Illinois State University

ISU ReD: Research and eData

2-26-2014

The Effect of Metalinguistic Strategy Instruction on the Oral and Written Expression of School-Aged Children

Karen Lara Dudek
Illinois State University, kldudek1980@gmail.com

THE EFFECT OF METALINGUISTIC STRATEGY INSTRUCTION ON THE ORAL AND WRITTEN EXPRESSION OF SCHOOL-AGED CHILDREN

Karen L. Dudek

208 Pages   May 2014

Vocabulary knowledge is critical for academic success; and research has indicated that students with low vocabularies can learn metalinguistic strategies that can improve their performance in school. In this study, I investigated the impact of metalinguistic strategy instruction on the oral and written expression abilities of elementary-aged children (third grade). The strategy was the Expanding Expression Tool (EET; Smith), which can help students to increase their oral and written expression by describing words using semantic features. I used a pretest-posttest-posttest between groups design to investigate the impact of three experimental conditions on oral and written expression over a nine week period: 1) Condition T_1, which received metalinguistic strategy instruction twice per week, 2) Condition T_2, which received metalinguistic strategy instruction four times per week, and 3) the control condition, which received standard enrichment in the area of reading and literacy.

Results indicated that all conditions were effective in increasing the oral describing abilities across testing time; although the two groups who received EET instruction improved slightly more than the control condition. Students in all conditions

improved the number of semantic features used in written expression; however the treatment conditions were more effective than the control condition in improving the diversity of semantic features students used. Condition T_1 was equally effective in improving written describing abilities compared to Condition T_2; however students in Condition T_2 had better retention of the strategy following a brief cessation in treatment. I have explained implications, limitations, and needed future research relating to these findings in my discussion.

THE EFFECT OF METALINGUISTIC STRATEGY INSTRUCTION ON THE ORAL

AND WRITTEN EXPRESSION OF SCHOOL-AGED CHILDREN

KAREN L. DUDEK

A Dissertation Submitted in Partial
Fulfillment of the Requirements
for the Degree of

DOCTOR OF EDUCATION

Department of Special Education

ILLINOIS STATE UNIVERSITY

2014

THE EFFECT OF METALINGUISTIC STRATEGY INSTRUCTION ON THE ORAL

AND WRITTEN EXPRESSION OF SCHOOL-AGED CHILDREN

KAREN L. DUDEK

COMMITTEE MEMBERS:

Julia B. Stoner, Chair
Christina M. Borders
John K. Rugutt
Heidi M. Harbers
Karen H. Douglas

ACKNOWLEDGEMENTS

CONTENTS

TABLES

CHAPTER I

CONCEPTUAL FRAMEWORK AND STATEMENT OF THE PROBLEM

Introduction

Vocabulary skills are necessary for academic success, may be impaired in

students with language disabilities and/or learning disabilities (LD), are affected by

exposure and metalinguistic skills, and can be taught using direct instruction, and

facilitating the use of metalinguistic strategies. First, I have organized this chapter to

discuss my conceptual framework. Second, I have provided a statement of the problem,

which summarizes previous research that supports the need for my study. Third, I

describe the significance of my research and then, fourth, I discuss the purpose of my

research. Fifth, I present my specific research questions. Sixth, I provide a definition of

key terms, to specifically guide the reader in understanding terms that are unique and

used frequently in the research and in my study. Seventh, I follow this with the

limitations and assumptions of my study. Finally, I conclude this chapter with a brief

summary.

Conceptual Framework

Public Law (PL) 94-142, the Individuals with Disabilities in Education Act

(IDEA; 2004), contains language mandating states to provide a free and appropriate

education (FAPE) to all children with disabilities. In order to ensure FAPE, professionals in the school systems are obligated to provide instruction to students with disabilities in their least restrictive environment (LRE). For many students with disabilities, this means that they receive the majority of instruction in general education classrooms. Therefore, it is incumbent for both general and special education teachers to have knowledge of effective instructional methods to utilize in order to ensure students with disabilities can access the curriculum.

The IDEA recognizes that students with language disabilities and/or LD may struggle to learn new concepts, comprehend written texts, or describe new information in oral and written form; these skills are integral and necessary to obtain literacy. Literacy is perhaps the prime component of a successful academic education and can be significantly impacted by deficits in these areas. Additionally, students with disabilities in these areas may qualify for specialized services to address literacy skills. Specifically, many students with disabilities struggle academically due to delays in vocabulary and have difficulty applying independent word-learning strategies (Cunningham & Stanovich, 1997; Zipke, 2007).

However, research has demonstrated these students can improve their language skills and academic performance given appropriate instruction addressing semantic skills (e.g., Beck & McKeown, 2007; Cain, 2007). Vocabulary is a common area of weakness for students with disabilities and low vocabulary skills can impact literacy and academics. In addition, when a student with disabilities has poor vocabulary skills and does not received instruction in this area, the deficit increases as academics, with more difficult vocabulary, becomes more complex with each grade level. It is imperative that

2

teachers address this issue, through evidence-based practices (EBPs) in the area of vocabulary. In order for students with disabilities to receive scientifically based instructional strategies within their LRE, teachers must utilize EPBs in the area of vocabulary within the contexts of general education settings to ensure students receiving inclusive services continue to make adequate academic progress.

Statement of the Problem

Strong vocabulary skills are necessary for academic success in order for students to tie new information to prior knowledge and learn information independently. However, some students may come to school with significantly less vocabulary knowledge than their peers due to environmental factors or language and/or learning disabilities (Hart & Risley, 1995). These students need direct instruction in vocabulary or metalinguistic strategy use because certain contexts, such as written language in texts, can be difficult to comprehend (Frishkoff, Perfetti, Collins-Thompson, 2010). Vocabulary and metalinguistic awareness not only impacts reading comprehension and independent word learning, but also performance on expressive language tasks (Justice, Meier, & Walpole, 2005). Students with weak vocabulary knowledge are often less able to explain words using semantic features (Alt, Plante, & Creusure, 2004) and are more likely to struggle with writing composition (Gersten, Fuchs, Williams, & Baker, 2001).

Research has suggested that direct instruction of words, involving active use of language and analysis of semantic features results in robust word learning across ability levels is effective; however these methods may be time consuming and unrealistic when teaching all the academic vocabulary students need to attain (Biemiller & Boote, 2006; Justice et al., 2005; Marulis & Neuman, 2010). Fortunately, the literature has indicated

3

that students can successfully learn to use metalinguistic strategies to increase language expression (Beck & McKeown, 2007; Cain, 2007; Zipke, 2012), particularly those that incorporate multisensory mnemonic devices for memory of strategy steps (Graham & Perin, 2007). However, a number of questions remain concerning a teaching method that incorporates these methods (i.e., direct instruction, active use of vocabulary, semantic features, and metalinquisitc strategies) that can facilitate the attainment of robust vocabulary. Therefore, research is needed to determine an efficient and effective method for teaching students to learn new words and use these words in varying, generative contexts.

Significance of Research

Vocabulary, also known as semantic knowledge, refers to one's knowledge of words and their meanings (Beck, McKeown, & Kucan, 2002). Individuals who have strong vocabulary skills typically also have strong metalinguistic awareness skills, which enable them to have deep understandings of words and their meanings (Graves, 2006). In the school setting, teachers may rely on direct instruction to teach students specific words as well as metalinguistic strategies to improve independent word learning (Beck et al., 2002). Students can successfully learn words and word learning strategies given direct instruction (Justice et al., 2005). However their success in utilizing these skills and strategies can depend on a number of factors; such as the presence or absence of language and LDs (Cain, Lemmon, & Oakhill, 2004).

Students with disabilities, particularly those specific to language and vocabulary, may have varying capabilities that can affect their ability to learn new words. For example, students with disabilities may be less efficient in fast mapping, which

4

essentially means constructing a meaning for an unfamiliar word upon minimal exposure(s) (Alt et al., 2004). Additionally, students with disabilities may struggle to comprehend academic texts, complete written language tasks, explain and describe words, and utilize independent word learning strategies (Gersten et al., 2001; Zipke, 2007). When students with language and/or LDs are unable to retain new meanings of words at the same rate as their peers, further gaps may occur over time (Marulis & Neuman, 2010). In order to comply with the guidelines in IDEA (2004), educators must provide vocabulary instruction for students with disabilities in the LRE to close these gaps, thus warranting the need for research to determine effective vocabulary and metalinguistic strategy instruction for students of varying ability levels.

<div align="center">**Purpose of the Study**</div>

For the purpose of this study, I investigated the impact of a metalinguistic strategy instruction on the expressive language abilities of third grade students with and without disabilities. The strategy, known as the Expanding Expression Tool (EET; Smith, 2011), is a method that incorporates direct instruction of words through the use of a multisensory tool. Specifically, EET instruction involves exposing and encouraging students to produce oral and written explanations of words by explaining their semantic features (i.e., category, function, physical appearance, composition, associations or items, and location). In order to aid explanation, students use a mnemonic device to recall the features and steps to describe objects. During instruction, teachers expose students to the steps in the mnemonic device through a verbal chant and visual aids, thus incorporating multisensory information.

While the EET consists of a number of features proven effective by previous research (e.g., utilizing direct explanation of words, analysis of semantic features, metalinguistic strategy instruction, and multisensory mnemonic devices), there are currently no published studies documenting the effectiveness of the EET. The EET manual provides descriptions of how to use the strategy, however, more information is needed to determine the frequency and intensity of instruction required for maximal results. There is also no published research that documents the impact of whole-class instruction incorporating the EET method on students' oral and written language skills. Consequently, while EET instruction may be effective in increasing students' vocabulary it is not, at this point, an EPB due to the absence of empirical research on this method.

Research Questions

The research questions that I answered in my study are:

1) What is the impact of whole-class metalinguistic strategy instruction on the oral language skills of students in elementary school (grade 3)?

