamic Assessment of Vocabulary in Context (DAVIC) graduated prompting phase of participants who scored
more than one standard deviation below the mean (Non Pass) and those who did not score more than one standard deviation below the mean (Pass) on the EOWPVT?

2. Is there a significant relationship between EOWPVT standard score and the frequency of responses at each Level of Successful Prompting?

3. Is there a significant relationship between EOWPVT standard scores and DAVIC pretest scores?

4. Is there a significant relationship between EOWPVT standard scores and the DAVIC transfer task scores?

5. Is there a significant relationship between scores on the DAVIC pretest and scores on the DAVIC transfer task?

This research study has the potential to develop an assessment tool that supports SLPs in an RtI model by determining levels of prompting needed to address weaknesses in vocabulary for fourth grade students.
Chapter II

Review of the Literature

The purpose of this study was to use Graduated Prompting to examine the word learning abilities of low income fourth grade students during an authentic reading task using context clues as a word learning strategy. This study further intended to compare participant performance on a Graduated Prompting task with performance on a standardized vocabulary assessment. This chapter addresses the following topics: 1) Vocabulary and reading, 2) influences of the home environment, 3) word knowledge, 4) fourth grade slump, 5) standardized assessment, 6) response to intervention and dynamic assessment, 7) word learning and context clues and 8) a summary that includes this study’s research questions.

Vocabulary and Reading

Vocabulary is a skill that is necessary for successful reading (Zinar, 2000; Qian, 2002). According to the Family Partnership in Reading Initiative and the National Research Council, there are five evidence based building blocks for teaching children to read. These building blocks are based on the development of the following skills: phonemic awareness, phonics, fluency, text comprehension, and vocabulary (Armbruster, Lehr & Osborne, 2003).

Children’s vocabulary knowledge is strongly related to their ability to comprehend what they read (Davis, 1944, 1972). This relationship was examined when ninety-five fourth grade students completed a battery of tests in the areas of reading, comprehension monitoring and word identification. The results showed that word identification predicted reading comprehension (Zinar, 2000). In another study, 217 English language learners were given a series of tests which assessed reading comprehension, depth of vocabulary knowledge and vocabulary size. The research revealed that depth and size of vocabulary predicted reading comprehension (Qian,
Another area where reading and vocabulary are related is instruction. Specifically, instruction in vocabulary can result in improved reading comprehension (McKeown, et.al, 1983). During an intervention study, the vocabulary and reading comprehension skills of fourth grade students were examined. Participants received vocabulary instruction over a period of five months. Those who received the instruction showed significant improvements in reading comprehension when compared to the control group (Beck, Perfetti & McKeown, 1982). The improvement of reading comprehension that can occur after vocabulary instruction further illustrates the relationship between reading and vocabulary. There are factors that contribute to both reading and vocabulary. Prior experiences and background knowledge are such factors.

**Influences of the Home Environment**

A child’s literacy skills and overall school success are influenced by his or her home environment (Scarborough, Dobrich & Hager, 1991; Heath, 1983; Purcell-Gates, 1996). One area that is particularly influenced by the home environment is vocabulary. Vocabulary is learned through experience with words. Hart and Risley (1995) found that these experiences are shaped by environmental influences such as socio-economic status (SES). In their longitudinal study, Hart and Risley, (1995) examined family practices in regard to use of language with children from different economic (which ranged from professional to welfare) and ethnic backgrounds. Hart and Risley (1995) revealed that children who come from lower SES homes have smaller vocabularies than children who come from higher SES homes. More important than the actual SES, was the difference in vocabulary experiences encountered by the families in the different SES groups. A key difference noted in the study was frequency of experience. By the age of three, children from the welfare families had been exposed to much less vocabulary than children from professional families. For example, the research showed that children from low
income families experienced 616 words per hour while children from the working class families experienced 1,251 words per hour. The gap between the income groups became even more noticeable when the professional families were considered. Children from these homes experienced 2,153 words per hour. The researchers also found that the children from the professional families had experienced a higher quality of words than children from low income homes. One phenomenon revealed in the study was that words recorded in children’s vocabulary at age three were similar to the words recorded in their parents’ vocabulary, despite SES background. This important study highlights the importance of the home environment in the language development of young children. This home environment is additionally important because research shows that there is a steady relationship between the skills with which a child enters school and later performance. The researchers found that children’s rate of vocabulary growth at age three was related to their vocabulary at age nine and ten (when children are typically in the fourth grade). They also found that vocabulary at age three predicted language skills at age 9 and 10 (Hart & Risley, 1995).

Environmental influence is a key contributor to vocabulary performance among high and low income children. As Hart and Risley (1995) illustrate, children from varying SES backgrounds can have diverse experiences with words. These experiences may result in differing interpretations of encountered texts. There is fMRI research to show that these experiences may even result in different brain activation patterns during reading tasks. Specifically, during phonological tasks, children from lower SES backgrounds have higher levels of fusiform activity than children from higher SES backgrounds (Noble et al., 2006). The differences caused by home experiences, specifically SES, are made evident in the large number (49%) of low income fourth grade children who scored below a basic level of reading on the NAEP in 2009 (U.S.
Department of Education, 2009). Considering the performance gap between low and high income children on standardized assessments of reading (U.S. Department of education, 2009) and the fact that vocabulary predicts reading (Davis, 1944, 1972), it becomes necessary to explore both vocabulary and assessments among low income children.

**Word Knowledge**

There are different ways to describe word knowledge. It is often explained using a hierarchy of levels or tiers (Armbruster, Lehr, & Osborn, 2003; Beck, McKeown & Kucan, 2002). The Tier system describes the type of words that are encountered and their functionality. Tier One consists of basic level words which will likely not need to be taught. Tier Two consists of functional words that are encountered in one’s daily vocabulary at a high frequency. Tier Three words are words that are encountered at a low frequency and are specific to a particular subject area (Beck et al., 2002). These functional aspects of word knowledge become important in the school setting where vocabulary assessments are used to measure children’s knowledge and determine eligibility for special education services (Watkins & DeThorne, 2000). This is because word knowledge is based on exposure and many children may not have encountered words (such as Tier Three words) that are presented on vocabulary assessments. Understanding the ways that vocabulary knowledge is described may assist special educators in their service delivery, including both assessment and intervention. In order to describe a child’s vocabulary knowledge, one must first be aware of the complexity of word knowledge (Nagy & Scott, 2000). There are several features of word knowledge. Some of these features are that: word knowledge is multidimensional, word learning is incremental; words can be ambiguous, and words are connected (Nagy & Scott, 2000).
**Word knowledge is multidimensional.** In order to describe a child’s word knowledge, it must first be understood that a word has different dimensions. One way to describe these dimensions of word knowledge is by quality. Perfetti and Hart (2002) use the lexical quality hypothesis to distinguish the levels of representations one can have for words when reading. Lexical quality, which is defined in terms of high and low, is dependent on the retrieval of three constituents: orthographic (spelling), phonological (pronunciation) and semantic (meaning) specifications of a word. The lexical hypothesis states that every word has each of these constituents. A high quality representation is one in which all three constituents can be retrieved.

A reader with many high quality word representations is likely to be a more skilled reader than a reader with fewer high quality word representations (Perfetti & Hart, 2002). A representation of a word that lacks one of the three constituents is low quality. For example, a child may see a pony at a petting zoo and is able to identify the animal phonologically by name (/poni/). This child also demonstrates semantic knowledge of the term because they can identify the animal. However, if the child lacks the orthographic representation of the word: “p-o-n-y”, then they have a low lexical quality for that word for reading. When a child approaches a low lexical quality word during a reading task, they may still be able to comprehend the overall text by compensating with the other knowledge that they do have. For example, Recht and Leslie (1988) found that having prior knowledge about a topic can assist poor readers because they can compensate for their low decoding ability by activating their schema and recognizing important details. Therefore, using a tool such as dynamic assessment to evaluate word knowledge may give children an opportunity to activate their schema and derive word meaning from context.

**Word learning is incremental.** According to Beck, et al., (1979), there are three levels of word knowledge: Unknown, Acquainted and Established. A word described as unknown is a
word that one has no experience with and therefore no knowledge of. One who has unknown knowledge of a word is completely unfamiliar with the term and its meaning. One that is acquainted with a word has a basic understanding of the word’s meaning. With this level of knowledge, one may be able to understand the meaning but the knowledge is not established enough for automatic use. With established knowledge, however, one may easily and immediately recognize the meaning of the word and can use the word appropriately.

Usually, in order to gain an established level of knowledge of a word, multiple exposures to the word must occur (McKeown, Beck, Omanson & Pople, 1985). This was examined during an intervention study in which different instructional and frequency conditions were examined. Four classes of fourth grade students were instructed using two frequency conditions: four exposures and twelve exposures. The study showed that the participants who received twelve exposures performed better than those who received only four. Because word learning occurs in increments, using traditional assessments may not completely reveal a student’s true level of word knowledge. This is because traditional assessments don’t typically acknowledge the different increments of word learning. Instead, these assessments measure whether the word is established (fully known) or not. Therefore, using a tool such as dynamic assessment may be more useful in providing practical information about a child’s word knowledge. For example, Pena (2001) found that a dynamic assessment provided more information about a child’s word knowledge than using a standardized assessment alone.

**Words can be ambiguous.** There are some vocabulary concepts that contribute to the complex nature of word learning. One such concept is polysemy, which is a concept that describes how one word can have multiple “senses” (Garcia et al., 2007). For example, the word “feed” can refer to the verb (meaning eating) as in “to feed” or it can refer to a noun (meaning
what is eaten), such as “chicken feed”. The term “feed” in the verb sense is semantically similar to the word in its noun sense. Words that are polysemous are different from homonyms because homonyms are words that are spelled alike and sound alike but have completely different meanings (Klepousniotou, et al., 2008). Research has shown that comprehension is affected by words that have semantically overlapping senses (Klepousniotou, 2008). The word “bark” is an example of a polysemous word and a homonym. It can refer to the verb (ex. “I hear the dog bark.”) or the related noun (ex. “I hear the dog’s bark.”) or an unrelated noun (ex. “The tree has bark.”). The phonologically and orthographically identical words are semantically unrelated. Adding to the ambiguity of word learning are homophones (words that are phonologically identical but are orthographically and semantically different, ex. “meat”, “meet”) and homographs, which are orthographically identical but semantically and phonologically different such as “bow” which can refer to the nouns /bou/ and /bau/ or the verb /bau/ (Templeton, 2003).

**Words are connected.** Understanding the connected nature of words is much like understanding the connected nature of knowledge in general. Learning and understanding new concepts is largely dependent on the prior knowledge one has about related concepts. For example, in order for a child to understand the word “pet”, they must have previous exposure to some concepts about pets (e.g. They are furry, they have four legs, and they make noise). As the child gains more experience with items in the pet category (e.g. dog, cat, and hamster) they can understand what is meant by the term “pet”. With experience they could eventually learn that pets can also be reptiles. One’s definition is shaped and perfected with each pet encounter. This concept is known as building a schema. One’s schema for a particular concept develops and changes with experience. As this process occurs, more and more connections are made, causing each word to connect to the other in some fashion. The word “pet” would not have any meaning
without any knowledge of the actual animals in that category. It is the connections that occur between words that make it necessary to teach words in context. Because words are not isolated from one another, assessment and intervention techniques that treat them as individual entities are not truly authentic (Nagy & Scott, 2000). Dynamic assessment is an evaluation method which can be designed to model an authentic, classroom based task.

In summary, when assessing vocabulary skills, there are several factors to be considered: word knowledge is multidimensional; word learning is incremental; words can be ambiguous; and words are connected. Standardized vocabulary assessments such as the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997) evaluate word knowledge out of context and may not display a child’s true word learning ability when used alone (Burton & Watkins, 2007). Additionally, vocabulary knowledge is based on background experiences and factors related to income (Hart & Risley, 1995). Consequently, children from non-mainstream background (minorities and children from low income homes) tend to score lower on standardized vocabulary assessments (Thomas-Tate & Edwards, 2004; Tabors, Páez & Lopez, 2003; Gray, Plante, Vance, & Henrichsen, 1999) which may result in the over-diagnosis of a language disorder (Washington & Craig, 1992). And although standardized tests are useful measures that provide normative data, they may not be effective measures of vocabulary when used alone. To this point, several implications have been made: non-mainstream groups are overrepresented in special education because of poor performance on standardized tests (Washington & Craig, 1992; Blackorby, et al, 2010); non-mainstream groups perform poorly specifically on reading assessments (such as the NAEP); there is a relationship between vocabulary and reading (Zinar, 2000) and the complex nature of vocabulary knowledge prevents it from being assessed completely with standardized tests. Another factor to be considered in vocabulary assessment is
grade level. Difficulties encountered when children take standardized vocabulary tests could be the result of increased vocabulary demands on students as they increase in grade level.

**Fourth Grade Slump**

Research shows that vocabulary difficulties often arise in low-income students when they reach the fourth grade (Chall, Jacobs & Baldwin, 1990). The “fourth grade slump” is a phenomenon primarily seen among low income students who suddenly experience a decline in reading performance at grade four (Chall & Snow, 1988). The theory behind this occurrence is based on Chall’s (1983) stages of reading. Chall (1983) explains that there are five (0-5) stages of the reading process. Stage 0 is the pre-reading or the emergent literacy stage. During this time, children become familiar with the components of emergent literacy as they prepare to move forward to conventional reading. Stages 1-2 occur during grades one through three. During this time, children learn to read. They grasp the alphabetic principle which is necessary for decoding and they learn vocabulary which is familiar and occurs in their everyday setting. Stages 3-5 occur during grades four through eight. During this time, reading becomes more complex and children encounter unfamiliar, academic vocabulary. At this time, they are expected to read to learn.

Chall, Jacobs and Baldwin (1990) performed a seminal study which provides quantitative evidence of the “fourth grade slump” phenomenon. The researchers longitudinally examined reading development in low income children over a period of two years beginning at grades two, four and six. The results showed that at grades one through three, the children performed at the same level as the normative population. Despite early success, their performance began to drop abruptly at grade four, particularly affecting word meaning. Chall, et al. (1990) showed that the participants had difficulty defining abstract, academic and literary words. As the participants
progressed in grade level, they continued to decline strongly in word meaning and then in overall reading ability which included both reading comprehension and oral reading. The study demonstrates the overall importance of vocabulary for later reading success as well as the importance of the fourth grade transition. Participants who demonstrated vocabulary difficulties at grade 4 continued to have vocabulary and ultimately other reading difficulties at grade seven. Additionally, even though some children performed well in grades one through three, they experienced a sudden decline in word meaning in grade four.