2) What is the impact of whole-class metalinguistic strategy instruction on the written language skills of students in elementary school (grade 3)?

3) What is the impact of treatment dosage of metalinguistic strategy instruction on the oral and written language skills of students in elementary school?

Definitions and Terms

In the following sections, I have explained key terms relevant to understanding the literature and research questions of the study. These terms relate to the underlying theories and strategies associated with metalinguistic strategy instruction.

6

Semantic Features

Vocabulary knowledge goes beyond simply being able to identify or repeat words (Graves, 2006). Children who have well-developed vocabularies are not only able to provide definitions of words, but are also able to describe semantic features of words and how they relate to other constructs (Beck et al., 2002; Taylor, Mraz, Nichols, Rickelman, & Wood, 2009). "Semantic features" may include descriptive attributes such as the function, location, physical make-up, physical appearance, or other physical attributes (Beck & McKeown, 2007; Neuman & Dwyer, 2009). Students' semantic representations of words consist of the knowledge they have of these features of words in their schema (Alt et al., 2004).

Fast Mapping

Fast mapping is the process through which we learn words. This phenomenon occurs when we rapidly construct meanings of words and store them for later retrieval based on initial encounters (Alt et al., 2004; McGregor, Newman, Reilly, & Capone, 2002). Our semantic representations of words after initial encounters may often be inaccurate or incomplete; however may change over time with repeated exposures to the words (Perfetti, 2007). A number of factors can influence the effectiveness and efficiency of fast mapping. One critical factor is existing vocabulary knowledge and language skills (McGregor et al., 2002).

The context of initial exposure may also impact fast mapping, in that exposure within less informative contexts may result in incomplete semantic representations (Frishkoff et al., 2010). For example, even students with strong vocabularies may struggle to learn meanings of unfamiliar words in certain academic texts that contain

7

inadequate contextual information; thus necessitating the need for explicit explanation of some academic vocabulary.

Students with speech and language impairments or weak vocabulary skills often fast map semantic features of words less effectively than their typically developing peers (Alt et al. 2004). They may therefore have limited semantic knowledge because they have attached fewer or less accurate semantic attributes to words. Consequently, students with low vocabulary and less effective fast mapping skills require effective and efficient teaching methods, to facilitate vocabulary development.

Working Memory and Multisensory Processing

When considering fast mapping and recall of words and their features, it is necessary to consider the way our brains process and store information (Sousa, 2011). Baddeley (2003) explained working memory as a three-part system involving a phonological loop which codes auditory signals into working memory, a visuospatial sketchpad which encodes visual information into working memory, and a central control mechanism which coordinates phonological and visuospatial information and facilitates transfer to long-term memory.

These working memory processes are relevant to vocabulary learning in order to determine the most effective means of transferring semantic information from working memory to long term memory, considering the role of multisensory input.   Some students may process information more effectively through one type of sensory information compared to others, such as those with weak phonological skills who process visual information more effectively than auditory information (Kibby, Marks, Morgan, & Long, 2004). Therefore, it is important to ensure that children receive information

8

through various senses when exposed to vocabulary words, in order to increase the chances that they store information in long term memory; this becomes especially crucial when students have language and/or learning disabilities.

The Matthew Effect

As was mentioned in the preceding section, some children learn words less efficiently than typically developing peers due to poor background knowledge or inefficient fast mapping abilities resulting in knowledge gaps over time (McGregor et al., 2002). This phenomenon is known as the Matthew Effect (Stanovich, 1986). The Matthew Effect occurs for a number of reasons. First of all, children with poor background knowledge know fewer words than their typically developing peers (Hart & Risley, 1995). They are therefore less able to tie new information to existing knowledge, which results in difficulty retaining new information at the same rate as their peers. Additionally, because they may fast map less effectively than other children, they have weaker semantic representations of words, resulting in poor metalinguistic awareness skills (Zipke, 2007). One important metalinguistic strategy in learning new words is the ability to infer word meanings through context. While some texts may be difficult for even high achieving students, a large amount of vocabulary learning occurs during reading (Graves, 2006). Children with poor vocabularies may have weak reading comprehension skills, making it difficult for them to both derive meanings from texts and infer meanings of words through contexts.

The Matthew Effect also occurs because children with weak vocabularies tend to read less effectively than their typically developing peers (Zipke, 2012). Texts consist of more sophisticated language than conversational speech. Therefore, children who read

9

frequently encounter more difficult vocabulary, thus improving their language skills. Students who shy away from reading because it is difficult for them may encounter even greater learning gaps if they are not exposed to the complexities of written language (Beck et al., 2002). Also, direct explanation of words may be critical for students who have poor vocabulary skills. Additionally, because they struggle to learn words independently, students need explicit instruction to teach them to effectively use metalinguistic strategies to increase the rate at which they learn language (Cain, 2007). Once again, if there is a language and/or LD present in a student, these issues become all the more important.

Metacognitive and Metalinguistic Strategies

Because teachers cannot feasibly provide direct explanations and extended instruction of words for all academic vocabulary students will need to learn, they must provide them with the skills they need to both learn new words and expand their expression independently (Zipoli, Coyne, & McCoach, 2010). I have referred to these strategies in the preceding and following sections as "metacognitive" or "metalinguistic" strategies. While the two are not synonymous terms, they are interrelated constructs. The term "cognitive" refers to thought processes, while the term "meta" refers to explicit awareness and analysis (Roehr, 2008). Therefore, "metacognition" refers to one's ability to analyze one's own thought processes.

Metacognitive strategies include any techniques used to purposefully improve memory, retention, and processing of information, such as rehearsal, using planners or calendars to recall appointments, or using visual aids to recall protocol steps. "Linguistic" refers to language, while "metalinguistic" refers to the awareness and

10

analysis of language uses and features. Metacognition is an umbrella term that includes metalinguistic awareness (Graves, 2006; Roehr, 2008). Therefore metalinguistic strategies are specific types of metacognitive strategies specifically focused on language processing. They can include analysis of grammatical, syntactic, phonological or morphological rules. Metalinguistic strategies specific to vocabulary knowledge could include analysis of semantic features, word associations, figurative language, or multiple meanings of words (Zipke, 2012).

Mnemonic Devices

Mnemonic devices are tools often utilized within metalinguistic and metacognitive strategies, and consist of some type of tool to assist in memorizing strategy steps. This may involve using some type of acronym or visual aid (Graham & Perin, 2007). For example, students can use what is known as Self-Regulated Strategy Development (SRSD; Gersten et al., 2001; Little, Lane, Gersten, & Sandmel, 2010). SRSD is a metacognitive strategy that involves using acronyms such as DARE (*D*evelop topic sentence, *A*dd supporting details, *R*eject arguments from the other side, *E*nd with a conclusion) to monitor one's language use during academic tasks such as writing compositions (Chalk, Hagan-Burke, & Burke, 2005). The mnemonic device in this case would be the acronym, because it provides a memory aid to assist in recalling the steps. Other mnemonic devices could include the use of chants or rhymes to recall strategy steps (Smith, 2011; Elliot & Gentile, 1986).

<div align="center">

Limitations

</div>

A limitation to this study was that intervention was only conducted at one school and in one grade level. This may limit generalizability of findings to populations in other

<div align="center">

11

</div>

geographic regions and across other grade levels. Further research is therefore needed to replicate findings and determine the impact of the EET metalinguistic strategy instruction on the skills of students in other settings and across different ages. A second limitation of the study was that some of the participants had some exposure to the strategy within small group instruction in the previous school years leading up to the study. However, no participants had received any whole class instruction related to the strategy prior to intervention. Additionally, they had not received instruction following the specific protocol outlined in the study. I established that groups were equivalent in expressive and receptive vocabulary prior to initiating the study by administering the Peabody Picture Vocabulary Test-4th Edition (PPVT-4) and the Expressive Vocabulary Test-2nd Edition (EVT-2).

An additional limitation existed related to instrumentation. Participants were tested on their ability to describe words using semantic features of words. Words on the pretest, posttest, and delayed posttest were the same, thus creating the chance that pretesting effects could have occurred. However, I utilized a control group and pretest-posttest-posttest design to determine if significant differences in performance occurred over time, and if differences were statistically significant across conditions. One final limitation is that the teachers in the study did not utilize all supplemental materials in the EET instructional kit. They did, however, utilize all materials necessary to implement the method with fidelity and with all necessary components (i.e., mnemonic device, picture cues of all semantic features, and directions in the manual). The number of materials was limited to this list in order to ensure consistency across conditions for the study. Future

12

research is therefore needed to investigate the impact of the EET method incorporating other supplemental materials.

Assumptions

For the purpose of this study, I made the assumption that the participating teachers understood the importance of language and vocabulary skills to academic success. I also made the assumption that the teachers were familiar with the appropriate vocabulary expected in the third grade curriculum. Additionally, I assumed that the teachers understood the importance of delivering evidence-based practices in their classrooms to improve students' language skills. I made the assumptions that both teacher and student participants filled out surveys associated with this study honestly and to the best of their ability. Finally, I assumed that the teachers delivering interventions were familiar with appropriate reading and language enrichment activities typically conducted in third grade classrooms for the purpose of developing the control condition.

Summary

In order to comply with IDEA (2004) guidelines, professionals working with students with language and/or LDs must have adequate knowledge of evidence based instructional practices that will enhance the academic success of these students. Students with disabilities often have poor vocabulary and metalinguistic awareness skills which can negatively impact their success in school (Zipke, 2012). Educators must therefore implement scientifically based instructional methods to support the needs of students with disabilities. Students with poor vocabulary skills can successfully learn new words and apply word learning strategies given effective instruction (Beck & McKeown, 2007); thus warranting the need for further research documenting the effectiveness of instruction

13

focused on vocabulary and metalinguistic strategy use on students' expressive language skills. The purpose of this study was therefore to investigate the impact of the EET instructional method designed to improve metalinguistic strategy use on students' oral and written language skills.