**Lack of expository text experience.** Some experts attribute the fourth grade slump to a lack of expository text experience in the primary grades (Duke, Bennett-Armistead & Roberts, 2003). For example, in a survey, Yopp and Yopp (2000) found that only fourteen percent of materials read aloud in class by primary grade teachers were expository. There is an overall shortage of informational text experienced by first grade students. Classroom observations show that first grade students are exposed to informational texts for an average of 3.6 minutes per day (Duke, 2000). The dearth becomes even more evident in low income schools, with first grade students being exposed to informational texts on an average of 1.9 minutes per day (Duke, 2000). The shortage of expository material may also carry into the older grades. A study of second, fourth and sixth grade basal readers found that most of the text at each grade was fictional, followed by poetry (Moss & Newton, 2002). This can be a problem as children take standardized tests because research shows that children are tested on expository materials more than they are exposed to this text in their basal readers (Flood & Lapp, 1986). The lack of expository text may result in a lack of exposure to necessary academic vocabulary. Research shows that incidental word learning can occur during the reading process (Swanborn & Glopper, 1999). Students are exposed to academic vocabulary in expository texts (Armbruster & Nagy, 1992). The
implication here is that if children are not exposed to expository texts, then the opportunity for incidental academic word learning is missed.

Another explanation for the fourth grade slump (as it relates to exposure to expository texts) could be the change in cognitive demand (Willson & Rupley, 1997). Students in grades two and three, who encounter more narrative texts, use phonemic and background knowledge to comprehend text. However, students in the fourth grade, who begin to experience expository text, tend to use strategy knowledge (knowing what and how to read) in order to comprehend texts (Willson & Rupley, 1997). Strategies focused on vocabulary may be particularly instrumental when reading an expository passage. Research shows that readers (whether good or poor comprehenders) tend to focus on vocabulary in order to read an expository text. For example, Kletzien (1991) examined the strategies that 48 high school students used to read expository text passages. The results showed that “focusing on vocabulary” was the strategy most used by all participants, despite comprehension ability of the participants. As a result, it may be useful to provide strategies surrounding vocabulary in order to assist a child with reading an expository passage.

Children from low-income homes tend to experience a decline in vocabulary skills (as measured by standardized tests) when they enter the fourth grade (Chall, et al., 1990). Statistics show that this demographic group performs poorly on the NAEP (49% below basic) (U.S. Department of Education, 2009). This may be the result of a lack of expository text experience. Because vocabulary skills are important for reading comprehension (Zinar, 2000) and children can learn words by reading (Swanborn & Glopper, 1999), it is important that children are exposed to academic vocabulary multiple times during the reading process (McKeown, 1985). Despite this, children receive more exposure to expository text (and ultimately academic
vocabulary) from standardized tests than they do from their basal readers (Moss & Newton, 2002, Flood & Lapp, 1986). Assessing vocabulary is complex (Nagy & Scott, 2000) because of the multidimensional, incremental, ambiguous and connected nature of words. It is further complicated because vocabulary is influenced by prior experiences and income. When examining vocabulary skills, both complexity and experiences related to income are factors.

**Standardized Assessment**

The influence of cultural and environmental factors has been particularly noted in research regarding standardized assessments (Noble, et al., 2006). Minorities and children from low income backgrounds are at risk for reading difficulties (Snow, Burns & Griffin, 1998) and often perform poorly on standardized assessments (Thomas-Tate, Washington & Edwards, 2004; Tabors, Paez & Lopez, 2003). As a result, these underrepresented children may be categorized as having academic disabilities, causing them to be overrepresented in special education (Washington & Craig, 1992; U.S. Department of Education, 2006).

There is research that shows that minorities and children from low-income homes perform lower than the normative group on standardized assessments (Thomas-Tate, Washington & Edwards, 2004; Craig, Thompson, Washington & Potter, 2003). This could be the result of biases caused by cultural and linguistic differences. For example, Scholmerich et al. (2008) used a standardized assessment to evaluate a variety of skills (including gross motor, fine motor, memory, action strategies, categorization, body awareness, receptive language, expressive language, social development and emotional development) in Kindergarten children. The study compared the scores of migrant/minority children with those of non migrant children. Although most of the migrant children received the tests in their first language, they required more time to take the tests. Also, the migrant children scored lower than the non migrant children on all the
measures. These results suggest that children from diverse backgrounds may be subject to standardized test biases based on their linguistic and cultural differences.

Standardized measures of vocabulary have been particularly criticized for not accurately measuring word learning ability among different populations (Burton & Watkins, 2007). Blachowicz, Fisher, Ogle and Watts-Taffe (2006) discuss how standardized vocabulary assessments have changed little over the years and do not truly assess a child’s word knowledge. One reason for this assessment problem is that word knowledge is based on experience. Children from varying backgrounds enter the general education classroom with different experiences. Furthermore, standardized assessments do not capture what the child can learn. Instead, they measure what the child already knows. As a result, children with fewer world experiences are penalized for not having the same learning opportunities. These unequal learning experiences become particularly significant for children raised in low income homes. Beitchman et al., (2008) longitudinally examined vocabulary development and growth in individuals over a period of twenty years (from age 5 until 25). When considered separately from other factors, family income was found to be a significant predictor of vocabulary growth as measured by the Peabody Picture Vocabulary Test (PPVT-III; Dunn & Dunn, 1997). Other predictors of PPVT performance were the mother’s age at birth as well as mother’s years of education. Research shows that there is a gap in the standard scores of low risk (M=102.33, SD=11.44) and high risk students (M=89.83, SD=10.01) on the PPVT-III (Burton & Watkins, 2007). The relationship between low income status and poor performance on vocabulary assessments is not limited to the PPVT-III. Gray et. al., (1999) assessed 31 four and five year old children with four standardized vocabulary tests: PPVT-III (Dunn & Dunn, 1997), Receptive One Word Vocabulary Test (Gardner, 1985), Expressive Vocabulary Test (Williams, 1997), and the Expressive One Word
Picture Vocabulary Test-Revised ([EOWPVT]; Gardner, 1990). The researchers found that none of the tests were successful identifiers of children with specific language impairment (SLI). The researchers concluded that standardized vocabulary assessments should not be the sole tool used to identify a child as language impaired. Washington and Craig (1992) assessed 105 African American preschool and kindergarten children using the Peabody Picture Vocabulary Test-Revised ([PPVT-R]; Dunn & Dunn, 1981). Participant scores were compared to the normative data. The results showed that 91% of the children scored below the mean on the assessment. The authors suggest that the participants’ poor performance on the PPVT-R was due to problems in the norming procedure, in that there was an absence of PPVT-R vocabulary items in the lexicon of participants.

The complex nature of vocabulary in general provides yet another reason for the vocabulary assessment problem. As stated previously, some of the complex features of word knowledge are that it is multidimensional, incremental, ambiguous and connected (Nagy & Scott, 2000). Words are not simply “known” or “unknown” but the knowledge about their meanings occurs in stages. Having a child name or point to a picture of an item does not reveal knowledge about whether or not a child can use the word. Having a child define a word does not provide information about that child’s knowledge of the word in context. Further, neither of those evaluation methods informs intervention by providing information about strategies the child used to determine their answer. There is valuable word learning information that is not captured by traditional standardized tests. In order to gain rich information regarding vocabulary knowledge, a more effective method of assessment should be implemented. Using an alternate form of assessment may result in fewer students unfairly identified as have a learning disability.
Differences in performance on standardized tests could also be the result of an overall difference in brain activation patterns. Noble et al. (2006) examined fMRI scans among children from different economic backgrounds during a reading task. The researchers found that the brain activation patterns were varied among income groups, suggesting that experiences related to income levels may influence brain activation patterns. The results showed that neurobiological, social and cognitive correlates of learning work together as a system. The authors suggest that the differences in activation patterns may influence standardized test performance.

In summary, a child’s literacy skills are important for their school success. Minorities and children from low-income homes typically have difficulties in the area of reading (Snow, Burns & Griffin, 1998), which become particularly evident in the fourth grade. Vocabulary is a skill that predicts reading ability (Zinar, 2000) and is a necessary skill for children to comprehend what they read. Deficiencies in vocabulary become particularly evident in the fourth grade (Chall, Jacobs & Baldwin, 1990). Children are able to learn vocabulary through multiple exposures but they are not frequently exposed to the vocabulary found in expository texts (Duke, 2000). They are, however, exposed to expository texts on standardized tests (Moss & Newton, 2002). Studies show that minorities and children from low income homes often perform poorly on standardized assessments, specifically those focused on vocabulary (Beitchman et al., 2008, Gray et al., 1999). It is important that children described as ‘at-risk’ for reading difficulties are neither over-identified, nor under-identified for special education services. Testing is necessary to identify and assist children who need extra support. This, along with the complex nature of vocabulary is reason to examine other ways of measuring vocabulary skills.
Response to Intervention and Dynamic Assessment

Response to intervention. Response to Intervention (RtI) is an instructional model which was implemented to change how disabilities are identified and targeted in the school by using frequent assessments and evidence-based intervention methods in the general education curriculum (Justice, 2006). Some RtI models are described in tiers with each tier representing a different level of instruction. One common RtI model is the three tiered model (Ehren and Whitmire, 2009). In a three tiered model, the first level is Tier One, which represents Primary instruction. Within this Tier, evidence-based instruction is provided within the general education classroom. Children who experience difficulties in this Tier are moved to Tier two, which is called Secondary intervention. At this level, direct instruction is provided, and children are assessed frequently to monitor their response to the intervention. Children who need further intervention after this level are moved to Tier three, where they are evaluated for special education services (Ehren and Whitmire, 2009). The theory behind RtI is that if evidence-based instruction (paired with frequent assessment) is provided at each level, then fewer children will require special education services (Vellutino et al., 2006; VanDerHeyden, 2007).

The efficacy of RtI has been examined in research. One evidence based tier model is the System to Enhance Educational Performance (STEEP; Witt, Daly, & Noell, 2000). Within STEEP (which is an adapted RtI program), classroom intervention and school wide screening occurs at the first tier. Children who score below the standard move on to tier two. At this tier, students are given an “incentive” for exceeding their previous screening performance. A student who does not improve (as determined by another assessment) moves on to tier three and receives individual intervention by the classroom teacher (or other educator) in the regular education classroom. The child is then assessed weekly to determine progress. The expectation is that most
children will respond to intervention and will not need further instruction (Witt, Daly, & Noell, 2000). This model was examined in five schools (grades 1-5) to determine its effects on the number and percentage of evaluations that qualified for special education services and the identification rates when compared to baseline data. The study showed that the RtI model resulted in fewer evaluations (speech, cognitive and behavior) being conducted. Of the evaluations conducted, more students were identified for services (suggesting that more students were appropriately referred for an evaluation) (VanDerHeyden, Witt & Gilberton, 2007). The results suggest that using a tier based RtI can reduce the number of students who need special education services and thereby increase the accuracy of the identification system.

In a longitudinal study of Kindergarten and first grade students, participants received a series of assessments and intervention which began at their Kindergarten year. Students who were defined as “at risk” for reading difficulties (as determined by the first screening) were placed either in a treatment group or a non-treatment group. The treatment group received intervention twice a week for thirty minutes per session in a small group. When compared to the non-treatment group, the treatment group performed better on literacy screenings, and fewer children were still poor readers in the first grade. Results also showed that 73% of the children who received kindergarten intervention had average skills at the end of the third grade (Vellutino, Scanlon, Small & Fanuele, 2006). The results also show that pairing assessment with intervention is an effective method for measuring a child’s progress.

Dynamic assessment may be a good tool to aid in the RtI model because evaluation and instruction are incorporated into both processes. Dynamic assessment is a method of evaluation which implements a variety of instructional methods (verbal feedback, prompting, questioning, etc.) within the assessment process (Lidz and Pena, 2009). It is based on Vygotsky’s concept of
the Zone of Proximal Development (ZPD) (Vygotsky, 1978). The Zone of Proximal Development is the difference between how a child performs unsupported and how a child performs with assistance. By providing a child with assistance (through prompts, questioning, etc.) during an assessment, the examiner is able to determine a child’s performance with no assistance and their performance with some form of intervention (or their ZPD) (Gutierrez-Clellen & Pena, 2001). The Individuals with Disabilities Education Improvement Act (IDEIA) of 2004 allows schools to use a child’s ZPD or response to research-based intervention as part of their assessment procedures to identify students with specific learning disabilities.

**Types of Dynamic Assessment.** There are three basic types of dynamic assessment: testing the limits, test-teach-retest and graduated prompting (Gutierrez-Clellen & Pena, 2001).

**Testing the limits.** Testing the limits is a dynamic assessment approach in which differentiated levels of feedback (through questioning or interviewing) are used to determine if a child understands the task and material presented. Feedback is described as simple or elaborated (Carlson & Wiedl, 1992). With simple feedback, the child is told whether or not their answer was correct. With elaborated feedback, the child is told whether or not their answer is correct as well as why the answer was correct or incorrect (Carlson & Wiedl, 1992). This method has been used in studies by Carlson and Wiedl (1976, 1978) and Carlson and Dillon (1978). Carlson and Dillon (1978) found that feedback was useful when using a static measure to assess children who are deaf. This method has also been used to measure semantic knowledge. In a case study, Pena et. al. (2001) examined the vocabulary performance of two children with normal language skills on the EOWPVT-R (Gardner, 1990). One participant received elaborated feedback on the static assessment while the other participant received no feedback. The child who received feedback received a score of 91 on the assessment while the other participant received a 74. Further,
testing in the feedback condition provided information about the child’s overall semantic abilities (Pena, et al., 2001).

**Test-teach-retest.** The test-teach-retest model is a traditional method of dynamic assessment which can be integrated within the RtI model (Lidz & Pena, 2009). The first step of this model is administering a pretest and determining the child’s area of difficulty. After this, a mediated learning experience occurs in which instruction is provided in the target area. Following the mediated learning experience, the child is assessed again and his or her progress is noted (Guiterrez-Clellan & Pena, 2001). This method can also be used to differentiate between a language disorder and a language difference (Gutierrez-Clellan & Pena, 2001). Test-teach-retest was used to assess the language learning abilities of 23 Native American children. The participants were taught categorization skills during a mediated learning experience. The results showed that the test-teach-retest procedure was useful in differentiating stronger and weaker language learners. The test-teach-retest procedure has also been used to examine vocabulary skills. Pena, Iglesias and Lidz (2001) used this method to examine word learning ability in 79 Head Start children with typical and low language ability. The results showed that the change observed from the pretest to posttest scores was more useful in predicting word learning ability than a static measure.

**Graduated prompting.** Graduated prompting is a dynamic assessment method in which a series of predetermined probing questions or prompts are presented to determine a student’s immediate learning potential (Oslwang & Bain, 1996), their level of independence in answering a question and their ability to transfer concepts to a similar but different task (Gutierrez-Clellen & Pena, 2001; Grigorenko, 2009). Learning ability is measured by transfer, which is the distance between the child’s ability to perform the original task and their application of concepts to new
tasks. This distance can be measured using a scale which ranges from no transfer, to near transfer, to far transfer, to very far transfer (Gutierrez-Clellen & Pena, 2001). The definition of each level of transfer is unique to the test given. However, based on prior research (Campione & Brown, 1987) a “no transfer” determination could be made when a child exhibits the same behaviors during a new task as they did in the original task, while a “very far transfer” could be determined when a child demonstrates that they have mastered the skill targeted in the assessment. A near or far transfer would be determined based on criteria set by the examiner.

Graduated prompting can predict immediate learning. Bain and Olswang (1995) examined the validity (construct, concurrent and predictive) of Graduated Prompting as a method of assessment. Fifteen children with specific expressive language impairment participated in this 9 week (3 weeks of assessment, 3 weeks of treatment, and 3 weeks of follow-up) study. During the assessment phase, verbal cues were presented in a hierarchic (minimal to maximal) manner in order to elicit a two word utterance from the participants. The results revealed construct and predictive validity of the graduated prompting procedure. The Graduated Prompting procedure was determined to have construct validity because participants were able to produce two word utterances as they received more supportive prompts. The Graduated Prompting method was also found to have predictive validity because children who performed better on the dynamic assessment showed greater changes in language production over time. Concurrent validity was examined by comparing Graduated Prompting to a language sample measure. The thirty minute language sample was more effective at eliciting two word utterances than the fifteen minute Graduated Prompting assessment and therefore concurrent validity was not found. This could be the result of more opportunities for two word utterances to occur during the language sample. Despite this, Bain and Olswang (1995) demonstrated that Graduated Prompting can be used to
predict change in language abilities. This theory was tested later when Olswang and Bain (1996) examined the language production of 21 preschool children. Participants were examined using a dynamic and a static assessment. The dynamic assessment consisted of verbal prompts which were given in a hierarchal order (from minimal to maximal) until a participant was able to produce a two word utterance. Results showed that graduated prompting was correlated with immediate change of language produced. Graduated prompting was also correlated with participants’ success during treatment. This study shows that Graduated Prompting can be used to determine a child’s zone of proximal development and predict how they will respond to intervention.

Graduated Prompting method has also been shown to be useful when assessing vocabulary skills. Larsen and Nippold (2007) used the Graduated Prompting dynamic assessment method to assess morphological analysis as a word learning strategy in typically developing sixth grade children. The Dynamic Assessment Task of Morphological Analysis (DATMA) was used in this study. The participants were asked to define 15 target words. The participants were provided with predetermined verbal prompts (based on morphological analysis) which were presented in a hierarchy (from minimal to maximal) until the participants were able to define the target word. In order to define the target word, some participants needed the word to be broken down into its smaller parts, while others needed the word to be presented in a sentence. The participants were scored based on the level of prompting that they needed in order to elicit the appropriate response. In other words, a child who responded after only one prompt received the most points (in this case, 5), whereas a child who needed the maximal number of prompts (6) in order to define the word would receive 1 point. A child who could not define the target word with maximal prompts received zero points.
Larsen and Nippold (2007) demonstrated that participants needed a varying amount of scaffolding in order to use morphological analysis as a word learning strategy. Participant performance on the Graduated prompting measure was related to word knowledge and reading comprehension scores which suggests that the amount of scaffolding a child needs in order to complete a task may be representative of their skill level (Larsen & Nippold, 2007). Furthermore, the relationship between response to prompting and skill level suggests that dynamic assessment may be a useful tool that could be used for planning interventions (Laing & Kamhi, 2003). For example, information regarding a child’s skill level (as determined through graduated prompting) may be used as baseline data. The level of prompting that the child required could be implemented as an intervention, and the child could be assessed frequently to determine progress. This process of recurrent assessment and intervention mirrors the assessment/intervention method found in the RtI model. Therefore, the information provided by a dynamic assessment could be used to demonstrate how a child responds to intervention.

There are key differences between standardized assessments and dynamic assessments. The major difference between the assessment methods is the incorporation of intervention within the dynamic assessment process. When using a standardized assessment method, intervention is considered as a separate process. Dynamic assessment can be used to inform intervention and to decrease test bias caused by cultural or environmental differences. This is because a child’s success on a dynamic assessment is not measured by their prior knowledge on a particular subject, rather the assessment measures how a child responds after being exposed to a concept (Grigorenko, 2009). As a result, when measuring a complex skill such as vocabulary (which is based on prior knowledge and environmental exposure) it may be useful to implement an assessment procedure which takes influencing factors into account. Also, dynamic assessment
allows for the creation of an authentic reading experience that models word learning in the
classroom environment. Considering the complexities of vocabulary, dynamic assessment may
prove to be a valuable tool in gaining multidimensional information about a child’s word
knowledge. While standardized assessments measure “whether or not” a student knows a word,
a dynamic assessment can be more beneficial by measuring how well the student knows the
word. This could be achieved by presenting vocabulary instructional based prompts.

In summary, Dynamic assessment has shown to be useful when measuring vocabulary
(Pena, et al., 2001). Rather than assessing a child’s particular knowledge of a word, dynamic
assessment allows the examiner to assess a child’s learning process and their responsiveness to
different levels of intervention. This information may be useful within the RtI model by
indicating the amount of support a child needs and informing intervention. Graduated Prompting
has shown to be a valid method of predicting how a child will respond to intervention and
determining their word learning potential (Bain & Olswang, 1995). When compared to
standardized assessments, Graduated Prompting has shown to be more useful in determining a
child’s immediate learning abilities. Further, research reveals a gap in performance between high
and low risk (as determined by income, maternal education and school demographics) students
on standardized vocabulary tests (Burton & Watkins, 2007). Because vocabulary skills predict
reading ability, it is important that children are able to utilize strategies that can help them
understand the unknown words that they encounter when reading a text. Graduated prompting
can be used to examine morphological analysis as a word learning strategy (Larsen & Nippold,
2007). Research needs to show how Graduated Prompting can be used to examine other word
learning strategies.
Word Learning and Context Clues

Words are learned in context (Swanborn & Glopper, 1999) and are not isolated (Nagy & Scott, 2000). For example, Rosenblatt’s transactional theory of reading states that every reading act is a transaction in which the reader engages with the text to make meaning (Rosenblatt, 1994). If words are decontextualized, then a reader is afforded little assistance in making meaning of the words that they encounter. Although many words may have morphological clues that could assist a student with defining a decontextualized term, there are some words which may be completely unfamiliar without the help of surrounding context (i.e. “yacht”, “humid”). Additionally, in the classroom setting, a student would likely be provided with context clues to assist with their overall comprehension of the text as well as with individual words.

Research shows that incidental word learning can occur during natural reading (Swanborn & Glopper, 1999) and that readers can gain partial word knowledge after one exposure to a word in context (Wagovich & Newhoff, 2004). Nagy, Anderson and Herman (1987) examined incidental word learning among third, fifth and seventh grade students. Participants read narrative or expository passages that contained target words without assistance. A week later, the participants were tested on their knowledge of the target words using a multiple choice test. The researchers found that word meanings can be learned from context during normal reading tasks. Additionally, word knowledge that occurs during normal reading tasks is able to be generalized. While there is evidence that it is possible for children to learn words incidentally during the normal reading process (Nagy, et al., 1987), the chance that a child will actually learn the word is small and ranges from 5% to 15% (Nagy et. al, 1987; Swanborn & Glopper, 1999). There are factors which can influence a child’s chances of learning a word from context. One factor is the readability of the text which is determined by using a mathematical
formula such as the Flesch-Kincaid readability test (Flesch, 1948; Kincaid, Fishburne, Rogers & Chissom, 1975). A child will learn fewer words from context if the text is difficult to read or above their grade level (Nagy, et. al, 1987). Another factor that can influence a child’s ability to learn a word during reading is the supportiveness of the text. Beck, McKeown and McCaslin (1983) explain that there are different types of contexts (natural and pedagogical) in which readers are presented with unknown vocabulary. There are four natural reading contexts which can be described along a continuum: misdirective, nondirective, general and directive. A misdirective context is one that leads a reader to gather an incorrect meaning of a target word. A nondirective context provides no guidance in determining the meaning of a target word. A general context provides basic information that may lead a reader in determining a basic category for a target word. And finally, a directive context leads a reader to the appropriate meaning of the target word. The directive context, which is a natural text, is not to be confused with a pedagogical context. The author’s purpose of a pedagogical context is to teach the meaning of unknown words (Beck et al., 1983; Gardner, 2007) whereas in a directive text, the author does not intend to convey word meaning. As children get older, texts are not structured around target words (as in pedagogical texts), and they are expected to gather word meanings from the clues in natural contexts (Beck et al., 1983).

A context clue refers to the information about a target word encountered during reading that assists with the comprehension of the text. Students who excel at decoding are more likely to use context clues to understand a text. A reader that is not able to use context clues may have difficulty decoding and as a result may need assistance in reading comprehension (Barton, 1997). If a child is not able to independently (and incidentally) determine the meaning of an unknown word in a natural context, then Graduated Prompting may be a way of determining the level of
assistance the child needs to comprehend the target word. Using dynamic assessment to direct readers toward context clues has shown to be a useful method for determining their word learning potential. Gardner (2007) used a dynamic assessment (test-teach-retest procedure) to assess the vocabulary skills of 95 elementary students during a reading task. During treatment, participants received a handout which described different types of context clues (forward cues—which occur after the target word in context and backward cues—which occur before the target word in context). The participants completed a post test using the treatment handout. Results showed that readers who received guidance in finding context clues significantly improved their post test scores (Gardner, 2007). In addition to teaching how to find context clues, there are other strategies that have been shown to improve a reader’s ability to derive word meaning from context. One method that has been examined is the SCANR (Jenkins, Matlock & Slocum, 1989), which is an acronym that stands for: Substitute a word or expression for the unknown word; Check to find context clues that support your idea; Ask if the substitution fits all of the context clues, Need a new idea?; and Revise idea to fit the context. The SCANR strategy was examined in a study with 135 fifth grade students. Participants were placed into two groups, one group was instructed on how to derive word meanings from context (SCANR) and the other group was instructed in specific word meanings. Results showed that SCANR participants scored better on the deriving meaning measure than the group who received specific word instruction. Other research shows that practice in deriving word meaning from context may be more effective than teaching specific strategies (Kuhn & Stahl, 1998). Therefore, providing multiple opportunities to derive word meaning from text may improve a reader’s ability to use context clues.

In summary, it is possible to incidentally learn the meaning of unknown words through reading natural contexts (Swanborn & Glopper, 1999). Readers who are unable to identify
context clues independently may be able to do so with assistance (Gardner, 2007). The difference between what a child can do independently and what a child can do with assistance is called the ZPD which can be determined using Graduated Prompting. Graduated prompting can be useful in measuring word learning ability (Larsen & Nippold, 2007), but this method of dynamic assessment has not been researched during a natural reading task or with using context clues as a word learning strategy.

Summary

Although having the ability to read is essential for educational achievement, there are many students who do not become proficient at this skill. One factor that can influence reading comprehension is vocabulary knowledge. Vocabulary is a complex skill that (when assessed using a standardized assessment) is influenced by income and prior experiences. The factors related to income and their influences on vocabulary skill may not become evident until a child reaches the fourth grade which is when children use more strategy knowledge, such as context clues (Kletzien, 1991), to comprehend text (Willson & Rupley, 1997). Because vocabulary knowledge is based on prior knowledge, along with the performance gap between high and low income students on standardized vocabulary tests, studies show that dynamic assessment may be more useful in determining a child’s vocabulary abilities. Dynamic assessments, such as the Graduated Prompting procedure, can be used to determine a child’s response to intervention. In this way, Graduated Prompting can be used to determine a child’s potential to learn new words (Larsen & Nippold, 2007), rather than just their current word knowledge (Watkins & DeThorne, 2000). One way to demonstrate a child’s ability to learn new words is through directing them to use context clues (Gardner, 2007). This strategy, however, has not been observed using Graduated Prompting.
The purpose of the current research study was to assess the word learning potential of low income fourth grade students within a natural reading text by presenting them with graduated prompts directed at context clues. The aims were met by answering the following research questions:

**Research Questions**

1. Is there a significant difference between the scores on the DAVIC graduated prompting phase of participants who scored more than one standard deviation below the mean (Non Pass) and those who did not score more than one standard deviation below the mean (Pass) on the EOWPVT?
2. Is there a significant relationship between EOWPVT standard score and the frequency of responses at each Level of Successful Prompting?
3. Is there a significant relationship between EOWPVT standard scores and DAVIC pretest scores?
4. Is there a significant relationship between EOWPVT standard scores and the DAVIC transfer task scores?
5. Is there a significant relationship between scores on the DAVIC pretest and scores on the DAVIC transfer task?
Chapter III

Method

The purpose of the current research study was to assess the word learning potential of low income fourth grade students within a natural reading text by presenting them with graduated prompts directed at context clues. This study aimed to achieve this by determining the participants’ level of successful prompting. The level of successful prompting is the level (1-5) at which a student can successfully express the meaning of a target word. The performance of two groups (Pass and Non-Pass) on both a static assessment and a dynamic, graduated prompting, assessment were compared.

Setting

This research was conducted at a public school in a southern Ohio county in the middle of the academic year. According to public records, this school serves a primarily low income population, with over 95% of the students receiving free or reduced lunches. This school serves grades Kindergarten through eight. Each child was tested individually in a quiet environment (speech therapy room) in the school.