CHAPTER II

REVIEW OF THE RELATED LITERATURE

Introduction

In this chapter I present a review of the research documenting the relationship of vocabulary and metalinguistic awareness on academic performance, as well as the features of effective vocabulary and metalinguistic strategy instruction. In considering the components of effective instruction, I discuss the role of multisensory information and its role in transferring knowledge to memory. I also review literature on direct instruction of words; including that which emphasizes active language use and analysis of semantic features. The final sections of my review included studies documenting the effectiveness of metalinguistic and metacognitive strategies in enhancing language expression, and the impact of mnemonic devices in recalling strategy steps. To locate published studies on related topics I conducted keyword searches in the Academic Search Complete database using the following search terms: *vocabulary, semantic interventions, semantic features, direct instruction, metalinguistic, metacognitive, writing, multisensory, mnemonic*, and *memory*. I also conducted ancestral searches of articles which contained relevant information to any of these topics.

Impact of Vocabulary and Metalinguistic Awareness on Academic Performance

Student's language skills have a significant impact on their performance in school. One aspect of language that has received attention in the literature is that of vocabulary skills (Beck et al., 2002). This area has a profound impact on all academic areas because curricular standards require students to learn concepts at a rapid rate, tie new semantic representations to prior knowledge, and apply them across multiple contexts and tasks (Graves, 2006). Students must also read independently to gain new information, requiring them to have adequate background knowledge to comprehend language in academic texts (Beck et al., 2002).

Vocabulary Knowledge

Research has indicated a strong relationship between vocabulary knowledge and reading comprehension abilities (Nagy & Scott, 2000; Scarborough, 2001; Zipke, 2007). Vocabulary knowledge in early elementary school can reliably predict which students will develop stronger comprehension skills later in their academic careers (Cunningham & Stanovich, 1997; Scarborough, 2001). Reading comprehension depends on a student already knowing between 90 and 95% of the words in a text (Nagy & Scott, 2000). When readers understand more than 90% of the text, they are often able to learn the remaining percentage of words that are unfamiliar by inferring meanings from the surrounding context (Biemiller, 2001). Contrarily, students who understand less than 90% of words may struggle to comprehend what they are reading, and may also have difficulty inferring meanings of unfamiliar vocabulary (Cain et al., 2004). The process of inferring meanings of new concepts is an important metalinguistic strategy students can utilize to independently learn new information (Graves, 2006).

16

Metalinguistic skills, in addition to vocabulary skills, are also necessary and predictive of academic success (Dreher & Zenge,1990; Zipke, 2007; Zipke, Ehri, & Cairns, 2009). Metalinguistic knowledge, or the awareness of the features and use of language, is predictive of future academic success in that it impacts students' ability to apply strategies to academic tasks such as reading texts or writing compositions (Dreher & Zenge, 1990; Gersten et al. 2001; Zipke, 2007). In the following section, I have discussed the literature focused on the predictive quality of metalinguistic skills, as well as factors which impact students' ability to utilize such strategies.

Metalinguistic Awareness and Strategy Use

Dreher and Zenge (1990) investigated the relationship between metalinguistic skills on reading comprehension scores with 65 children in a mid-Atlantic county school system of varying socioeconomic class. Researchers conducted 20-item interviews with children when they were in first grade to measure their metalinguistic awareness. They documented metalinguistic awareness by asking the children questions to determine the following: (a) their understanding of reading as a meaning gathering process, (b) their ability to identify language segments (e.g., letters, words, sentences), and (c) their ability to define instructional terms specific to reading. They measured reading comprehension in third and fifth grades using the California Achievement Test, and also measured academic aptitude using the Short Form Test of Academic Aptitudes. Using a regression analysis, they determined that academic aptitude accounted for 20% of the variance in reading comprehension in third grade and 35% in fifth grade; both were statistically significant. Metalinguistic awareness accounted for 17% of the variance in reading comprehension scores in third grade and 26% in fifth grade, which were not statistically

17

significant. However, combining metalinguistic awareness with academic aptitude strengthened the predictive quality of both measures, resulting in statistically significant variances at both grade levels. In combination, metalinguistic awareness skills and aptitude accounted for 28% of the variance in reading comprehension scores in third grade and 47% of the scores at fifth grade. The researchers concluded that metalinguistic awareness skills provided useful predictive information regarding future reading comprehension abilities. They further suggested that weak metalinguistic awareness skills will result in a Matthew Effect over time, in that students with higher skills learned at a faster rate than students with weaker skills.

Zipke (2007) conducted a similar study during which she investigated the relationship of vocabulary and metalinguistic skills to reading comprehension with 105 6th and 7th graders from a small charter school in Delaware in a middle socioeconomic neighborhood. Participants ranged from 11 years 1 month to 14 years 10 months. The researcher measured metalinguistic skills through sentence ambiguity tasks and a riddle-solving task. For the sentence ambiguity task, she presented students with 40 written sentences with either a lexical or structural ambiguity. Participants were to indicate whether or not the sentence had one or two possible meanings. For the riddle task, participants read 25 different riddles and were to choose the correct solution from two possible choices. Zipke measured reading comprehension and vocabulary by administering the reading comprehension and vocabulary subtests of the Gates-McGinitie Reading Test-4th Edition (GMRT-4).

Results indicated a strong correlation between reading comprehension and metalinguistic awareness as measured by the riddle task. There was also a correlation

18

between the sentence ambiguity task and reading comprehension, although it was not as strong as the riddle task. A regression analysis indicated that vocabulary subtest scores accounted for 59% of the variance in scores on the reading comprehension task. The riddle solving task and vocabulary scores in combination accounted for 63% of the variation in reading comprehension scores, while adding the sentence ambiguity scores to the equation did not significantly change the correlation coefficient. These results indicated that knowledge of word relationships and meanings, as measured by the riddle task and vocabulary scores, may have been critical skills for the academic success of students in this study.

The findings of Zipke (2007) are consistent with findings of Dreher and Zenge's (1990) study, which showed a correlation between metalinguistic skills and future reading abilities. Together, these studies have suggested that metalinguistic skills such as word defining abilities and knowledge of language features (e.g., letters, sounds, words, and sentences) may be key components to academic success. While Dreher and Zenge (1990) and Zipke (2007) studied the predictive quality of metalinguistic strategies, Cain et al. (2004) investigated the impact of students' current reading abilities and their ability to apply strategies to learn new words. Specifically, they measured the ability of 9- and 10-year-old students to infer meanings of novel vocabulary words from context in text in a two-part experiment.

In Study 1, Cain et al., (2004) selected students for the study using the Gates-MacGinitie Primary Two Vocabulary Test. They eliminated students who scored extremely high and very low, selecting the middle 74 students out of 227 tested. They then administered the Neale Analysis of Reading Ability to identify 12 students skilled in

19

comprehension (performing at or above average) and 13 students less skilled in comprehension (performing at least 8 months below their age group norms). For the vocabulary inferencing task, experimenters read eight short stories containing novel words and asked students to read the story and explain the word meanings. They also asked students to complete a working memory task during which students completed sentences with the correct words and recalled the words at a later time. Results indicated a significant main effect for the skill group, with skilled comprehenders inferring definitions more accurately than less skilled comprehenders. Proximity of words also interacted with skill group in that less skilled comprehenders had more difficulty inferring word meanings than skilled comprehenders when contextual information relating to meanings of words was further away from novel words. Less skilled comprehenders had more difficulty than skilled comprehenders on the working memory task.

Study 2, the experimenters selected 12 students skilled in comprehension, 12 students less skilled in comprehension, 12 students less skilled in comprehension with weak vocabulary skills using similar selection procedures as the Study 1. They identified students with low vocabulary skills and weak comprehension skills using the word associations subtest of the Clinical Evaluation of Language Fundamentals-Revised (CELF-R) and the Graded Nonword Reading Test. For the direct instruction task, experimenters explained novel word meanings to students and asked them to read passages containing the words and subsequently explain what words meant. The vocabulary inferencing subtest was the same as the first study. Participants also completed a short-term memory forward digit span test, and two working memory

20

listening span procedures; the working memory task from the first study and an additional task with numbers.

Results indicated a significant effect for the skill group for the direct instruction and vocabulary inferencing tasks. Students with weaker vocabulary skills required significantly more word explanations in the direct instruction condition to recall word meanings. There was no difference between groups in ability to recall details from the reading passages, but less skilled comprehenders and less skilled comprehenders with weak vocabulary had more difficulty inferring word meanings. Proximity also interacted with the skill group, as the skilled group performed better than the other groups when the contextual information was far from the novel word in the text. There were no significant differences in short term memory or the number working memory task, but the skilled comprehenders performed better than the other groups on the word listening span working memory task. Results from Cain et al.'s (2004) study further highlighted the importance of vocabulary knowledge and metalinguistic awareness in word learning and academic success. Because students who have weak vocabulary skills struggle to use metalinguistic strategies, they learn new information at a slower rate than those with stronger skills who effectively utilize these strategies.

Gersten et al. (2001) also conducted research investigating the effectiveness of students with learning problems in utilizing metacognitive strategies during academic tasks, as well as strategies specific to language use. Specifically, they completed a comprehensive literature review across a 20-year span reporting the impact of comprehension strategies for students with learning disabilities. They reviewed studies published before 1999 including experimental or quasi-experimental designs that focused

21

on school-aged students with learning disabilities and included at least one quantitative measure of reading comprehension.

Gertsen et al. (2001) drew several conclusions based on their review. First of all, research suggested that students with learning disabilities tend to not engage in self-talk typically involved with utilizing metacognitive strategies that enhance comprehension and retention of concepts. Instead, they often require explicit teaching of these strategies. Students with learning disabilities tend to have breakdowns in strategic processing and problem solving associated with metacognitive strategies.   However, findings of this study indicated that when students with learning disabilities receive direct instruction on metacognitive strategies, they can successfully utilize these strategies for a number of skills, such as inferring word meanings in context, or expanding verbal or written expression.