Participants

The participants included thirty students in the middle of their fourth grade year. The study was described to the school principal, classroom teachers, parents and students. Consent to participate was sent to all students in the fourth grade (approximately 75 students). Both child and parent consent were collected. Parents were asked to report information regarding their child’s age in order to gather demographic data.

Consent and Participant Identification. Each student who consented was alphabetically assigned a number based on their name (e.x. A participant whose last name is “Aaa” was assigned to number 1, while student whose last name is “Aab” was assigned to
number 2. There were 28 participants who participated in both phases of testing. The participants ranged in age from 9 years 3 months to 11 years, 1 month. The mean age of participants was 9;9. There were 19 female participants and 9 male participants. Parents were not asked to report their racial/ethnic backgrounds because this question was not included in the university Institutional Review Board application. However, according to public data, the school population is as follows: 93.3% African American and 6.7% non-African American. Examiner observation suggests a similar ratio among participants in this study.

**Exclusionary criteria.** The Dynamic Assessment of Vocabulary in Context (DAVIC) is designed for use within a school-based Response to Intervention (RtI) model which can provide data to help with deciding if a child will require a full evaluation. Therefore, in order to model the typical RtI process, students who were already receiving special education services under an Individualized Education Program (IEP) were not included in this study. There were no other exclusionary criteria.

**Randomization.** This study aimed to have an even number of participants per group (Pass and Non Pass). This was to be achieved by randomly selecting participants (based on their performance on the Expressive One Word Picture Vocabulary Test ([EOWPVT]; Brownell, 2000) to be administered the DAVIC. However, this step was not needed because an even number of students scored into the two groups.

All students who consented to participate were assessed using the EOWPVT. There was an attrition of one participant who withdrew from the school. Another student was excluded from the study due to being labeled as having a speech-language impairment in the middle of the study. Twenty-eight fourth grade students participated in and completed the study.
Procedure

This study used a two-stage, experimental design. Thirty students were assessed using the EOWPVT. Twenty eight of the thirty participants were administered the DAVIC.

Training. The EOWPVT was administered by two undergraduate students. They were trained on how to administer the EOWPVT during a two hour training session. To ensure fidelity of the DAVIC protocol, the primary investigator (PI) administered the dynamic assessment to all participants.

Power Analysis. Sample size determination was based on detecting a difference between the mean scores on a prior study (Burton and Watkins, 2007). Mean scores and standard deviations between low risk (M=8.08, SD=1.62) and high risk (M=7.17, SD=1.59) groups on a dynamic vocabulary assessment were used to determine statistical power. A mean difference may be detected as significant (two tailed alpha=.05) with 57% power for 12 subjects per group. Results were based on an unpaired t-test.

Standardized Vocabulary Testing. The EOWPVT was administered to thirty students. Due to attrition and exclusionary criteria, the data of twenty-eight participants will be reported. Participants were divided into two groups based on their performance on the EOWPVT. One group consisted of participants who scored no more than one standard deviation below the mean score for the assessment. This group was referred to as the Pass group. The second group consisted of participants who scored more than one standard deviation below mean on this assessment. This group was referred to as the Non-Pass group. Undergraduate students scored the EOWPVT and recorded Pass and Non-Pass group status so that the P.I. would not have this information until after all testing procedures were completed. Based on EOWPVT scores, 14 participants were placed into the Pass group (scored 85 and higher) and 14 participants were
placed into the Non-Pass group (scored below 85). The even number of participants per group occurred naturally, and randomization was not necessary to form the groups.

**Dynamic Assessment.** Twenty-eight Pass and Non-Pass participants were assessed using a dynamic assessment which was created by the P.I. for this study entitled the Dynamic Assessment of Vocabulary in Context (DAVIC). Pass and Non-Pass status information was not available to the P.I. until after the testing was completed. Participants were assessed in a random order. The purpose of the DAVIC is to determine each participant’s level of successful prompting. The level of successful prompting is the point at which a student is able to express the meaning of a word. The level of successful prompting was determined by using a graduated prompting method similar to Larsen and Nippold (2007) which is explained below. The subject matter used in the screening tool is based on the Ohio Academic Content Standards. Grade appropriate academic vocabulary was used as defined by Marzano and Pickering (2005).

**Preparation of text.** The DAVIC is used to examine the level of support that a child needs in order to derive the meaning of unknown words in a given context. The target words in the DAVIC were presented in two expository reading passages: “Properties of Water” and “Changing Rocks”. The reading passages used in the DAVIC were written by the P.I. using grade appropriate academic vocabulary words. All vocabulary words used in the DAVIC passages were selected from a fourth grade science textbook. Each passage contained five target words. The following words were targeted in “Properties of Water”: 1. Cycle, 2. Sleet, 3. Evaporation, 4. Condensation, 5. Precipitation. The following words were targeting in “Changing Rocks”: 1. Weathering, 2. Erosion, 3. Sediment, 4. Sand dune 5. Glacier.

The subject matter of the reading passage was selected based on Ohio academic content standards (Ohio Department of Education, 2002, p. 81) which describe what students should
know and be able to do by grade level. In order to ensure an appropriate reading level, the passages were written using a fourth grade science textbook as a model. Further, the readability of each passage was determined using the Flesch-Kincaid Grade Level Readability Formula (Flesch, 1948; Kincaid et al., 1975) which was calculated to be 4.3 using computer software.

**Coding of Passages.** Four raters read the text to determine the level of supportiveness of each context clue: misdirective, nondirective, general and directive (Beck, McKeown & McCaslin, 1983). For the purpose of this study, each context clue in the reading passage was written to be rated as either general or directive. As stated previously, a general context provides basic information that may lead a reader in determining a basic category for a target word, and a directive context leads a reader to the appropriate meaning of the target word. Using the criteria explained in Appendix E, four raters found each context clue to be either general or directive.

**Testing Procedures.** The administration of the DAVIC occurred in three stages: (1) Pretest, (2) Reading passage and prompting and (3) Transfer task. An outline describing the steps of the DAVIC procedure is located in Appendix G. The DAVIC procedure is described in detail below.

**Phase One: Pretest.** The first stage of the DAVIC is a pretest (Appendix A). The purpose of the pretest was to determine participants’ prior knowledge of each target word meaning. During the pretest, participants were asked to provide the meaning of each of the words that would be encountered in the reading passage. The PI wrote down each response. DAVIC pretest responses were coded and scored after all testing for the study was complete.

**Phase Two: Graduated Prompting Phase.** After pretesting, each participant was presented with the expository text passages one at a time. Half of the participants were randomly chosen to receive the “Properties of Water” passage first while the other half were randomly
chosen to receive the “Changing Rocks” passage first. Participants were instructed to read the passage aloud.

The goal of the DAVIC is to determine how much prompting a child needs in order to derive word meaning. For the purpose of this study, participants needed to be able to decode the passages effectively in order to evaluate their ability to achieve this goal. Therefore, in order to decrease bias caused by a child’s difficulty in decoding, participants were given minimal decoding assistance when they could not decode a word. When a participant demonstrated that they were unable to decode a word (i.e. stopped reading the passage), the word was stated for them one time.

After the participants read the passage, they were presented with a scripted series of graduated prompts in order to determine the level of successful prompting. The conceptual framework of the prompting system used for the DAVIC was based on a graduated prompting system used in the Dynamic Assessment Task of Morphological Analysis (DATMA) (Larsen & Nippold, 2007). Prompts for the DAVIC were selected based on Ohio academic content standards (Ohio Department of Education, 2001) and prior research which have shown the strategies to result in improved in word learning (Jenkins, Matlock & Slocum, 1989; Gardner, 2007; Kuhn & Stahl, 1998). According to the academic content standards fourth grade students in Ohio should be able to use context clues to determine the meaning of unknown vocabulary items (Ohio Department of Education, 2001, p. 61-62). The research based strategies have been modified for the purposes of this study. The strategies listed below were presented in a script format (Appendix B).

2. Strategy Two: Practice in deriving meaning (Kuhn & Stahl, 1998). Direct the child to the word by pointing and allow them to independently derive meaning.


5. Strategy Five: Modified version of SCANR (Jenkins, Matlock & Slocum, 1989). Ask student to substitute a word for the unknown word to determine if it makes sense in the context.

At each level, the participant was asked to provide a definition of the term based on what they found in the context. When a correct response was not elicited, and word meaning was not determined, the definition was provided and the next item was administered. A response during this phase of the DAVIC was deemed “correct” when it met the “Established” requirements as defined in Appendix D. Therefore, during the prompting phase, all participants had to demonstrate an Established Level of knowledge prior to completing a transfer task for that word. Participants who did not answer with an Established response for a prompt were provided with graduated prompts until this occurred. If an established response was not elicited, then they were given a score of 0 (No Level of Successful Prompting [NLSP]) and did not complete a transfer task for that item. The reading passage and prompting phase were audio recorded.

**Phase Three: Transfer Task.** As stated previously, within Graduated Prompting, learning ability is measured by transfer, which is the distance between the child’s ability to perform the original task and their application of concepts to new tasks. This distance can be measured using a scale which ranges from no transfer, up to very far transfer (Gutierrez-Clellen
When the participant provided a correct response for a word in the Graduated Prompting Phase, they moved on to Phase Three (transfer task) for that particular item. During the transfer task, participants were asked to use the target word in a sentence. After the transfer task was completed for the target word, the next word was presented within the Graduated Prompting phase. This process (graduated prompting phase then transfer task) was repeated for each word until all ten words were complete. In other words, the transfer task was elicited within the Graduated Prompting phase, but only occurred after the participant provided a correct response for the word. DAVIC transfer task responses were coded and scored after all testing for the study was complete.

**Coding for the DAVIC pretest and the DAVIC transfer task.** After all data was collected, the pretest and transfer task responses were coded for word knowledge based on levels defined by Beck, McKeown, McCaslin and Burkes (1979). These levels are: Unknown, Acquainted, and Established. A DAVIC pretest response was coded as “Unknown” if it had no relation to the stated definition of the term and resembled “Unknown” responses as outlined in Appendix D. A response was coded as “Acquainted” if it displayed a basic understanding of the word and resembled example “Acquainted” responses as outlined in Appendix D. A response was coded as “Established” if the word was immediately recognized and was used appropriately and resembled example “Established” responses as outlined in Appendix D. Coding for the DAVIC transfer task (sentence production) was similar to that of the DAVIC pretest (definition production). However, because sentence productions may not contain information commonly found in definitions (e.g. “We may get a lot of precipitation this week.”), syntactic along with semantic knowledge was considered when determining level of word knowledge for the transfer task. The criteria for coding the sentence productions can be found in Appendix F.
Participants could score up to 3 points for each item for a possible total of 30 on both the pretest and the transfer task. If the participant did not complete and item on the pretest or the transfer task (e.g. correct definition not given during the graduated prompting phase) then this occurrence was automatically coded as Unknown. Responses coded as Unknown, Acquainted and Established, were scored as 1, 2 or 3 respectively. Therefore, the range of scores for both tests was 10-30. Scores from the transfer task were compared to scores from the pretest to determine the level of transfer.

**Agreement.** The Principal Investigator (PI) coded all DAVIC pretest and DAVIC transfer task responses. Ten percent of the responses were coded by a trained Research Assistant (RA) to determine a percentage of agreement. The target for this study was 90% agreement between the PI and the RA for both the DAVIC pretest and the DAVIC transfer task. The agreement procedure occurred in three steps: 1.) The RA was trained on the coding key (Appendices D and F). 2.) The RA practiced coding 10% of the pretest responses. The RA was given feedback regarding the codes by the PI. 3.) The RA coded another 10% of the sample to determine the percentage of agreement. This procedure was completed to determine the percentage of agreement for both the pretest and the transfer task. Agreement for the pretest was calculated at 97%. Agreement for the transfer task was calculated at 90%.

**Scoring of the Graduated Prompting Phase.** The scoring system for the DAVIC Graduated Prompting Phase is similar to the scoring system found in the Larsen and Nippold (2007) study and was based on the participants’ level of successful prompting. There are ten target words on the DAVIC for a possible total of 50 points. The system (see Appendix C) is described below: Scores (1-5) were inversely given based on the level of successful prompting (1-5). In other words, a participant who was successful at Level 5 received a score of 1, while a
participant who was successful at Level 1 received a score of 5. With a total score of 50 (possible 5 points each for 10 words), a higher score represented a participant that did not require many prompts, while a lower score represented a participant who required several prompts to derive word meaning. Therefore, the range of possible scores on the DAVIC Graduated Prompting Phase was 10-50 (or 0-100% after being converted to percentages.

**Data Collection and Timeline.** Each participant was assigned a subject number and a folder which included all of their test forms including: signed consent, demographic data form, DAVIC, EOWPVT, and score form. Total testing time for this study was two months. Each student was tested on 2 different occasions. On the first session, they were administered the EOWPVT. On the second session, they were administered the DAVIC.

**Data Analysis**

In order to determine the influence of graduated prompting on participant word learning, possible influencing variables had to be ruled out as contributing factors. Analyses were performed to determine whether variables such as gender, type of context clue (forward and backward), supportiveness of text (directive and general) and order of target word presentation (based on which passage the participant received first) contributed to the performance of students on the assessments. The analyses performed included Chi-square Tests for Independence and Independent measures t-tests. The results of the above analyses are described in the Results section.

This study addressed five research questions:

Question One. Is there a significant difference between the scores on the DAVIC graduated prompting phase of participants who scored more than one standard deviation below the
mean (Non Pass) and those who did not score more than one standard deviation below the mean (Pass) on the EOWPVT?

The purpose of this question was to determine if there was an overall difference between the two groups and to examine how they performed on the DAVIC Graduated Prompting Phase. In order to answer this question, standard scores from the EOWPVT were coded to determine the two groups: Pass (scored no more than one standard deviation below mean of 100) and Non-Pass (scored more than one standard deviation below mean of 100). The DAVIC Graduated prompting scores could range from 10-50 (Ten target words worth a possible 1-5 points each). Scores on the DAVIC graduated prompting phase (along with the pretest and transfer task) were converted to a percentage (50=100%) so that the scores from all three phases on the DAVIC would be on a similar scale. An independent measures t-test was used to determine if there was a significant difference between the percentage scores of the two groups on the DAVIC Graduated Prompting Phase.

H0: There is no significant difference between the population means.

Question Two. Is there a significant relationship between EOWPVT standard score and the frequency of responses at each Level of Successful Prompting?