Furthermore, Gersten et al. (2001) suggested that when students allocate too much attention to low-level processing than students recruit to complete tasks such as decoding or determining word meanings, comprehension often decreased. Studies from the literature review suggested that students could learn the meanings of words when they encountered them in texts after approximately six to ten exposures; however they could learn them with only two exposures if teachers presented a definition for words prior to reading. The literature review indicated that students with diagnosed learning problems often struggled to use step-by-step strategies and processes during written compositions or steps to take to self-monitor reading comprehension. However, when students learned to use mnemonics for recalling steps to metacognitive strategies, their retention and use of these strategies increased.

22

As studies have indicated (e.g., Cain et al., 2004; Dreher & Zenge, 1990; Gersten et al., 2001) word knowledge and ability to utilize metacognitive and metalinguistic strategies are critical to academic success because they are necessary for reading comprehension, independent word learning, and self-regulation during language-rich processes such as writing. Educators must be aware that students with language and learning disabilities may not independently utilize these strategies and may potentially need direct instruction and practice to apply them successfully. Additionally, all students may require direct instruction of metalinguistic strategies due to the complexities of language in academic texts. Texts may consist of language difficult for even high achievers, simply because contextual information necessary to infer meanings of unfamiliar words may be absent in some circumstances (Beck et al., 2012; Brown, 2010).

Impact of Context on Vocabulary Learning

The cognitive load placed on all learners, regardless of age and skill level, may impact their ability to learn new information when reading if there is inadequate contextual information. Frishkoff et al. (2010) investigated the impact of contextual conditions on word learning. Specifically, they measured the impact of high versus low context word training conditions on cortical responses, known as event-related potentials (ERPs), in the brains of 15 adult male participants (mean age 19.1 years). Researchers conducted pre and posttests to compare the effects of high versus low contextual information on word learning. Pre and posttests included multiple choice items assessing the participants' ability to identify synonyms of 90 target words (assigned to either the low or high context conditions). They also measured ERPs during various semantic tasks over two separate sessions to investigate cortical activity during semantic learning.

23

The researchers conducted two separate experimental sessions. Session 1 consisted of Task 1A, Task 1B, and then a repetition of Task 1A. During Task 1A, referred to as the "meaning generation task", researchers exposed participants to rare words in isolation and asked them to generate meanings of words by typing synonyms. This included 60 rare words and 30 familiar words. In Task 1B, referred to as the "training task", researchers exposed participants to rare words in meaningful sentences and asked them to generate meanings of words by typing synonyms. This included the same 60 words from Task 1A, which were referred to as "trained rare words" for this experiment. Task 1B included two different types of sentences; those with high contextual information and those with low contextual information. Sentences with high contextual information provided context clues relating to meanings of rare words, while those with low contextual information provided little or no information about meanings of rare words. Following Task 1B, participants repeated Task 1A. Participants ERPs were measured during all tasks in Session 1.

Session 2 was conducted two days later, which consisted of a repetition of Task 1A and Task 2. Procedures for Task 1A were the same in Session 2 as they were in Session 1. For Task 2, participants completed a semantic priming task. Participants were exposed to two words and asked to indicate whether or not words were semantically related by pressing buttons on a keypad. Words presented during Task 2 included the 60 rare trained words from Task 1B, 30 untrained rare words, and 30 untrained familiar words. The researchers measured ERPs during both tasks in Session 2.

Frishkoff et al. (2010) conducted pre and posttest measures three times throughout the experiment. They completed the initial pretest before experimental sessions,

24

immediately following Session 1, and immediately following Session 2. Data collection for pre and posttests included a 90 item written multiple choice test during which the participants were to select words closest to the meaning of the target words. The test consisted of 60 trained rare words (the same words used during Task 1A and 1B), and 30 untrained familiar words as controls.

Results of posttest measures revealed significantly more gains from pre to post test for words trained in sentences with high contextual information compared to those with low contextual information; however word learning occurred within both conditions. Delayed posttest measures (conducted after Session 2) indicated that participants retained word learning following initial training (Task 1B).

Researchers reported ERP data taken during meaning generation tasks (Task 1A) following training tasks (Task 1B). These results showed significantly higher cortical responses associated with effortful word retrieval for rare trained words in low context conditions as compared to rare trained words in high context conditions. These cortical responses suggested that word learning in the low context condition was incomplete, indicating more robust learning in the high context condition. ERPs taken during the meaning generation task (Task 1A) in Session 2 indicated shift in responses from frontal to posterior regions. Responses in the frontal regions are associated with early learning, while shifts to posterior regions are associated with long term memory and robust word learning. Therefore, this frontal to posterior shift indicated that word learning occurred in both low and high context conditions. However, ERP data indicated increased responses associated with effortful retrieval for words in the low context condition; indicating that word learning was more robust for words presented with high contextual information.

25

Therefore, this research indicated that for adult learners with no identified disabilities, word learning is significantly greater when it occurs within contexts which provide sufficient semantic information about words. Therefore, there is a chance that we may observe this same phenomenon with younger individuals, suggesting that contexts of texts with minimal semantic information may be difficult for typically developing students as well as students with low vocabularies.

In summary, research has indicated that vocabulary skills and metalinguistic awareness are key predictors of academic success (Cain et al., 2004; Gersten et al., 2001). Unfortunately, the complexities of text language may increase the processing load even for high achievers, resulting in inaccurate or partial learning of new words (Frishkoff et al., 2010). Therefore, we must be aware of effective techniques that enable individuals of all ability levels to improve their vocabulary knowledge and use of metalinguistic strategies in order to assist them across contexts.

Components of Effective Instruction

In order to consider ways to improve students' language skills, it is first necessary to consider how we process communicative behaviors, recall information, and develop semantic representations of words. In the following section, I have discussed the process of word learning, as well as literature documenting the importance of working memory and its ability to process multisensory information and transfer it to long term memory. I have followed this discussion with a review of literature documenting effective methods for improving students' vocabulary and increasing their use of metacognitive and metalinguistic awareness strategies (e.g., Cain, 2007; Graham & Perin, 2007; Zipke, 2012). Metacognitive strategies are those which involve explicit analysis of thought

26

processes; while metalinguistic strategies are analysis strategies that apply specific to language use (Graves, 2006). Both types of strategies apply to successful language use in that they are complex processes requiring students to remember multiple steps while utilizing them. I have therefore concluded my literature review with a discussion of literature documenting effective techniques for enhancing memory of strategy steps, such as the use of mnemonic devices.

Multisensory Information and Memory

Individuals learn words through the process of fast mapping, during which we rapidly attach semantic representations to words based on initial exposures (Alt et al., 2004). This information is then stored for later retrieval, and individuals often fine-tune representations of words during subsequent exposures (Perfetti, 2007). Effective and accurate information processing during fast mapping results in more sophisticated vocabulary knowledge and increased ability to use words in varied contexts. When individuals are exposed to words, semantic features are observed through multiple senses; gaining visual, phonological, auditory and other information. As Baddeley (2003) explained, auditory signals are processed in the brain through what is known as the phonological loop, while visual information is processed through the brain's visuospatial sketchpad. The central control mechanism in the brain coordinates both visual and phonological information to transfer it from working memory to long term memory. Therefore, when considering the most effective practices for word learning, it is relevant to consider the brain's responses to various sensory stimuli. It is also important to discuss the relationship of sensory information and working memory capacity, and the process of storing information in short-term and long-term memory (Sousa, 2011).

27

Sugihara, Diltz, Averbeck, and Romanski (2006) investigated the integration of auditory and visual information in the brains of three rhesus monkeys (one female, two male). This research is relevant in that cortical responses in the brain, such as those observed in this study, are critical to the development of communication and vocabulary skills. Specifically, Sugihara et al. measured the response of neurons in the ventrolateral prefrontal cortex by implanting recording cylinders in the subjects' brains and presenting them with a series of stimuli via pictures, movie recordings, and auditory recordings. They completed a within-subjects design by comparing responses to auditory, visual, and paired audiovisual stimuli. Visual stimuli consisted of static pictures of monkeys or humans, short digital movies of humans or monkeys moving and vocalizing, or objects with accompanying sounds. Vocalizations of humans consisted of common vowel sounds, while vocalizations of monkeys included common sounds such as grunts, barks, screams, or coos. There were a total of 10 auditory stimuli, 10 visual stimuli, and 10 audiovisual stimuli.

Results indicated a higher number of unimodal visual neurons than unimodal auditory neurons. Forty-six percent of the neurons in the ventrolateral prefrontal cortex had a multisensory response. Researchers found that many unimodal neurons previously believed to emit unimodal only responses had robust responses to multisensory information. Sugihara et al. (2006) concluded that audiovisual information relevant to communication reaches the ventrolateral prefrontal cortex in rhesus monkeys. Additionally, they commented that integration of auditory and visual information is crucial for tasks such as object recognition and communicating effectively, and that the multimodal responses in the rhesus monkeys indicates the importance of multisensory

28

integration in completing such tasks. Further, Sugihara et al. (2006) concluded that this multimodal effect may be observed in humans, suggesting that humans communicate and learn object information through multisensory input as well.

While Sugihara et al.'s (2006) research provided information on cortical responses associated with exposure to multimodal communication behaviors, Li et al. (2011) studied the impact of audiovisual stimuli specific to semantic information on brain activity measured by functional magnetic resonance imaging signals. Participants were nine Chinese males (mean age 31.5) with no known disabilities. Researchers presented participants with 80 different pictures associated with two different semantic categories (i.e., young people, old people) under four different conditions; (a) semantically congruent audiovisual stimuli (pictures and spoken words consistent with pictures), (b) semantically incongruent audiovisual stimuli (pictures and spoken words inconsistent with pictures), (c) visual stimuli (pictures only), and (d) auditory stimuli (spoken words only). Visual stimuli consisted of pictures of faces that either fell into the semantic categories of "young people" or "old people". Auditory stimuli consisted of spoken words either stating "young people" or "old people." The congruent audiovisual condition consisted of presentation of a picture with an auditory label consistent with the picture, such as the presentation of a "young" face, paired with the spoken words "young people." The incongruent audiovisual condition consisted of presentation of an auditory label (i.e., young people, old people) inconsistent with the picture. Researchers conducted a within-group design, exposing all participants to all four conditions. Participants were instructed to watch and/or listen to the stimulus and press a button to indicate that they had seen it.

29

Results indicated higher levels of brain activity in the congruent audiovisual condition compared to the other stimuli. There were no significant differences between the two unimodal conditions (i.e., visual, auditory). Specifically, there were significant differences in brain activity in the superior temporal sulcus and middle temporal gyrus in the audiovisual congruent condition compared to the unimodal and incongruent conditions. These two areas of the brain are associated with integration of semantic features through multisensory information. When the brain decodes information, it creates semantic-category specific information which is observable through fMRI signals. Li et al. (2011) found that decoding accuracy was significantly higher in the congruent audiovisual condition compared to the other three conditions. This indicated that cross-modal sensory integration may facilitate increased neural activity and this increased neural activity could, in turn, facilitate increased representation of semantic categories.

While Sugihara et al. (2006) and Li et al. (2011) studied the impact of audiovisual stimuli through studying cortical responses; Delogu, Raffone, & Belardinelli, (2009) investigated the impact of auditory and visual stimuli on serial verbal recall. This study is relevant because it provides information regarding the impact of multisensory information on memory. Verbal recall is an important task to consider in relation to word learning. In order for an individual to repeat a word, they must have adequate representations of its phonological and articulatory features (Perfetti, 2007). The awareness of how words *feel* and *sound* are important components to *knowing* a word. So, while ability to complete serial recall tasks cannot guarantee detailed semantic knowledge, this is an important indicator of how we process information during language learning.

30

Participants in Delogu et al.'s (2009) study included 80 university students ranging from 19 to 35 years of age. Researchers compared the following three conditions: (a) visual, (b) auditory, and (c) bimodal (audio-visual). For all conditions, participants watched and/or listened to a series of 40 items on a 14-inch monitor, and verbally repeated what they saw or heard. Items included 20 verbal and 20 nonverbal items. For the visual mode, participants saw 20 pictures with written words for the verbal items, and 20 pictures of environmental objects for nonverbal items. For the auditory mode, participants heard 20 spoken words for the verbal items, and heard 20 environmental sounds for nonverbal items (e.g., the sound of a doorbell, cat meowing). For the bimodal condition, the participants observed 20 pictures with written words paired with an auditory stimulus stating the word. They also saw 20 pictures of environmental objects with corresponding sounds (e.g., the sound of a cat meowing paired with a picture of a cat). Delogu et al. also divided the nonverbal and verbal conditions into two subgroups: one group that completed articulatory suppression (subvocalizing syllables repeatedly), and one that did not.

Delogu et al. (2009) measured serial verbal recall for each block of stimuli using an adapted version of the digit span task from the Weschler Adult Intelligence Scale-III. Results of the within-groups analysis for modality indicated that bimodal presentation was superior to unimodal presentation. A post hoc analysis indicated that this was only the case in the nonverbal condition. There was also a significant main effect for articulatory suppression, with participants demonstrating longer serial recall with its presence. Auditory presentation was better than visual alone for the verbal condition only. There was no significant main effect for format (verbal versus nonverbal); however

31

a post hoc analysis revealed that there was a higher span for auditory than visual specifically in the verbal condition, and that performance in the nonverbal condition was better than in the unimodal condition. This information suggests that for some types of information, multisensory presentation may be superior to information appealing to only one of the senses. Additionally, these findings suggest that some individuals may retain information more effectively through auditory presentation than visual alone. The fact that subvocalizing syllables repeatedly increased recall span emphasizes the importance of phonological and articulatory representations of words in working memory.

Kibby et al. (2004) investigated Baddeley's (2003) three-part working memory model, comparing the visual and verbal memory skills of students with and without reading disabilities in the third through seventh grades ranging from age 9 to 13. Participants included 20 students with reading disabilities and 20 students without reading disabilities; both groups comparable in age, gender, and intelligence as measured by formal intelligence quotient tests. The researchers compared the participant's performance on a number of verbal and visual memory tasks. The first task assessed verbal working memory, during which the participants had to recall lists of words presented visually. Lists varied in length (i.e., three, five, or seven words), word length (i.e., one or three syllable words), or phonological similarity (i.e., similar or dissimilar). An additional task assessed visual spatial working memory, during which participants were to memorize a list of spatial positions on a matrix. Conditions varied based on the number of positions they recalled (e.g., three, five, or seven positions). Participants completed both verbal and visual spatial working memory tasks with three different levels of central executive load (i.e., none, moderate, high). For the "none" condition,

32

participants completed the verbal or visual spatial task in isolation. For the "moderate" load condition, they completed the verbal or visual spatial task while also repeating a motor task requiring them to reproduce a tapping sequence. For the high load condition, researchers asked them to reproduce the tapping sequence backwards while completing the visual spatial or verbal memory tasks. Finally, researchers measured articulation rate, during which participants repeated pairs of words 10 times as fast as they could.

Results for verbal working memory tasks indicated that performance was better on shorter versus longer lists of words. Performance increased as central executive load decreased. Post hoc analyses indicated that significant differences existed between groups in that children with reading disabilities displayed better performance for lists with short words regardless of phonological similarity. Contrarily, phonological similarity positively impacted the performance of children without reading disabilities on long lists of words. Results also indicated that for visual spatial working memory tasks, children performed better on shorter sequences than longer, and that performance improved as central executive load decreased. No significant differences existed between groups on visual tasks. The researchers drew the conclusions that visual working memory remained intact both for students with and without reading disabilities. However, students with reading disabilities did not improve performance for longer lists of words in the verbal task given phonologically similar words, indicating that this group had weaker verbal working memory skills. While Delogu et al.'s (2009) findings indicated that some individuals recall words more effectively through verbal rehearsal, Kibby et al.'s (2004) findings indicated that others may have stronger visual processing skills.

33

In summary, research has suggested that exposure to both auditory and visual information stimulates increased cortical responses in the brain when compared to multimodal stimuli (Li et al., 2011; Sugihara et al., 2006). The impact of auditory, verbal, and visual information on working memory tasks may vary dependent on individual learner characteristics, such as the presence or absence of learning disabilities (Kibby et al., 2004). Some individuals may be more likely to transfer information from working to long term memory given both auditory and visual information. Others may prefer one mode to the other, however exposing students to multisensory information will increase the chances that children will recall information over time.

Direct Instruction of Individual Words

While we can successfully learn information through exposure, fast mapping alone may result in inaccurate or incomplete representations of more difficult vocabulary (Frishkoff et al., 2010). In relation to word learning, it is important not only to consider how our brains process sensory information, but also specific educational practices proven effective in the area of vocabulary instruction. Therefore, it is important to identify ways we can facilitate more detailed semantic representations of words. Research has suggested that school-aged children learn words most effectively when they receive explicit explanations of word meanings (e.g., Beck & McKeown, 2007; Biemiller & Boote, 2006; Marulis & Neuman, 2010). In this section, I have discussed the literature addressing the impact of direct instruction on vocabulary skills.

Marulis and Neuman (2010) completed a meta-analysis on 67 studies investigating the impact of vocabulary instruction with students in pre-Kindergarten and Kindergarten. They included studies that involved specific vocabulary training,

34

intervention, specific teaching techniques, with empirical designs, had participants with no mental, physical, or sensory handicaps between the ages of birth through nine, conducted with English words, and had identified independent and dependent variables. Results indicated that interventions consisting of explicit instruction, or combinations of explicit and implicit instruction were more effective than those that included implicit instruction alone. Students with at least one identified risk factor made gains similar to those with no identified risk factors; however students from low socioeconomic status made fewer gains than those of middle to high socioeconomic status, indicating that the Matthew Effect was present.

Justice et al. (2005) also investigated the impact of direct instruction of words. They examined the impact of exposure to novel words during storybook reading on the vocabulary skills of 57 Kindergarteners from two urban elementary schools using a pretest- posttest comparison group design. Researchers randomly assigned students to two conditions: those who listened to storybook reading without direct explanations of target words (e.g., elaborated condition), and those who listened to storybook reading with direct explanations of word meanings (e.g., nonelaborated condition). They also identified students as having low vocabularies or high vocabularies according to norm-referenced language assessments. The researchers measured progress with a criterion referenced pretest-posttest during which they asked students to define a list of 30 target words by stating a synonym for the given word.

Each intervention condition consisted of 20 small group reading sessions during which students listened to novel storybook readings. Students in the nonelaborated condition heard storybook readings containing the 30 target words over the 20 sessions.

35

When researchers encountered one of these target words in the storybooks, they provided the children with a definition of the word and an example of the word used in a supported context. The children in the nonelaborated condition listened to storybook readings containing 30 target words as well, but did not hear any direct explanations of word meanings. The researchers found that the students with high vocabularies significantly improved their ability to define words for both the elaborated and non-elaborated conditions; however students with low vocabularies significantly improved their vocabularies in the elaborated condition only, indicating that they did not significantly improve their defining skills given exposures to new words. These findings are consistent with other research that suggests students with high vocabularies can more successfully infer meanings of words through exposure alone compared to students with poor background knowledge (e.g, Cain et al., 2004; Marulis & Neuman, 2010). However it is encouraging that the students with low vocabularies benefitted when provided with direct word instruction.