The purpose of this question was to determine if there was a relationship between the participants’ scores on the EOWPVT and their frequency of responses at each level of prompting. This was achieved by first finding each participant’s frequency of occurrence for the prompt categories. For analyses, the prompt categories were coded as: LSP 1 (Prompt 1), LSP 2 (Prompt 2), LSP 3 (Prompt 3), LSP 4 (Prompt 4), LSP 5 (Prompt 5), and NLSP (No level of successful prompting reached ). For example, out of the ten target words, one participant reached Prompt 1 for one word, Prompt 2 for eight words, and Prompt 3 for one word. This
question aimed to demonstrate whether a relationship occurred between how the participant performed on the EOWPVT, and how frequently they responded at each prompt category. Six separate Pearson Correlations were used to compare the standard scores on the EOWPVT and the frequency of occurrence at each prompt level per participant. One analysis could not be used to analyze the data because the participants did not have an equal opportunity to reach each level. For example, a participant who reached a level 1 prompt for a word would no longer have the opportunity to reach the level 5 prompt. This is because after a level of successful prompting is reached for a word, no further prompting is necessary and the next target item is presented. As a result of this, six separate Pearson Correlations were used to show the relationship between EOWPVT and how frequently a participant reached each level. It was assumed that participants who scored lower on the EOWPVT would require more prompts and would respond more frequently to higher level prompts. It was similarly assumed that a participant who scored higher on the EOWPVT would require fewer prompts and would therefore respond more frequently to the lower level prompts.

H₀: There is no relationship between EOWPVT scores and the frequency of responses at LSP 1, LSP 2, LSP 3, LSP 4, LSP 5 and NLSP.

Question Three: Is there a significant relationship between EOWPVT standard scores and DAVIC pretest scores?

The purpose of this question was to determine if there was a relationship between how participants performed on the standardized assessment, and how they performed on the DAVIC without Graduated Prompting assistance. The DAVIC pretest score was determined by coding the responses as Unknown, Acquainted and Established as outlined in Appendix D. The highest score possible was 30. Because participants were given a score of 1 for Unknown responses, the
lowest possible score was 10. Therefore the range of scores for the DAVIC pretest was 10-30. This number was converted into a percentage (30=100%). The standard scores on the EOWPVT were compared to the percentage score. The Pearson Correlation was used to determine if a significant relationship existed between the two variables.

In order to determine if there was a group difference on the pretest, a sub-question was addressed: Is there a significant difference between Pass and Non Pass participants on the DAVIC pretest? The purpose of this sub-question was to determine if there were group (Pass and Non Pass) differences in performance on the DAVIC pretest. Pass and Non Pass groups were defined by using participant standard scores from the EOWPVT. An independent measures t-test was used to compare the percentage scores of the two groups.

H0: There is no relationship between performance on the EOWPVT and the DAVIC pretest score.

H0: There is no difference between performance of Pass and Non pass groups on the DAVIC pretest.

Question Four: Is there a significant relationship between EOWPVT standard scores and the DAVIC transfer task scores?

The purpose of this question was to determine if there was a relationship between how participants performed on the EOWPVT and how they performed on the DAVIC after Graduated Prompting assistance. The DAVIC transfer task score was determined by coding the responses as Unknown, Acquainted and Established as outlined in Appendix F. The highest score possible was 30. This number was converted into a percentage (30=100%). The standard scores of the EOWPVT were compared to the percentage score on the DAVIC. The Pearson Correlation was used to determine if a significant relationship existed between the two variables.
In order to determine if there was a group difference on the transfer task, a sub-question was addressed: Is there a significant difference between performance of Pass and Non Pass participants on the DAVIC transfer task? The purpose of this sub-question was to determine if there were group (Pass and Non Pass) differences in performance on the DAVIC transfer task. Pass and Non Pass groups were defined by using participant standard scores from the EOWPVT. An independent measures t-test was used to compare the percentage scores of the two groups on the DAVIC transfer task.

H0: There is no relationship between performance on the EOWPVT and the DAVIC transfer task.

H0: There is no difference between performance of the Pass and Non pass groups on the DAVIC transfer task.

Question Five. Is there a significant relationship between scores on the DAVIC pretest and scores on the DAVIC transfer task?

The purpose of this question was to determine if a significant pre-test to post test improvement was observed among participants after completing a Graduated Prompting task. DAVIC Pretest data and DAVIC transfer task (post test) responses were coded as Unknown, Acquainted, and Established, and were scored as 1, 2, and 3 respectively. The responses for the DAVIC pretest were coded based on the criteria outlined in Appendix D. Responses for the DAVIC transfer task were coded based on the criteria outlined in Appendix F. A total of 30 points were possible for both the DAVIC pretest and the DAVIC transfer task. The scores from the DAVIC pretest and DAVIC transfer task were converted to a percentage (30=100%). A paired samples t-test was used to determine if there was a significant difference between the two scores. Cohen’s D was computed in order to determine effect size.
H₀: There is no significant difference between DAVIC pretest scores on the DAVIC and DAVIC transfer task scores.
Chapter IV

Results

The purpose of this research study was to assess the word learning potential of low income fourth grade students within a natural reading text by presenting them with graduated prompts directed at context clues. There were potential confounding variables (supportiveness of text, type of cue, gender and order of passage presentation) in this study. Therefore, a series of analyses were performed in order to examine whether the variables contributed to the outcomes. A chi-square test of independence was performed to examine the relationship between supportiveness of text (general or directive) and frequency of responses at each Level of Successful Promting (LSP). The relationship between these variables was not significant, \( \chi^2 = 7.347, df = 5, p < .196 \). Therefore, the supportiveness of text (general and directive context clues) did not contribute to the frequency of responses at each prompt level. A chi-square test of independence was performed to examine the relationship between type of cue (forward or backward) and frequency of responses at each prompt level. The relationship between these variables was not significant, \( \chi^2 = 7.869, df = 5, p < .164 \). Therefore, the type of cue (forward or backward) did not contribute to frequency of responses at each prompt level. An independent measures t test showed that performance on the Dynamic Assessment of Vocabulary in Context (DAVIC) of female participants (\( M = 85.05, SD = 15.42 \)) and male participants (\( M = 86.33, SD = 6.91 \)) did not differ significantly \( t(26) = -2.236, p = .815, d = -0.11 \). An independent measures t test showed that there was no significant difference \( t(26) = -0.092, p = .927, d = 0.03 \) between performance on the DAVIC of participants who received the “Properties of Water” passage first (\( M = 66.00, SD = 14.23 \)) and participants who received the “Changing Rocks” passage first (\( M = 65.57, SD = 10.08 \)). The above analyses suggest that changes observed in this study are the
result of the graduated prompting procedure and are not due to other potentially confounding factors.

This study addressed six research questions which are focused on participant performance on the DAVIC and the Expressive One Word Picture Vocabulary Test (EOWPVT). Table 1 shows participant descriptive statistics on the two measures administered.

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Pass Group (n=14)</th>
<th>Non-Pass Group (n=14)</th>
<th>Total Sample (n=28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOWPVT</td>
<td>95.5</td>
<td>75.42</td>
<td>85.46</td>
</tr>
<tr>
<td>SD</td>
<td>8.93</td>
<td>7.92</td>
<td>13.15</td>
</tr>
<tr>
<td>Range</td>
<td>85-112</td>
<td>63-84</td>
<td>63-112</td>
</tr>
<tr>
<td>DAVIC Pretest</td>
<td>40</td>
<td>35.9</td>
<td>37.96</td>
</tr>
<tr>
<td>SD</td>
<td>6.83</td>
<td>3.97</td>
<td>5.87</td>
</tr>
<tr>
<td>Range</td>
<td>33-53</td>
<td>33-47</td>
<td>33-53</td>
</tr>
<tr>
<td>DAVIC Graduated Prompting Phase</td>
<td>67</td>
<td>64.57</td>
<td>65.79</td>
</tr>
<tr>
<td>SD</td>
<td>14.5</td>
<td>9.52</td>
<td>12.1</td>
</tr>
<tr>
<td>Range</td>
<td>34-90</td>
<td>50-82</td>
<td>34-90</td>
</tr>
<tr>
<td>DAVIC Transfer Task</td>
<td>63.21</td>
<td>61.42</td>
<td>62.32</td>
</tr>
<tr>
<td>SD</td>
<td>11.07</td>
<td>10.46</td>
<td>10.61</td>
</tr>
<tr>
<td>Range</td>
<td>47-93</td>
<td>33-77</td>
<td>33-93</td>
</tr>
</tbody>
</table>

Research Question 1: Is there a significant difference between the scores on the DAVIC graduated prompting phase of participants who scored more than one standard deviation below the mean (Non Pass) and those who did not score more than one standard deviation below the mean (Pass) on the EOWPVT? The Levene’s Test was used to test whether the two groups have equal variances on the DAVIC. The result (F=2.308, p=.141) suggests that the two groups have equal variances so the assumption is met. An independent measures t test showed that there was no significant difference between students who scored more than one standard deviation above the mean (Pass group) ($M=67.00, SD=14.502$) and students who scored more than one standard deviation below the mean the Non Pass group ($M=64.57, SD=9.525$) on the DAVIC graduated
prompting task, $t(26)=.524, p=.605, d=.19$. This means that there was no relationship between the participants’ level of word knowledge (as measured by the EOWPVT) and their word learning ability (as measured by the DAVIC Graduated Prompting Phase). Because the DAVIC Graduated Prompting phase was scored based on the levels of prompts needed to derive word meaning, this question shows that there is no relationship between the amount of prompts needed to derive word meaning, and participant EOWPVT score.

Research Question 2: Is there a significant relationship between EOWPVT standard score and the frequency of responses at each Level of Successful Prompting? A Pearson Correlation was used to compare the standard scores on the EOWPVT and the frequency of responses at each at each level of successful prompting per participant. The analysis revealed the following correlations: EOWPVT scores and Prompt 1 frequency were significantly and positively correlated, $r=.429, (p=0.05)$. This means that participants who scored higher on the EOWPVT responded more frequently at the level one prompt (incidental word learning). The correlation between EOWPVT score and Prompt 2 (context only) frequency was not significant, $r=.123 (p=.535)$. This suggests that participants responded similarly at the level two prompt, despite their performance on the EOWPVT. The EOWPVT score and Prompt 3 (direct instruction) frequency were significantly correlated, $r=-.376, (p=0.05)$. The negative correlation suggests that participants who scored higher on the EOWPVT responded less frequently to the level three prompt. Likewise, this negative correlation indicates that students who scored lower on the EOWPVT responded more frequently to the direct instruction which occurred at the level three prompt. The correlation between EOWPVT score and Prompt 4 (modeling) frequency was not significant, $r=-.038 (p=.850)$. This suggests that participants responded similarly at the level four prompt, despite their performance on the EOWPVT. The correlation between EOWPVT
scores and Prompt 5 (substitution) frequency was not significant, \( r = 0.008 \) (\( p = 0.968 \)). This suggests that participants responded similarly at the level five prompt, despite their performance on the EOWPVT. The correlation between EOWPVT score and NLSP (no level of successful prompting) frequency was not significant \( r = -0.335 \) (\( p = 0.082 \)). This suggests that the inability to derive word meaning was not related to the students’ performance on a standardized vocabulary test. Table 2 displays the correlations at each prompt level.

Table 2: Correlations Between EOWPVT standard scores and frequency of responses at each prompt level.

<table>
<thead>
<tr>
<th>EOWPVT Score</th>
<th>Prompt 1</th>
<th>Prompt 2</th>
<th>Prompt 3</th>
<th>Prompt 4</th>
<th>Prompt 5</th>
<th>NLSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>EOWPVT Score</td>
<td>1</td>
<td>0.429*</td>
<td>0.123</td>
<td>0.376*</td>
<td>-0.038</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.429*</td>
<td>1</td>
<td>0.123</td>
<td>0.376*</td>
<td>-0.038</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.123</td>
<td>0.123</td>
<td>1</td>
<td>0.376*</td>
<td>-0.038</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>0.376*</td>
<td>0.376*</td>
<td>0.376*</td>
<td>1</td>
<td>-0.038</td>
<td>0.008</td>
</tr>
<tr>
<td></td>
<td>-0.038</td>
<td>-0.038</td>
<td>-0.038</td>
<td>-0.038</td>
<td>1</td>
<td>-0.335</td>
</tr>
<tr>
<td></td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>0.008</td>
<td>-0.335</td>
<td>1</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed)

Research Question 3: Is there a significant relationship between EOWPVT standard scores and DAVIC pretest scores? A Pearson Correlation was used to examine the relationship between EOWPVT scores and DAVIC pretest scores. The Pearson Correlation showed that scores on EOWPVT and DAVIC pretest scores were significantly correlated, \( r = 0.599, p = 0.01 \). This positive correlation means that students who scored higher on the EOWPVT also scored higher on the pretest. Both the pretest (students defined target words) and the EOWPVT assessed the participants’ vocabulary knowledge without any additional supports. Therefore, the results suggest that the two assessments may be similar measures of vocabulary knowledge.

Sub-question: Is there a significant difference between scores of the Pass and Non Pass groups on the DAVIC pretest? An independent measures t test revealed that there was no significant difference between scores of the Pass group (\( M = 40, SD = 6.84 \)) and scores of
the Non Pass group (M=35.9, SD=3.97) on the DAVIC pretest, t(26)=1.93, p=.065, d=.73. This means that despite the positive correlation between the scores, there was no difference between Pass and Non Pass groups on the DAVIC pretest. The positive correlation between the EOWPVT scores and pretest scores along with the lack of a significant relationship between groups suggests that the Pass and Non Pass groups may have been too similar to detect a difference.

Research Question 4: Is there a significant relationship between EOWPVT standard scores and the DAVIC transfer task scores? The Pearson Correlation showed that there was not a significant correlation r=.184 p=.184, between scores on the EOWPVT and DAVIC transfer task scores. The EOWPVT is a measure of participant vocabulary knowledge without any scaffolding while the transfer task is a measure of vocabulary knowledge after participants have received prompting. The lack of a significant relationship between standard scores and the DAVIC transfer task shows that there is no relationship between a child’s vocabulary knowledge without assistance and a child’s vocabulary knowledge after prompting.