Biemiller and Boote (2006) conducted a similar two-part study investigating the impact of repeated storybook reading, direct explanation of words, and pre-testing effects with students in primary grades. In the first study, they compared the impact of pretesting, repeated readings, and direct explanations on the word explaining skills of 112 students in Kindergarten, first grade, and second grade. Each grade level had a total of 48 target words which researchers presented during storybook readings. They established equivalency of cohorts using a 40-word vocabulary test during which they asked students to explain what words meant. The researchers used this same procedure to pretest half of the target words.

36

Researchers assigned the students to one of the following four conditions: exposure to target words via reading books two times without explanation of target words, exposure to target words via reading books four times without explanation of target words, exposure to target words via reading books two times with explanation of words, exposure to target words via reading books four times with explanation of target words. For both conditions, the researchers read the same three books. Each grade level had different selected books containing target vocabulary. Results indicated no significant difference in performance on posttests when comparing the words that were pre and posttested. Furthermore, overall there were no significant improvements in defining skills when researchers read books two times versus four times; however when examining the impact on each grade level the students in kindergarten and first grade did improve word explaining skills significantly more when they read books four versus two times. All grade levels improved significantly more when given explanations of target words regardless of the number of times they read books.

In the second part of this study, the Biemiller and Boote (2006) examined the impact of explanations of target words during book reading with 107 students in Kindergarten, first grade, and second grade. They used the same procedures for pre and post-testing as the first study to test students' ability to define all 48 selected target words per grade level. They found that students who listened to storybook readings with brief explanations of target words significantly improved their ability to explain target words over time (a 41% increase) compared to a non-treatment control group. Additionally, during a delayed posttest two weeks later, students who received storybook readings with explanations of words continued to gain vocabulary skills (6% increase over initial post-

37

test), indicating retention and carryover of skills learned. On average, students in the study learned between 8 and 12 words per week.

Active Use and Extended Instruction

Research indicates that students not only benefit from direct instruction of individual words, but benefit specifically from instruction requiring them to actively engage with and use words (e.g., Christ & Wang, 2011; Ewers & Brownson, 1999; Walsh & Blewitt, 2006). Ewers and Brownson (1999) measured the impact of direct instruction (comparing passive versus active participation conditions) on the vocabulary acquisition of Kindergarten students during single storybook readings with 66 students (with a mean age of 6 years, 0 months) from 4 schools in Central New York. School populations were primarily Caucasian with students from middle class backgrounds.

Researchers used a pretest-posttest between groups design to compare the effect of passive participation versus active participation on vocabulary identification. Ewers and Brownson administered the PPVT-R to assign students to either the low or high vocabulary groups. Students in both the low and high vocabulary groups were then assigned to either the passive participation or the active participation conditions. Prior to testing the researchers also administered the Children's Test of Nonword Repetition (CNRep) to measure working memory skills, and administered the Senechal Vocabulary Test-Adapted to measure receptive vocabulary. The CNRep required students to repeat nonwords, and the Senechal Vocabulary Test-Adapted required students to identify pictures representing target words from the storybook.

Ewers and Brownson (1999) read the same storybook to all participants individually, pointing to pictures representing the target words as encountered in the

38

story. In the passive participation condition, the researchers provided the students with a recast, which consisted of direct word explanations consisting of familiar synonyms of the target words. In the active participation condition, the researchers asked the student a "what" or "where" question about the target word. Results indicated that there was a Matthew Effect in that students with high vocabularies made significantly more gains from pretest to posttest (as measured by the Senechal Vocabulary Test-Adapted) than students with low vocabularies regardless of treatment condition. Students in the active participation group made greater gains that those in the passive participation group indicating that asking questions requiring active engagement was more effective that passive recasts. Additionally, a strong correlation between receptive vocabulary (PPVT-R scores) and working memory (CNRep scores) was noted.

Walsh and Blewitt (2006) also investigated the impact of direct instruction, whichencouraged active word use, by studying the impact of questioning style during storybook reading on vocabulary acquisition of 35 3-year olds from middle to upper middle class families in preschool with a pre-test-post-test comparison groups design. The researchers conducted a total of four intervention sessions over a 6-week period. Prior to the intervention, Walsh and Blewitt gave each child the PPVT-III. They randomly assigned the three highest scoring children to intervention groups. They then assigned the next three highest scoring students to one of the three conditions; and continued to use this procedure to place children to ensure group equivalency. As a pretest-posttest measure of vocabulary skills, they administered the New Word Production Test to assess children's ability to label nine target words which appeared in storybooks used during the intervention. Participants received one of three types of

39

instruction during storybook reading; (a) a condition during with the experimenters asked vocabulary eliciting questions (those which required participants to directly use or explain target words in the story), (b) a condition during which the experimenters ask noneliciting questions (those which asked participants a question exposing the children to target words but not requiring them to use or explain the words), or (c) storybook reading with no questions at all.

Results indicated significant effects for treatment conditions as measured by the New Word Production Test, with both questioning conditions having higher vocabulary scores than the non questioning conditions; however noneliciting questions were just as effective as eliciting questions. There was no interaction of prior vocabulary knowledge (as measured by the pretest) on ability to learn new words, indicating that children with low vocabularies improved just as much as children with high vocabularies. However, this absence of the Matthew Effect may be due to the fact that all children in the study were equally unfamiliar with the novel words. Results indicated that children with low vocabularies are capable of learning vocabulary given direct instruction, and so are children with high vocabularies. However, children with higher vocabularies may often be more familiar with academic vocabulary than children with low vocabularies, suggesting that the same rate of progress may not always be present across students with different vocabulary levels. The results are still encouraging because they suggest that early direct instruction in vocabulary can be helpful for students of all ability levels.

Coyne, McCoach, and Kapp (2007) investigated the impact of extended vocabulary instruction during book reading with Kindergarten students in a two part pretest-posttest-posttest study. The first study compared vocabulary learning during

40

extended instruction versus incidental exposures using 31 students randomly selected from a small elementary school in the Northeastern United States. Researchers delivered intervention during three small group book reading sessions 20 to 30 minutes in length. The extended vocabulary instruction condition consisted of the following procedures: having children pronounce target words prior to book reading, rereading and drawing attention to target words when encountered in the story, providing definitions of words, rereading sentences substituting the definition of target words, prompting children to repeat target words, and engaging in deep processing activities (e.g., encouraging children to use words in generative contexts, asking children questions containing the words) following storybook readings. In the incidental exposure condition, children heard the words during the story but did not receive any other instruction. The researchers used a within-groups analysis, randomly assigning the 30 target words to one of the two experimental conditions (i.e., extended vocabulary instruction or incidental exposure).

Pre and posttesting included the PPVT-III, and the following measures specific to the target words: (a) an expressive definitions measure (using a 0- to 2- point rating scale to score responses), (b) a receptive measure of definitions during which children answered two yes/no questions about the word, and (c) a receptive measure of word knowledge in context during which children answered two yes/no questions about the word in relation to the story context. Delayed posttesting was six weeks following the immediate posttest. Results indicated that performance on expressive defining and both receptive measures of target words was significantly higher for words in the extended instruction condition compared to the incidental exposure condition. Performance on

41

expressive definitions of words in the extended instruction condition decreased from posttest to delayed posttest, but no other significant effects of time were identified.

The second study compared the effects of extended instruction and embedded instruction with 32 participants from the same elementary school utilized in the first study using the same storybook reading procedures. The extended instruction condition was the same as in the first study. For embedded instruction, Coyne et al. (2007) reread sentences with target words during storybook readings, and reread sentences substituting definitions. Results indicated students received higher scores for expressive definitions and receptive measures for target words in the extended instruction condition as compared to the embedded instruction condition. No significant effects of time existed for the expressive defining and receptive word knowledge measures, indicating that students retained some expressive and receptive knowledge of words.

However, there was a significant interaction of time and type of intervention for the receptive knowledge of words in context measure, with knowledge of words taught in the extended condition decreasing from posttest to delayed posttest, and knowledge of words in the embedded condition increasing from posttest to delayed posttest. Further analysis revealed that children with higher PPVT-III scores made higher gains than children with lower PPVT-III scores, indicating that children with stronger vocabulary skills benefited more from both extended and embedded instruction than children with weaker receptive word knowledge. However, the analysis also indicated that children with low vocabularies still benefited from intervention, particularly in the extended instruction condition.

42

Christ and Wang (2011) did a meta-analysis of 31 studies examining the impact of instructional methods in early childhood classrooms on the vocabulary skills of preschool-aged children. They included studies that involved empirical research on vocabulary practices, that were conducted in English-speaking general education pre-Kindergarten and Kindergarten classrooms, and that focused on early childhood classroom practices. All studies had publication dates between 1986 and 2008. Christ and Wang found three common practices among studies, (a) purposeful exposure to vocabulary, (b) direct instruction of words, and (c) multiple methods of vocabulary instruction. The authors drew the conclusions that children learned words most effectively when teachers strategically exposed them to words in conjunction with direct instruction and teachers used a number of means to expose children to words, such as read-alouds, illustrations, theme-based units, video presentations, or interactive discussions.

Research indicated that recasting, definitional information, or direct questioning related to target words has a positive impact on word learning. Providing multiple exposures to vocabulary words is also an effective method of word learning, but children typically have greater gains given some type of direct instruction (e.g., labeling, explaining, recasting, questioning). Single exposures to words in conjunction with direct teaching methods can result in the same amount of word learning as exposing children to words multiple times in the absence of direct explanations.

Christ and Wang (2011) also determined that the type of direct instruction impacted the extent to which children learned words. Extended analytical discussions which required children to use words in generative contexts or discuss word meanings

43

resulted in greater word learning than brief exposures and word explanations. Brief explanations, such as recasts or short definitional explanations were still an effective means of teaching words, but the strength of word learning was not as robust.

A final observation of Christ and Wang (2011) was the differential effect of teaching methods on children of high versus low socioeconomic status. An encouraging finding was that children of all abilities and economic backgrounds made gains given purposeful and direct vocabulary instruction. However a discouraging finding was that a Matthew Effect was present in many of the studies in Christ and Wang's meta-analysis. The authors concluded that gaps in vocabulary learning still existed, with students of lower socioeconomic status making fewer gains than those of higher socioeconomic status.

Therefore, research has indicated that children not only benefit from instruction that involves direct explanations of words, (Biemiller & Boote, 2006; Justice et al., 2005), but also intervention that encourages active use of new words. Requiring students to ask or answer questions (Ewers & Brownson, 1999; Walsh & Blewitt, 2006) may be more effective than activities involving passive participation only. A Matthew Effect was present in a number of studies (Christ & Wang, 2011; Walsh & Blewitt, 2006) however students with low vocabularies still incurred benefits when actively engaged in learning words. For all ability levels, learning was greater, both quantitatively and qualitatively, with direct instruction that allowed for extended discussion and use of words when compared to brief explanations of words.

Direct Instruction with Attention to Semantic Features

Not only does active use of words during direct instruction improve word learning (Christ & Wang, 2011); there is also evidence that direct instruction specifically focused on semantic features of words is an effective way to expand children's vocabularies (Beck & McKeown, 2007; Munro, Lee, & Baker, 2008). Munro et al. (2008) conducted research to investigate the impact of hybrid language intervention focused on phonological and semantic features of words on the language skills of 17 preschool and early school-aged children (ranging from 4 years, 8 months to 6 years, 5 months) with specific language impairments using a pre-test-post-test within-groups design. The following pre and posttest measures were used, The Token Test for Children to measure listening comprehension, the Hundred Picture Naming Test to measure expressive vocabulary, The Bus Story to measure oral narrative production, the Preschool and Primary Inventory of Phonological Awareness to measure phonological awareness, and the Developmental Test of Visual Motor Integration to measure drawing skills. The authors also documented performance on two experimental tasks pre- and post-intervention; (a) a word attribute identification task, which included questioning to test knowledge of semantic functions, semantic attributes, associative features, rhyming abilities, and use of alliteration, and (b) a word association task during which the child was to name any words they thought of when given stimulus words. Munro et al. coded children's responses into the following categories: syntagmatic (words typically found near stimulus words within sentences or discourse), paradigmatic (words with a taxonomic relationship such as synonyms or antonyms), or clang (words with rhyming or alliterative features). Syntagmatic, paradigmatic, and clang words tend to indicate strong

45

metalinguistic awareness skills. Munro et al. also documented multiword responses, repeats (repeating the stimulus word), and not related (words irrelevant to the stimulus word) responses, which were all indicative of poor metalinguistic awareness.

Munro et al. (2007) conducted intervention with students individually one time per week for 60 minutes. Intervention consisted of storybook reading, during which researchers followed a script including embedded statements related to phonological and semantic features of words. They also gave the participants' parents the storybooks with a set of instructions, suggesting they complete at-home activities for 10-15 minutes reviewing the phonological and semantic features of words. Results indicated large effect sizes from pre- to post-intervention on all phonological awareness and language measures. Students' measures of knowledge of semantic functions, semantic attributes, rhyme recognition, and alliteration all significantly increased. The only measure that did not significantly increase was the semantic associations. For the word association task, students increased their paradigmatic, syntagmatic, and clang responses; responses indicative of rich metalinguistic awareness. Students decreased their repeated, unrelated, and multiword responses; responses which suggest poor metalinguistic awareness. Researchers concluded that the intervention improved both vocabulary and metalinguistic skills.

Duff, Fieldsend, Bowyer-Crane, Hulme, Smith, and Gibbs (2008) evaluated the impact of a 9-week reading intervention using a pretest posttest AB design (focusing on phonological and vocabulary skills) on 12 8-year olds who failed to respond to phonemic awareness reading interventions as measured by performance on the British Abilities Scale-2nd Edition. The researchers conducted assessments six months prior to

46

intervention to establish a baseline period, immediately before initiating intervention, and immediately following intervention. Assessments included the following measures: letter identification (identifying letters corresponding with letter sounds), spelling of trisyllabic words, decoding as measured by the British Abilities Scale-2nd Edition, Sound Linkage Test of Phonological Awareness (phoneme blending, segmenting, and deletion), CNRep (nonword repetition), Weschler Abbreviated Scale of Intelligence (defining words orally), Action Picture Test (answering questions eliciting various grammatical structures), processing speed, measured by the symbol searching and coding subtests of the Weschler Intelligence Scale for Children, and the Strengths and Difficulties Questionnaire behavioral profile. Additionally, researchers measured defining ability of 30 target words for the intervention using a 3-point rating scale for scoring.

The intervention consisted of two daily individual sessions that were 15 minutes in length. Session A consisted of 5 minutes of reading, 5 minutes of rich vocabulary instruction, and 5 minutes of narrative writing. Session B consisted of 3 minutes reviewing vocabulary from Session A, 5 minutes of phonological awareness training, 3 minutes of reading, and 3 minutes focused on reviewing vocabulary. Phonological awareness instruction taught segmenting, blending, and deletion of initial, medial, and final phonemes. Vocabulary instruction was consistent with the Reading with Vocabulary Intervention framework (Beck et al., 2002). Teaching assistants delivering interventions taught sophisticated tier two words using the following steps: contextualizing words by using them in contexts relevant to books, developing phonological representation of words by having the child repeat words, providing definitions of words, providing examples of word use in alternative contexts, encouraging

47

the child to use the word in alternative contexts, and having the child repeat the word to reinforce phonological representation.

Results of t-tests indicated significant effect sizes for word reading (p < .01), letter-sound knowledge (p<.05), phoneme awareness (p < .01), grammar (p<.05), and word defining (p<.001), with all of these measures improving after intervention. Additionally, analyses indicated no significant changes in pretest scores before and after the 6-month baseline period, indicating that gains in performance did not occur until the intervention period. Therefore, Duff et al. (2008) had similar findings to that of Munro et al. (2011); both found an increase in vocabulary skills after implementing an intervention focused on explicitly teaching semantic features of words.

Beck and McKeown (2007) conducted a two-part study which investigated the impact of interventions focused on detailed semantic features of words. Specifically, the impact of "rich instruction" and "more rich instruction" during story read-alouds on the vocabulary skills of Kindergarten and first grade students was investigated. In Study 1, a between-groups pre-test-post-test design was implemented with 98 students in Kindergarten and first grade in a small urban school district consisting of a high number of students of low socioeconomic status (all students were African American). The participant sample consisted of eight classes; four from each grade level. Two classes per grade level received the intervention condition and the other two functioned as a comparison group. The intervention for the study was Text Talk, during which teachers complete the following steps after story read alouds: provide contextualization of target word, define words, provide examples of word use in other contexts, encourage children to judge use of words in multiple contexts, instruct children to construct their own use of

48

words in varying contexts, and provide additional phonological and meaning information. The teachers in the study implemented Text Talk, which the researchers referred to as "rich instruction", for a 10-week period. Pre and posttests consisted of the PPVT and a researcher-designed vocabulary measure for the target words which included a picture identification task, similar to the PPVT, focused on semantic features of words (e.g., "Which shows someone who is___?"). Results indicated a significant difference in word learning across conditions; those students who received rich instruction learned words better than those without rich instruction.

In Study 2, Beck and McKeown (2007) investigated the difference between "rich instruction" and "more rich instruction" with 76 Kindergarten and first grade students from the same district and drawn from six different classes. They utilized a within-groups design, during which they randomly assigned words to one of two conditions; (a) the "rich instruction" which consisted of the same Text Talk protocol as Study 1, and (b) the "more rich instruction" condition which consisted of the same procedures as "rich instruction", plus additional extension activities. These additional extension activities were one additional review of words following initial presentations, plus two additional cycles during which children received two more exposures to words. Participants received five exposures to words during "rich instruction", and 20 exposures during "more rich instruction." Pre and post-tests included the PPVT and the same type of receptive vocabulary test designed for the target words as in the first study. Testing also included an additional verbal component during which children were to answer two yes/no questions about target words. One question asked about synonymous words (e.g. "Does ____ mean the same thing as ____?"), and one asked about the appropriateness of

49

contextual use words (e.g., "Would it be extraordinary to see a monkey at the zoo?", p. 11). Results indicated that students in both grades improved significantly more on verbal and picture identification tasks in the "more rich instruction" condition when compared to the "rich instruction condition", although children made significant gains from pre to post-test in both conditions

In Study 1, Beck and McKeown (2007) found that "rich instruction", which emphasized semantic features of words, was significantly more effective than no direct instruction during read-alouds. In the Study 2, they found that "more rich instruction" was more effective than "rich instruction" when incorporated into story read-alouds. Results from the first study suggested that drawing children's attention to semantic features of words was more effective than simply exposing children to new words through stories. Additionally, results from the second study suggested that in-depth emphasis on semantic features was more effective than instruction that was less intensive and detail-focused. Beck and McKeown drew the conclusion that the most effective vocabulary instruction may involve detailed analysis and focus on features associated with words.

Zipoli et al. (2010) investigated the impact of several different methods of vocabulary instruction on word learning with 80 Kindergarten students who attended three different urban schools with high percentages of students from low socioeconomic status. For an 18-week period, they conducted a within groups pretest-posttest design examining the impact of three different vocabulary treatment methods implemented during storybook reading; in the first condition, the "no review" condition, the experimenters explained and defined meanings of target words during initial storybook

50

readings and extended instruction (e.g., reintroduction of words, pronunciation of words, repeat definition of words, words presented again in an anchor sentence) immediately following initial storybook readings. Children received no other review of vocabulary words. In the second "embedded review" condition, participants received an explanations and definitions of words during initial storybook readings, completed extension activities, and provided additional review and definitions embedded within additional storybook readings following initial readings. In the third and final "semantic related review" condition, participants received explanations and definitions during initial storybook readings, and then completed additional extension activities emphasizing semantic features. These extension activities were loosely based on semantic feature analysis, and focused on the following features: sound, appearance, sensation, action, association, and location. Students also completed other activities focusing on additional semantic features and word associations. A total of 54 words, with a total of 18 words per condition, were used and randomly assigned to one of the three conditions.