Sub-question: Is there a significant difference between scores of Pass and Non Pass participants on the DAVIC transfer task? An independent measures t test revealed that there was no significant difference between scores of the Pass group (M=63.21, SD=11.07) and scores of the Non Pass group (M=61.43, SD=10.46) on the DAVIC transfer task, t(26)=.439, p=.665, d=.17. The lack of a significant difference between Pass and Non-Pass groups may be the result of group similarities (Pass=85 and higher; Non-Pass=84 and lower). However, the lack of a significant correlation between the DAVIC transfer task and EOWPVT scores, along with the lack of a significant difference
between the Pass and Non Pass groups on the transfer task may suggest that performance on the EOWPVT is not related to how a child can perform after they receive assistance.

Research Question 5: Is there a significant relationship between scores on the DAVIC pretest and scores on the DAVIC transfer task? A paired samples t test showed that the participants scored significantly higher on the transfer task \( (M=62.32, SD=10.61) \) than on the pretest \((M=37.96, SD=5.87)\); \( t=-10.73, p<.001, d=2.84 \). This reveals that participant scores improved significantly from the pretest to the transfer task, with a large effect found for the graduated prompting procedure. The DAVIC pretest and the DAVIC transfer task are both measures of vocabulary knowledge. The improvement in scores suggests that the interventions that occurred during the graduated prompting phase are effective at improving vocabulary knowledge.
Discussion

The purpose of this study was to use graduated prompting to examine the word learning potential of low income fourth grade students during an authentic reading task using context clues as a word learning strategy. This research sought to determine if vocabulary knowledge and word learning potential could be assessed within a graduated prompting procedure. The purpose was also to determine if word knowledge would improve as a result of graduated prompting and to measure students’ level of independence in deriving word meaning from context. Finally, the purpose of this study was to determine if there is a relationship between word knowledge (as measured by a standardized assessment) and word learning potential (as measured by graduated prompting. This chapter will address the following sections: 1) vocabulary knowledge and word learning ability, 2) unknown, acquainted and established, 3) congruence with prior research, 4) teaching a word learning strategy, 5) future research, 6) limitations, and 7) implications and conclusion.

Vocabulary Knowledge and Word Learning Ability

There is a distinction between vocabulary knowledge and word learning potential. Vocabulary knowledge is the information that one has about a word based on their prior experiences with the term. The complexity of vocabulary knowledge can be revealed within levels: Unknown, Acquainted and Established (Beck et. al, 1979). Word learning potential can be defined as one's ability to derive word meaning. Researchers have suggested that dynamic assessment may be a useful tool in examining word learning potential (Watkins & DeThorne, 2000). This study provides evidence supporting this concept.

Relationship between vocabulary knowledge and word learning potential. A major finding in this study was that there was no relationship between vocabulary knowledge (as
shown by group performance on the standardized test) and word learning potential (as measured by scores on the Dynamic Assessment of Vocabulary in Context [DAVIC] Graduated Prompting Phase). The implication is that using a standardized assessment alone to examine vocabulary skills does not provide information regarding a child's learning abilities. For example, after prompting, one Non-Pass (scored more than one standard deviation below the mean) participant who scored a 66 on the Expressive One Word Picture Vocabulary Test (EOWPVT) was able to demonstrate both Acquainted and Established responses (i.e. Acquainted-When we went to the beach we seen a glacier floating in the sea.”; Established- “When we went to the beach we built a sand castle and the water knocked it down and there were wet sand dunes on the beach.”). Therefore, assessment procedures, such as those used to identify students with disabilities, should include some form of scaffolding to determine how a child responds to intervention.

Conflicts with prior research. The absence of a relationship between the DAVIC and EOWPVT scores conflicts with prior research (Larsen and Nippold, 2007). Larsen and Nippold (2007) administered the Dynamic Assessment Task of Morphological Analysis (DATMA) which is a measure designed to evaluate a child’s ability to use morphological analysis as a word learning strategy. The DATMA was found to be related to scores on a standardized vocabulary measure (PPVT-III). In other words, word knowledge was determined to be related to a child’s ability to use morphological analysis. A major assumption that could be made here is that similarities in groups in the current study confounded the findings. However, the significant correlation found between the DAVIC pretest and EOWPVT scores, along with the lack of a significant correlation between the DAVIC transfer task and EOWPVT support the finding that there is no relationship between vocabulary knowledge and word learning ability. Although the
group similarities did not contribute to the conflicting findings, there are factors that may have contributed to the differences.

First, the differences in testing populations may have contributed to the varying results. Larsen and Nippold assessed predominantly White American students in the mid-SES range while the population of the current study included predominantly African American, low income students. Students in the current study were more likely to encounter standardized test bias because of factors caused by income and cultural differences. This is illustrated by the mean standard scores on the assessments administered in each study. In the DATMA study, the mean standard score on the PPVT-III was 112.72 which is much higher than the mean of EOWPVT standard scores (M=85.5) found in the DAVIC study. This number (85.5) represents the average score of the participants (n=28) who participated in both phases of the study. It is noted that the two participants who were removed from the DAVIC study were both Non-Pass (scored below 85) participants, which would therefore cause the average to be lower. This means that the students in the current study started with poorer word knowledge (measured by the EOWPVT which was significantly correlated with the DAVIC pretest) and ultimately had room for growth to occur.

Additionally, one of the purposes of this study was to measure improvement in word learning which was why a pretest and a transfer task were included in this study. This was not the purpose of the Larsen and Nippold (2007) study therefore word learning improvement (by included a pretest and a post test) was not assessed. Rather, the study examined how students performed on the graduated prompting measure and how this related to the standardized assessment. Also, the current study examined context clues while Larsen and Nippold (2007) examined morphological analysis. All of these factors may have contributed to the differences
in findings. The finding of the current study that there is no relationship between vocabulary knowledge and word learning potential is therefore maintained.

**Assessing vocabulary knowledge and word learning potential.** This study showed that graduated prompting can be used to examine both vocabulary knowledge and word learning potential. The information obtained from the EOWPVT could not be used to specify the participants’ level of word knowledge (i.e., Unknown, Acquainted, Established). The data obtained from the EOWPVT showed only whether the word was known or not. Both the DAVIC Pretest and the DAVIC Transfer Task, however, provided specific information pertaining to the participants’ level of word knowledge. Also, through the graduated prompting procedure, information regarding participant word learning potential was revealed.

**DAVIC Pretest and EOWPVT.** Vocabulary knowledge was assessed with the EOWPVT, the DAVIC pretest (define the target word), and the transfer task (use the target word in a sentence). The results showed that there was a relationship between EOWPVT scores and the DAVIC pretest (the higher the EOWPVT score, the higher the pretest score) which may be the result of both assessments examining vocabulary knowledge without context and without scaffolding support. Further, the pretest, like the EOWPVT, examined the participants' prior experiences with each target word. It is noted, however, that despite the significant correlation between EOWPVT and DAVIC pretest scores, there was no significant difference noted between groups on the DAVIC pretest. This may be a result of the criteria used to define group status (84 and below=Non Pass, 85 and higher=Pass). The marginal one point difference between group status on the EOWPVT may have prevented group differences (Pass, M=95.5; Non-Pass, M=75.42) from being detected which subsequently is a limitation of this study. Figure 1 illustrates the spread of EOWPVT and DAVIC Pretest scores.
Pass and Non Pass Similarities. Group status in the current study (Pass and Non-Pass) was based on scores from the EOWPVT. Students who scored 85 and above were placed in the Pass Group while students who scored below 85 were placed in the Non-Pass group. The overall range of scores on the DAVIC for the Pass group was 85-112. The overall range of scores (found in Table 1) on the DAVIC for the Non Pass group was 63-84. Both groups included borderline participants (students who scored between 84 and 86 on the EOWPVT). This meant that there
was only a marginal difference of one point between some students in the Pass Group and in the Non Pass group. The marginal difference between Pass and Non Pass groups may not be clinically relevant, however, because performance on standardized assessments is based on prior knowledge, while performance on the DAVIC is not. The range of EOWPVT scores by group is illustrated in Figure 2. The t-tests showed that there was no significant difference between the Pass and Non-Pass group on the Pretest, the Graduated Prompting Phase or the Transfer Task of the DAVIC.

Figure 2: Expressive One Word Picture Vocabulary Test Standard Scores by Pass, Non Pass Group status

**DAVIC Transfer Task and EOWPVT.** There was no relationship found between the EOWPVT scores and the transfer task suggesting that after experience with the word in context and scaffolding support, children of high and low vocabulary knowledge (displayed by Pass and Non Pass group status) can perform comparably. Figure 3 illustrates the scores on the EOWPVT and the DAVIC transfer task. The significant correlation found between the EOWPVT and the DAVIC pretest along with the lack of a significant correlation found between the EOWPVT and the DAVIC transfer task suggests that the comparable performance between the Pass and Non
Pass Groups on the transfer task is the result of the Graduated Prompting procedure. This further supports the finding that vocabulary knowledge is not related to word learning ability.

In summary, the DAVIC framework (pretest, graduated prompting, transfer task) was more useful than the EOWPVT in examining the overall complexities of vocabulary knowledge and word learning. During the DAVIC pretest, baseline information regarding each participant’s word knowledge was collected. During the DAVIC prompting phase (where students were asked to define the target words), information was collected regarding how much scaffolding was
required to derive the word meaning. Finally, student progress was measured during the DAVIC transfer task.

**Unknown, Acquainted and Established**

Using the “Unknown, Acquainted, Established” coding system, several word learning patterns were observed among the participants within the DAVIC pretest and DAVIC transfer tasks. During the DAVIC pretest (Definition) portion of the assessment, most participant answers (89%) displayed an Unknown level of knowledge for the target words (i.e. Cycle-“A person that you just be, act like.”). Few answers (8%) displayed an Acquainted level of word knowledge for the target words (i.e. Glacier-“Big ice.”) and even fewer (3%) demonstrated an Established level of word knowledge for the target words (i.e. Cycle-“Something that repeats.”). Therefore, it was judged that in general, the participants demonstrated an Unknown level of word knowledge for the target words prior to being exposed to the terms in context. Following exposure in context and with scaffolding provided through Graduated Prompting, most participant productions (55%) demonstrated an Acquainted level of knowledge (i.e “I see evaporation when my mom makes tea.”). Additionally, more productions were found to exhibit an Established level of word knowledge (i.e. “The story has the word evaporation which mean to turn liquid into gas.”) after Graduated Prompting (16%) than before Graduated Prompting (3%). Finally, the number of productions that showed an Unknown level of word knowledge (i.e. “If it erosion I wouldn't know what to do.”) greatly decreased from 89% to 29%. The statistics suggest that the significant difference that was found between the DAVIC pretest to the DAVIC Transfer task may be attributed to the increase in Acquainted and Established Responses, paired with the decrease in Unknown responses. The DAVIC graduated prompting phase had a large effect (d=2.84) on the observed improvement. During the DAVIC, which is an instructional/assessment
method, direct instruction in deriving word meaning was provided which resulted in an improvement of word knowledge. DAVIC framework (pretest, prompting, transfer task) could be used to assess and provide intervention in a systematic manner within RtI models.

**Development of Sentence Coding Rules.** The codes created for the Definition productions elicited during the pretest (Appendix D) were not able to be applied to Sentence productions elicited during the transfer task. It was determined that the requirements for producing a sentence were quite different than the requirements for producing a definition. For example, according to Appendix D, in order to demonstrate an established level of knowledge of "precipitation", the participant had to mention rain, sleet or snow. However, many students used the word appropriately in a sentence but made no reference to an actual form of precipitation. For example: "In the winter precipitation happens”; and “There was a lot of precipitation during December and January.” These productions are acceptable uses of the term. As a result, rules (Appendix F) were created in order to code sentence productions. First it was determined that syntactic appropriateness would be considered when determining level of word knowledge. The inclusion of this concept is supported by the lexical quality hypothesis which states that the semantic constituent of a word includes word meaning in addition to grammatical information (Perfetti & Hart, 2002).

Definitions were created for Semantic Knowledge and Syntactic Knowledge:

*Semantic knowledge*-A sentence production will be described as demonstrating semantic knowledge when it references information specified in the Definition Coding Key. *Syntactic knowledge*-A sentence production will be described as demonstrating syntactic knowledge when the integrity of the word form is maintained, or the word form is changed and used in a semantically and syntactically acceptable manner.
Next, conditions were created to determine when each level (Unknown, Acquainted, Established) was met. Therefore, under the new conditions, the sentence: “There was a lot of precipitation during December and January.”, would satisfy the “Acquainted” rule under the justification that the sentence displays appropriate syntactic use. Table 3 provides examples of sentence productions coded in each category. A thorough explanation of each level (Unknown, Acquainted and Established) can be found in Appendix F, however in general: a production was deemed Unknown when it displayed no knowledge of the target word; a production was deemed Acquainted when it demonstrated either semantic or syntactic knowledge; and a production was deemed Established when it demonstrated both semantic and syntactic knowledge.

This information may be useful in the educational setting because it illustrates the importance of practical application of curricular concepts. This study shows that children are able to show that they have knowledge about a word without providing a specific definition. Also, if a child is able to use a word appropriately, then they are demonstrating that they have at least an Acquainted level of word knowledge which may be applied in the classroom setting. The practical application of vocabulary assessed with the DAVIC is not measured on the EOWPVT. Therefore, a graduated prompting framework (such as the DAVIC), could be used in RtI models in order to provide a thorough description of a child’s vocabulary and word learning abilities.

Table 3: Examples of participant responses and sentence coding for the transfer task at each level of word knowledge (Unknown, Acquainted and Established).

<table>
<thead>
<tr>
<th>Target Word</th>
<th>Unknown</th>
<th>Acquainted</th>
<th>Established</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle</td>
<td>Today I learned what the word cycle means.</td>
<td>A life cycle has steps.</td>
<td>A water cycle happens over and over again.</td>
</tr>
<tr>
<td>Sleet</td>
<td>I sleet ice cream into the freezer.</td>
<td>I was outside and it was sleet into cold ice.</td>
<td>Sleet is when rain turn into ice.</td>
</tr>
<tr>
<td>Evaporation</td>
<td>In a rap song somebody used evaporation.</td>
<td>I saw the water evaporate.</td>
<td>Evaporation is liquid turning into gas.</td>
</tr>
<tr>
<td>Condensation</td>
<td>I saw condensation at my friend house.</td>
<td>I saw the condensation on the pot.</td>
<td>Condensation is gas that turns into water.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
<td>Example</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------</td>
<td>----------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Precipitation</td>
<td>I don't know</td>
<td>We may get a lot of precipitation this week.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The weathering is telling us what the weather means on the news.</td>
<td>Precipitation is when rain and snow fall out of the clouds.</td>
<td></td>
</tr>
<tr>
<td>Weathering</td>
<td>The weathering is telling us what the weather means on the news.</td>
<td>The rocks were breaking down because it was weathering outside.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The rocks were breaking down because it was weathering outside.</td>
<td>Rocks that break down into pieces is called weathering.</td>
<td></td>
</tr>
<tr>
<td>Erosion</td>
<td>I erosion pieces of glass today at home.</td>
<td>Erosion makes sediment crack rocks.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The wind and water caused sediment.</td>
<td>No established responses for the target word</td>
<td></td>
</tr>
<tr>
<td>Sediment</td>
<td>I have never seen sediment before.</td>
<td>Sediment is water and wind that makes rock move from one place to another.</td>
<td></td>
</tr>
<tr>
<td>Sand dune</td>
<td>I don't know</td>
<td>Sand dune is a pile of sand.</td>
<td></td>
</tr>
<tr>
<td>Glacier</td>
<td>I see a glacier on TV.</td>
<td>Glaciers are ice that moves around in the sea.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>When I go to the sea I see glaciers.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Congruence with prior research**

**Standardized vocabulary tests.** Many of the findings in this study are consistent with prior research. For example, several studies show that low income children perform poorly on standardized vocabulary tests (Beitchman et al., 2008; Burton & Watkins, 2007; Washington and Craig, 1992). Because children with language disorders were excluded from this study, it was assumed that the current study sample consisted of children with normal language abilities. Despite this, out of the thirty participants tested for this study, more than half (16) had a standard score lower than 85 on the EOWPVT. Considering the assumption was that all participants had normal language skills (excluding the one participant found to have a speech-language impairment [SLI]), most of the students should have scored in the average range for the assessment. Many of these same children were able to respond to scaffolding and demonstrate an Acquainted level of word knowledge during the DAVIC transfer task. These results support the implementation of response to intervention (RtI) before diagnosing a child with a disability. This is because a child who is able to display an Acquainted level of word knowledge demonstrates
that the intervention which occurred is effective. The level of scaffolding required may be implemented within a long term intervention model, with the goal of reaching an Established level of knowledge. Also, without scaffolding, most students displayed low vocabulary skills (both on the DAVIC pretest and the EOWPVT). Therefore, assessing a child without some accompanying form of support does not provide information about a child’s true abilities. Rather, in the area of vocabulary, it only demonstrates what the child has been exposed to.

**Incidental word learning.** Another finding that is consistent with prior research is the occurrence of incidental word learning. Incidental word learning is the ability to derive word meaning after being exposed to it during the natural reading process. The results of the current study are congruent with prior research which shows that children can demonstrate vocabulary knowledge after a single exposure to a word. Based on research, there is a 5% (Nagy et. al, 1987) to 15% (Swanborn & Glopper, 1999) chance that a child will learn an unknown word incidentally after being exposed to it in context. In the current study, incidental word knowledge was measured at Prompt 1 (the passage was removed from sight and the participants were asked to define the target word). During the prompting phase, participants were able to demonstrate incidental word knowledge (correctly explain the word meaning to an established level at Prompt 1) 8.6 percent of the time. During the pretest (definition production), 3% of the responses demonstrated an Established level of word knowledge prior to the participant being exposed to the words in context which means that the observed incidental word learning was not merely the result of the children already knowing the words before testing. It was also observed that at Prompt 1 only two (9%) of the transfer task (sentence production) responses revealed Established knowledge (i.e. “The *evaporation* from the water made smoke.”; “I saw a *glacier* in the North Pole float through the water.”). This means that although children may demonstrate some word
learning after one exposure in context, they may not demonstrate an Established level of knowledge for the words learned. Despite the fact that only 9% of the Prompt 1 responses resulted in an Established sentence production, the current research shows that the ability to incidentally learn a word appears to be associated with vocabulary knowledge. This is illustrated by the positive correlation that was found between the EOWPVT score and Prompt 1. This showed that the higher the participant’s EOWPVT score, the more frequently they derived word meaning after one exposure in the text.

**Teaching a word learning strategy**

The results of this study suggest that the DAVIC shows promise as an instructional tool in the RtI model. The RtI model is a cycle of assessment and intervention. Likewise, there are two basic components of the DAVIC: assessment and intervention. These components are addressed within the three phases of the DAVIC: Pretest, Graduated Prompting phase and Transfer task. As a three part system, the DAVIC represents a structured, yet practical method of assessment, intervention and progress monitoring. It is within the Graduated Prompting Phase, however, that the intervention and growth occurs.

**Use of context and word learning.** The results of this study revealed that the participants were able to derive word meaning more frequently at Prompt 2 than at any other prompt level. At this level, participants were given the passage to look at and were told to define the word. As shown in Table 4, Pass and Non Pass participants performed almost equally in the Prompt 2 category which likely resulted in the lack of a significant correlation found between the EOWPVT score and Prompt 2. The data collected shows that most participants, despite their performance on a standardized test, were able to derive word meaning after being exposed to the target word in context (Prompt 2). This finding is consistent with other studies which suggest
that standardized assessments may not be the best tool when assessing vocabulary knowledge (Gray et. al, 1999). When context is not provided, children perform significantly poorer on vocabulary measures. Therefore, when assessing vocabulary knowledge, context should be provided so that word learning occurs.

**Direct instruction and word learning.** Another key finding in this study was the significant negative correlation found between the EOWPVT and Prompt 3 (direct instruction in locating context clues). This negative correlation suggests that participants who scored lower on the EOWPVT required more focused direction in finding context clues. In an RtI model, this information would be useful when determining a child who may require more intensive intervention. Within RtI, the DAVIC could be used to determine the type of scaffolds a child needs. For example, a child who needs direct instruction in locating context clues (Prompt 3) may require intervention in the areas of developing word learning strategies and locating context clues. This does not mean that a child who needs direct instruction in locating context clues (Prompt 3) on the DAVIC Graduated Prompting task will require special education services. It does suggest, however, that before using standardized test scores to make a special education determination, direct instruction in finding context clues should be provided to determine if the child responds to intervention. It also suggests that providing direct instruction helps support lower performing children and can improve their word learning skills. The DAVIC therefore, could be used for both progress monitoring as well as intervention.

**Intensive prompting and word learning.** There was no significant correlation found between Levels 4 or 5 and scores on the EOWPVT. There was also no significant correlation between EOWPVT score and “No Level of Successful Prompting” (the level at which the child never elicited an appropriate response). Table 4 shows that at Prompts 4, Prompt 5 and NLSP,
Pass and Non Pass participants performed similarly. This suggests that there is no relationship between performance on a standardized test and the need for a high level of prompts (or reaching no level of successful prompting). With the range of 63-112 for EOWPVT scores, the implication is that in the academic setting, despite ability level, children may require intensive intervention in order to comprehend some of the words they encounter. Further, because vocabulary is learned through multiple exposures, some children (despite ability level) may not respond to a brief intervention for words they encounter in text.

**Substitution as a word learning strategy.** Table 4 shows that out of a total of 280 responses from the participants, only one response was correct at Prompt 5. At level 5, the participants were provided with information on how to substitute a new word for the target word. No significant correlation was observed between EOWPVT score and Prompt 5 which suggests that there is no relationship between performance on the EOWPVT and a child’s ability to substitute a word to derive word meaning. The data reveals that participants reached Prompt five 18 times. Out of 18 opportunities, only one resulted in a successful prompt. The other 17 occurrences did not result in a correct definition production and were consequently given a code of NLSP (No level of Successful Prompting) for that item. There are implications regarding the insignificant relationship between EOWPVT and prompt 5.

First, the method of substitution used in this study may not have been an effective strategy when examining academic vocabulary. Although prior research regarding the SCANR method (Jenkins, Matlock & Slocum, 1989) has shown substitution to be effective as a word learning strategy, the current study used a modified substitution procedure which may have decreased it effectiveness.
Another reason for the ineffectiveness of the substitution strategy during this study could be the use of Tier 3 (content specific words). Academic vocabulary words such as “erosion, condensation, and weathering” are not easily substituted by other words. Therefore, successful implementation of the substitution strategy may have been influenced by the type of words chosen.

Also, because of the challenging nature of substitution (find a synonym), this strategy may have been more appropriately placed earlier in the graduated prompting hierarchy. Responses that were not deemed correct at Prompt 5 resulted in an automatic NSLP. Therefore, it cannot be determined if the child would have responded to a different type of instruction or not.

Another explanation for the lack of a significant relationship could be that students who do not respond at Prompt 4 (provide the context clue after it is directly stated) and require Prompt 5 (substitute a different word for the target word) may demonstrate abilities that require intensive, extended instruction in finding context clues. For example, the participant that derived a word meaning using the substitution prompt (Level 5) required four modeling prompts (Level 4), two “context only” (Level 2) prompts and had three unsuccessful attempts (NSLP) for their other nine words. As a whole this participant required intensive prompting (higher than a level 2) in order to derive meaning. It is noted that despite requiring a substitution prompt (Level 5) for the word “evaporation”, the participant reached an Acquainted level of knowledge as measured by his sentence production during the transfer task: “When you boil water it evaporates”. This illustrates that although a student requires a higher level of prompting (scaffolding) in order to derive word meaning, a child can still demonstrate at least an Acquainted level of word knowledge.
The ineffectiveness of the substitution strategy in this study illustrates the need for assessment/intervention cycles such as RtI and dynamic assessment. When a child is not responding to an intervention, it does not mean that they will require special education services. Rather, it could mean that the intervention method needs to be modified in order to give the child the opportunity to make improvements. This information cannot be detected using a standardized assessment alone, which further shows the importance of dynamic assessment in the RtI model.

Table 4: Frequency of word meanings derived at each prompt level

<table>
<thead>
<tr>
<th>Prompt 1</th>
<th>Prompt 2</th>
<th>Prompt 3</th>
<th>Prompt 4</th>
<th>Prompt 5</th>
<th>NLSP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total sample</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>139</td>
<td>46</td>
<td>53</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Pass Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>69</td>
<td>18</td>
<td>29</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Non Pass Participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>70</td>
<td>28</td>
<td>24</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>

Future Research

The DAVIC can be a valuable tool for assessing the word learning potential in diverse populations. This study examined primarily low income students. However, in the school setting, many children (despite SES) will encounter unknown words in context. The DAVIC could be used to determine the level of assistance that children from various groups need in order to derive word meaning. Research should examine other populations in order to determine if the observed trends are similar to those found in this study.

Most responses produced in the DAVIC transfer task demonstrated an Acquainted level of word knowledge. Future research should determine the usefulness of Acquainted word knowledge as it pertains to a child’s reading comprehension. A poor reader can activate their schema and use prior knowledge about a topic in order to comprehend a text (Recht and Leslie, 1988). Additionally, many children (despite reading ability) will likely encounter words for
which they do not have an Established level of word knowledge. Research should therefore examine how much a child can comprehend when they have an Acquainted level of knowledge for the target words.

This study showed that with scaffolding, many children can use context clues to derive word meaning. Prior research shows that instruction in vocabulary can improve reading comprehension (McKeown, Beck, Omanson & Perfetti, 1983). Therefore, research should determine if instruction in locating context clues can result in improved reading comprehension.

The DAVIC can be used to determine the level at which a child responds to prompting during a brief assessment procedure; however, the effectiveness of this prompting needs to be examined within an intervention model such as RtI. Within an RtI model, an intervention can last for several weeks. Therefore, future research should longitudinally examine the effectiveness of an intervention which was determined during Graduated Prompting. Likewise, research should examine if the word learning that occurs during one session of Graduated Prompting can last over time. This could be achieved by examining Graduated Prompting and the test-teach-retest methods of dynamic assessment together. The intervention revealed during Graduated Prompting could be implemented in a test-teach-retest procedure to determine effects over an extended period of time.

**Limitations**

One limitation of this study was the inclusion of borderline participants. Seven participants scored within the “borderline” range of 84-86 but were placed within Pass/Non Pass Groups. The criteria for being placed in either the Pass or Non Pass group was one standard deviation (15) below the mean (100). Students who scored 85 and higher were placed into the Pass group, while students who scored 84 and lower were placed into the Non-Pass Group. Prior
research (McKeown, 1985) has examined word learning among high and low ability children and eliminated participants whose scores fell in the middle of the distribution. The marginal difference between Pass and Non-Pass participants included in this study may have contributed to the similarities in group performance on the DAVIC.

There was a low frequency of responses that occurred at prompt 5 (substitute target word for another word) when compared to the other prompts. Most students (17 out of 18) who attempted Prompt 5 were not able to derive word meaning and were coded as NSLP (no level of successful prompting). Prompt 5 may have been too challenging to be considered a high level prompt. As a result, future Graduated Prompting research should use pilot data in order to determine the placement of prompts within the hierarchy. Future research is also needed to determine the effectiveness of substitution as a word learning strategy with Tier 3 (content specific) words.

Another limitation is that despite all of the words being judged to be general or directive by four coders, one word appeared to present a greater challenge to participants than the other target words. The word “cycle” resulted in the fewest number of Level 2 responses, with only two participants able to derive word meaning at Level 2. Level 2 was the most frequent Prompt level for every other target word. The level 2 prompt consisted of the participant being provided with the passage only and no verbal prompting. It is therefore implied that the context for “cycle” was not as supportive as the context for the other words. Future research should include a pilot study to determine if the target words and context are appropriate.

**Implications and Conclusion**

This study demonstrates that Graduated Prompting directed at context clues can result in improved vocabulary knowledge. Because Graduated Prompting accounted for 81% of pretest to
transfer task improvement, there is evidence to support its implementation within the RtI framework. Within RtI, evidenced based intervention is paired with frequent assessment in order to determine a child’s academic needs. Therefore, the graduated prompting method can be utilized in the RtI model by providing educators with different levels of information about a child’s vocabulary skills. First, by administering a pretest, educators are provided with information about a child’s initial vocabulary knowledge. Next, through Graduated Prompting, educators can determine the level of instruction that the child requires in order to successfully derive word meaning. This instruction can be implemented in order to determine if the child responds to intervention. Finally, the transfer task can be used to measure how much the child improved as a result of the brief intervention. It should be noted that the “intervention” that occurs during Graduated Prompting does not suffice as a complete intervention within the RtI model. Rather, Graduated Prompting may be useful for determining an effective intervention that could be implemented with the child over time.

A major finding in this study was that there is no relationship between vocabulary knowledge as measured by the EOWPVT and word learning potential as measured by the DAVIC. Standardized vocabulary assessments measure what a child has been exposed to up to the time of testing. The nature of word learning, however, is based on exposure. Therefore word learning potential should be examined in addition to vocabulary knowledge. Standardized assessments such as the EOWPVT measure vocabulary knowledge but do not provide information regarding the specific level of word knowledge (Unknown, Acquainted, Established) or information regarding word learning potential. Graduated prompting can be used to measure vocabulary knowledge and word learning potential. It can also be used to teach a word learning strategy, such as finding context clues, and can reveal the particular strategy that results in
improved vocabulary knowledge. By teaching this strategy within a Graduated Prompting model, most children can demonstrate at least an Acquainted level of word knowledge. Therefore, when assessing a child’s vocabulary skills, Graduated Prompting is more effective at determining a child’s abilities than when using a standardized assessment alone.
References


Statewide Assessment Accommodations (2009).

http://www.ode.state.oh.us/GD/Templates/Pages/ODE/ODEDetail.aspx?page=3&TopicRelationID=9&ContentID=47700&Content=81933


http://nationsreportcard.gov/reading_2009/nat_g4.asp?tab_id=tab2&subtab_id=Tab_5#tabsContainer


http://nces.ed.gov/pubs2009/100_largest/tables/table_02.asp


Weaver, C. (2002). Reading Process and Practice: Third Ed. Heinemann, Portsmouth, NH.


Appendix A

96
## Dynamic Assessment of Vocabulary in Context (DAVIC)

<table>
<thead>
<tr>
<th>Participant Number</th>
<th>Pretest Score</th>
<th>PPVT Score</th>
<th>Date</th>
<th>DAVIC Score</th>
<th>CELF Score</th>
<th>SCORE</th>
</tr>
</thead>
</table>

### PRE-TEST

<table>
<thead>
<tr>
<th>Vocabulary Term</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word 1</td>
<td></td>
</tr>
<tr>
<td>Word 2</td>
<td></td>
</tr>
<tr>
<td>Word 3</td>
<td></td>
</tr>
<tr>
<td>Word 4</td>
<td></td>
</tr>
<tr>
<td>Word 5</td>
<td></td>
</tr>
</tbody>
</table>

### Total

<table>
<thead>
<tr>
<th>Score Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Unknown</td>
</tr>
<tr>
<td>2 Acquainted</td>
</tr>
<tr>
<td>3 Established</td>
</tr>
</tbody>
</table>

Say to each student: "Tell me what the word_____means".
Properties of Water

The water on earth is always changing. It changes from one form to another. This happens over and over. This is called the water cycle.

There are three properties of water. They are: solid, liquid and gas.

When water freezes, it changes from a liquid to a solid. When rain freezes and turns into ice, it is called sleet.

Water can also turn into gas. When water gets too hot, it evaporates. You can see evaporation when you watch water boil. The smoke that you see is the water turning into gas.

Gas can turn back into water. If you put a plate on top of boiling water, you may see little drops of condensation on the plate. The gas has turned back into little drops of water! You can also see condensation when you look at clouds in the sky. The clouds are filled with condensation. When they get too full, precipitation happens. You see precipitation when rain and snow fall out of the clouds.

Water is always changing from one form to another!

Reading level: 4.3
The earth has a lot of rocks which can change into different forms. Sometimes when you look at a rock, it looks cracked and worn away. This is because of **weathering**. During weathering, rocks break down into smaller pieces.

Most weathering is caused by water. You can see weathering when you look at rocks in the river. The small rocks you see are broken down from larger rocks because of water. You can also see weathering when you see holes in the street. When it rains, water gets into the cracks in the street. When the water freezes, it can break the rock and turn into potholes.

After weathering breaks rock down, **erosion** happens. During erosion, pieces of **sediment** are moved from one place to another. Wind and water leave behind pieces of rock called sediment. In the desert, **sand dunes** are formed by erosion. The wind moves sediment which forms large piles of sand. **Glaciers** can cause erosion. As these large pieces of ice move around in the sea, they help shape lakes and plains. When the glaciers melt, they leave behind large pieces of sediment and a lot of water which fill the lakes.
## Appendix C
### DAVIC Script and score sheet

<table>
<thead>
<tr>
<th>Prompt</th>
<th>Script</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LSP 1</strong></td>
<td>Tell me what the word ____ means.</td>
</tr>
<tr>
<td><strong>LSP 2</strong></td>
<td>Show child the passage and point out the sentence where the word occurs. Say: &quot;Re-read this to find context clues. Tell me what the word ____ means&quot;</td>
</tr>
<tr>
<td><strong>LSP 3</strong></td>
<td>Say: &quot;Context clues are words in a sentence that help you figure out what another word in the sentence means. They can come before or after the word and sometimes have trigger words like “are” or “is”. Give child context clue handout. Say: &quot;Look at this example of BEFORE and AFTER clues.&quot; Point and say: &quot;This clue (read aloud) is before and this clue (read aloud) is after. Try and find the context clues. Tell me what the word ____ means&quot;.</td>
</tr>
<tr>
<td><strong>LSP 4</strong></td>
<td>Point to the context clue: Say: &quot;This is the context clue. It says&quot; (read context clue aloud). Say: &quot;Now tell me what the word ____ means&quot;.</td>
</tr>
<tr>
<td><strong>LSP 5</strong></td>
<td>Say: &quot;Sometimes you can substitute a word that you do know for the word that you don’t know.&quot; Show Substitution Handout. &quot;Look how this word was substituted for the harder word. Try and think of a word that you can substitute that makes sense with the context clues. Now, tell me what the word ____ means.&quot;</td>
</tr>
<tr>
<td><strong>NSLP</strong></td>
<td>Say &quot;The word ____ means ____.&quot; If the child reaches this level, then no transfer task should be attempted and move to next item.</td>
</tr>
<tr>
<td><strong>Transfer Task</strong></td>
<td>Say: &quot;Okay, now use the word ____ in your own sentence&quot;</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LSP 1</strong></td>
</tr>
<tr>
<td><strong>LSP 2</strong></td>
</tr>
<tr>
<td><strong>LSP 3</strong></td>
</tr>
<tr>
<td><strong>LSP 4</strong></td>
</tr>
<tr>
<td><strong>LSP 5</strong></td>
</tr>
<tr>
<td><strong>NSLP</strong></td>
</tr>
</tbody>
</table>

Unknown
Acquainted
Established
Appendix D

Coding Key 1

Unknown, Acquainted, Established

Unknown-A response will be coded as “unknown” when it bears no relation to the stated definition of the term. Acquainted-A response will be coded as “acquainted” when it reveals a basic understanding of the word. Established-A response will be coded as “established” when the word is immediately recognized and is used.

Code responses as: unknown, acquainted, and established, based on their resemblance to the following examples. For the pretest, judge responses based on their similarity to provided examples. For the transfer task, judge responses (sentences) based on their reference to the provided examples.

1. Cycle
   a. Unknown- water; no reference to changing or a repeating event
   b. Acquainted-changes from one form to another; solid liquid, gas
   c. Established-something that happens over and over

2. Sleet
   a. Unknown-no reference to ice
   b. Acquainted-ice, rain
   c. Established-frozen rain

3. Evaporation
   a. Unknown-water; no reference to gas
   b. Acquainted- gas; boiling water, heat smoke
   c. Established-water that turns into gas when it gets too hot; smoke from boiling water, mentions boiling water and air

4. Condensation
   a. Unknown- no reference to drops of water, plate
   b. Acquainted-drops of water, clouds, gas, water
   c. Established-gas that turns back into water/liquid, mentions both boiling water and drops of water

5. Precipitation
   a. Unknown-no reference to something that falls from clouds or the sky
   b. Acquainted-clouds
   c. Established-rain, sleet or snow
Appendix D

Coding Key 2

Unknown, Acquainted, Established

Unknown-A response will be coded as “unknown” when it bears no relation to the stated
definition of the term. Acquainted-A response will be coded as “acquainted” when it reveals a
basic understanding of the word. Established-A response will be coded as “established” when the
word is immediately recognized and is used.

Code responses as: unknown, acquainted, and established, based on their resemblance to the
following examples. For the pretest, judge responses based on their similarity to provided
examples. For the transfer task, judge responses (sentences) based on their reference to the
provided examples.

6. Weathering
   a. Unknown- a rock, no mention of cracking, wearing or breaking down
   b. Acquainted- a cracked rock
   c. Established-Breaking down rocks into smaller pieces, relation to erosion

7. Erosion
   a. Unknown- rocks breaking down, no mention of moving sediment
   b. Acquainted- sediment moving, rocks moving, caused by some type of weathering
   c. Established-the process of moving sediment from one place to another

8. Sediment
   a. Unknown- no mention of rock
   b. Acquainted- pieces of rock
   c. Established- pieces of rock (or “something”) left by wind or water

9. Sand dunes
   a. Unknown- no mention of sand or piles
   b. Acquainted- piles or sand, references a location such as desert or beach but doesn’t
      reference piles
   c. Established- large piles of sand

10. Glaciers
    a. Unknown- water; sea; a solid; no reference to ice
    b. Acquainted- a piece of ice; frozen water; referencing something floating in a body of
       water such as sea or ocean.
    c. Established- sea water that has frozen or turned into ice; references something
       floating in a frozen location
A misdirective context is one that leads a reader to gather an incorrect meaning of a target word. A nondirective context provides no guidance in determining the meaning of a target word. A general context provides basic information that may lead a reader in determining a basic category for a target word. And finally, a directive context leads a reader to the appropriate meaning of the target word. Beck, McKeown and McCaslin (1983). Read the passage: “Properties of Water”. Code responses as: misdirective, nondirective, general or directive, based on their relationship to the above definitions.

<table>
<thead>
<tr>
<th>Target word</th>
<th>Context Clue</th>
<th>Misdirective, Nondirective, General or directive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cycle</td>
<td>This happens over and over. This is called the water cycle.</td>
<td></td>
</tr>
<tr>
<td>Sleet</td>
<td>When rain freezes and turns into ice, it is called sleet.</td>
<td></td>
</tr>
<tr>
<td>Evaporation</td>
<td>When water gets too hot, it evaporates. You can see <strong>evaporation</strong> when you watch water boil. The smoke that you see is the water turning into gas</td>
<td></td>
</tr>
<tr>
<td>Condensation</td>
<td>If you put a plate on top of boiling water, you may see little drops of <strong>condensation</strong> on the plate. The gas has turned back into little drops water!</td>
<td></td>
</tr>
<tr>
<td>Precipitation</td>
<td>You see <strong>precipitation</strong> when rain and snow fall out of the clouds.</td>
<td></td>
</tr>
</tbody>
</table>
Definition Coding Key 2

Misdirective, Nondirective, General, and Directive

A misdirective context is one that leads a reader to gather an incorrect meaning of a target word. A nondirective context provides no guidance in determining the meaning of a target word. A general context provides basic information that may lead a reader in determining a basic category for a target word. And finally, a directive context leads a reader to the appropriate meaning of the target word. Beck, McKeown and McCaslin (1983). Read the passage: “Properties of Water”. Code responses as: misdirective, nondirective, general or directive, based on their relationship to the above definitions.

<table>
<thead>
<tr>
<th>Target word</th>
<th>Context Clue</th>
<th>Misdirective, Nondirective, General or directive?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weathering</td>
<td>Sometimes when you look at a rock, it looks cracked and worn away. This is because of weathering. During weathering, rocks break down into smaller pieces.</td>
<td></td>
</tr>
<tr>
<td>Erosion</td>
<td>During erosion, pieces of sediment are moved from one place to another.</td>
<td></td>
</tr>
<tr>
<td>Sediment</td>
<td>Wind and water leave behind pieces of rock called sediment.</td>
<td></td>
</tr>
<tr>
<td>Sand dunes</td>
<td>In the desert, sand dunes are formed by erosion. The wind moves sediment which forms large piles of sand.</td>
<td></td>
</tr>
<tr>
<td>Glaciers</td>
<td>Glaciers can cause erosion. As these large pieces of ice move around in the sea, they help shape lakes and plains.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix F

Sentence Production Coding Rules

The sentence productions are coded based on the participants’ demonstration of semantic and syntactic knowledge of each target word.

**Semantic knowledge**—A sentence production will be described as demonstrating semantic knowledge when it references information specified in the Definition Coding Key.

**Syntactic knowledge**—A sentence production will be described as demonstrating syntactic knowledge when the integrity of the word form is maintained, or the word form is changed and used in a semantically and syntactically acceptable manner.

**Unknown**

A sentence production will be coded as Unknown when:

1. The production reveals unknown syntactic knowledge about the target word (i.e. uses target word in an unacceptable word form: “I sleet ice cream into the freezer.”) and demonstrates unknown semantic knowledge as determined by the Definition Coding Key. OR

2. The production is syntactically acceptable but does not reveal the participant’s level of semantic knowledge and does not reveal the participant’s level of syntactic use of the word. (i.e. “Today I learned what the word cycle means.”) OR

3. The participant does not produce a sentence.

**Acquainted**

A sentence production will be coded as Acquainted when:

1. It reveals an acquainted level of semantic knowledge as specified in the Definition Coding Key. OR

2. The sentence reveals established or acquainted semantic knowledge but the term is not used in a syntactically appropriate manner. OR

3. The production reveals appropriate syntactic knowledge about the target word but does not reference the information specified in the Definition Coding Key.

**Established**

A sentence production will be coded as Established when:

1. The target word is used in a syntactically appropriate manner and reveals established semantic knowledge as specified in the Definition Coding Key.
Appendix G

The Dynamic Assessment of Vocabulary in Context (DAVIC) procedure

1. DAVIC Pretest
   - The participant is asked to define each word
   - Responses are written down by P.I

2. Graduated Prompting Phase
   - While audio recorded, the participant is given the first passage to read aloud
   - After the passage has been read aloud, the participant is asked to define the first target word.
   - Prompts are provided until the participant is able to define the word
   - After defining the target word, the participant immediately moves to the DAVIC transfer task word for that item. SEE DAVIC Transfer Task
   - After 10 words from passage are complete, the second passage is presented and the Graduated Prompting Phase is repeated.

3. DAVIC Transfer Task
   - This step occurs within the Graduated Prompting Phase
   - After the participant defines a word, they are asked to use that target word in a sentence
   - Response is written
   - The next word is presented in the Graduated Prompting Phase