For pre and posttesting, Zipoli et al. (2010) administered the PPVT-III and Expressive One-Word Picture Vocabulary Test. They also administered a Test of Word Knowledge for the target words, during which children were to state a definition of the word and answer a neutral context question (e.g., "What would you do if you were halting?", p. 137). The researchers scored the TWK with a 0- 2-point rubric. Pretest results indicated that the target words were unfamiliar to the participants prior to the intervention. Results of the pre to post-test analysis indicated a significant improvement in all words from pre to posttest.

51

The two conditions that included some type of systematic review (e.g., embedded instruction, semantically related review) resulted in greater student gains than the condition that did not include any systematic review (e.g., no review). The semantically related review condition resulted in greater student gains than the embedded review condition. Zipoli et al. (2010) also documented the average length of time spent teaching words (no review = 10 minutes 36 seconds, embedded review = 13 minutes 13 seconds, semantically related review = 20 min 29 seconds), and derived an instructional efficiency index for each intervention. The embedded review instruction was most efficient, when considering the amount of words learned in conjunction with the length of time spent teaching individual words, due to the fact that semantically related review was more time consuming. This research is consistent with findings of Coyne et al. (2007) that suggested that extended review and active use is more effective than brief word explanations; and is also consistent with other findings suggesting that emphasizing semantic features results in robust word learning (Beck & Mckeown, 2007).

While many of the studies I discussed thus far focused on instruction delivered through storybook reading or in small groups, Apthorp (2006) conducted research to investigate the impact of whole-class vocabulary instruction on both vocabulary and reading comprehension. She used a pretest-posttest comparison group design with 15 third grade classes from two separate Title 1 schools. Site A had a 92% African American student population with 90% of students receiving free or reduced lunches; while Site B had a 74% white student population with between 24 and 35% of the students receiving free or reduced lunches. Apthorp randomly assigned classrooms to one of two treatment conditions: the intervention group during which teachers

52

implemented the Elements of Reading (EOR) Vocabulary Level C lesson, and the control group during which teachers conducted typical classroom vocabulary instruction focused on deep and active processing of words. The EOR intervention consisted of 24 20-minute weekly lessons focused on 7 words per week. The lesson included read aloud and extended talk exercises which provided students with definitional and contextual information about words, opportunities to complete semantic relatedness activities, and review and assessment of written vocabulary comprehension within reading passages. The teachers conducted interventions over the course of an entire school year.

Pretest and posttest measures included the Early Reading Diagnostic Assessment which measured students' ability to name, identify, name synonyms, define words, and state multiple meanings of words. It also included the GMRT which examined the students' knowledge of printed words and reading comprehension (ability to answer questions about passages they read). Results indicated that students at Site A made significantly more gains in the intervention condition in the area of oral vocabulary (as measured by the Early Reading Diagnostic Assessment) compared to the control group. Students at Site A also made significantly more gains in reading vocabulary as measured by the GMRT compared to the control group. There was no significant impact of treatment condition in Site B.

Apthorp (2006) theorized that the intervention had a greater impact in Site A, which consisted of a higher population of students of lower socioeconomic status, because it may have exposed them to richer vocabulary than their typical daily environment. The treatment condition may have had less of a differential effect on the students in Site B, which consisted of fewer students from lower socioeconomic status,

53

because the overall student population may have been exposed to richer vocabulary in their daily environments. It is, however, encouraging that students in Site A incurred benefits because this suggests that students of lower socioeconomic status can improve their oral and reading vocabulary skills given robust vocabulary instruction.

After reviewing the relevant research, I conclude that directly explaining words can not only effectively improve students' learning of words, but is also a necessary supplement to simply exposing them to words through multisensory information (Biemiller & Boote, 2006; Justice et al., 2005). Direct instruction is effective when it includes student-friendly explanations of word learning, and results in more robust word learning when it involves active processing activities, such as those requiring meaning generation, answering direct questions, or analysis of semantic features (Beck & Mckeown, 2007; Coyne et al., 2007; Walsh & Blewitt, 2006). This type of instruction is effective delivered individually, in small groups, and through whole-class instruction (Apthorp, 2006; Duff et al., 2006; Zipoli et al., 2010).

Metacognitive and Metalinguistic Strategy Instruction

Although children can learn up to 5000 words per year in school, they are only likely to learn approximately 300 words through direct instruction (Biemiller & Boote, 2006; Justice et al., 2005). Reading is one of the most effective means of learning vocabulary; however students may struggle to infer meanings of unfamiliar words within texts if they have poor background knowledge or if texts contain inadequate contextual information (Cain et al., 2004; Frishkoff et al., 2010). Vocabulary skills can impact performance on written language tasks as well (Graham & Perin, 2007). Therefore, it is

54

necessary to teach students strategies to both independently learn new words and to expand their written expression.

Research has indicated that students of varying ages and ability levels can effectively learn metacognitive strategies that can increase performance on essay writing tasks (Therrien, Hughes, Kapelski, & Mokhtari, 2009) and enhance metalinguistic skills associated with strong vocabulary skills (Cain, 2007; Zipke, 2012). In the following section I discussed studies which investigated the impact of teaching specific metacognitive or metalinguistic strategies to enhance language expression. This included studies addressing techniques for increasing oral and written expression, as well as those focused specifically on semantic skills and word learning. I have also discussed the use of mnemonic devices and their impact on students' ability to apply metalinguistic strategies.

Use of metalinguistic strategies. Inferring word meanings from contexts of conversations or written material is an important metalinguistic strategy that can facilitate semantic knowledge. Because many students may struggle to utilize this strategy (Cain et al., 2004), it is important to consider research documenting the impact of direct instruction aimed at improving the use of this technique. Cain (2007) investigated the ability of 45 7 and 8 year olds from a middle class community in England to derive meanings of words from context in a reading passage using a posttest only between-groups study comparing the impact of three different treatment conditions. The researcher read short stories containing novel words, provided brief explanations of novel words, and asked the participants to define the new words after reading the passage. The intervention conditions for the study were as follows; (a) students in the FOR group were

55

asked to explain how they determined their definition, then received feedback about their definition, (b) students in the FER group were given feedback on their definitions, then provided with a definition by the adult and asked to explain how the adult determined the definition, and (c) students in the FO group were only given feedback on their definition. Prior to initiating the intervention conditions, Cain administered a modified version of the Neale Analysis of Reading Ability-Revised to measure the children's ability to answer questions about passages read to them, and administered the British Picture Vocabulary Scale-2nd Edition, to measure their ability to identify words. No significant differences existed in the language skills prior to intervention according to these assessments.

During the intervention, the adults reading the stories to children recorded their ability to define words using the following ratings, 0=definition incorrect, 1=definition partially correct, 2-definition correct. Results indicated that the FO group had significantly lower scores for their definitions than children in the FOR and FER group. Cain concluded that providing students instruction that involves direct explanation of words and requires them to actively engage in defining and explaining words is an effective means of vocabulary instruction. She did note that students in all conditions improved their ability to explain meanings of words over the course of the intervention; indicating that they are capable of utilizing context in passages to infer meanings of words with repeated practice with this strategy. Students were also more likely to use a strategy correctly if they consciously explained it, a requirement in the FOR and FER conditions.

In addition to inferring word meanings from context, metalinguistic techniques can also include analysis of multiple meaning words and figurative language or

56

comprehension monitoring.   Zipke et al. (2009) investigated the impact of metalinguistic instruction on metalinguistic skills, comprehension monitoring strategies, and reading comprehension of 46 3rd grade students (mean age 8 years 7 months) of low socioeconomic status. They used a between-groups pretest-posttest design by randomly assigning students to one of two conditions, (a) a control group focused on book reading and discussion, and (b) a treatment group focused on metalinguistic training (e.g., analysis of multiple meanings, words, sentences, and riddles). Four training sessions, each 45 minutes in length were implemented. Pre and posttesting of metalinguistic awareness consisted of a homonym definition task requiring students to explain meanings of homonyms, a sentence ambiguity task requiring students to explain as many meanings of sentences as possible (e.g., "The chicken is ready to eat."), and riddle-detection which required students to identify the correct riddle solution from two choices.

Comprehension was monitored with (a) heteronym pronunciation (words with more than one possible pronunciation), during which examiners documented whether or not children correctly read heteronyms in sentences, (b) anomaly detection during which children had to detect semantic inconsistencies in text (e.g., "Winter was coming soon, when the weather would be very hot."), and (c) miscue correction, where children were to self-correct reading errors. Reading comprehension was obtained by administering the subtest of the Woodcock Reading Master Test-2nd Edition and GRMT-4.

Results showed significant main effects for time of test, homonyms and sentence ambiguity training, indicating significantly higher scores on the posttest compared to the pretest. There was an 83% gain in performance of homonym definitions and a 91% gain in explaining ambiguous sentences, but no significant differences from pre to posttest for

57

riddle detection. Two-way ANOVAs indicated that metalinguistic training facilitated a significant effect on performance during the anomaly detection task. There were significantly higher scores on definitions for homonyms, ambiguous sentences, and anomaly detection for the treatment group compared to the control group. There was a significant interaction between time and treatment on Woodcock Reading Master Test-2nd Edition, indicating that reading comprehension scores increased more for those in the treatment group compared the control group. This same interaction was not present for the GMRT-2 scores, indicating that there was not a significant impact of treatment on this reading comprehension measure.

Zipke et al. (2009) documented participants' correct responses to homonym definitions, ambiguous sentences, and riddle detection questions to calculate success during training scores. Correlation coefficients showed that reading comprehension measures on the pretest predicted success during training scores, indicating that those with higher comprehension scores were more successful in responding to training. However, a negative correlation was present between the GMRT-4 and homonym definitions and the Woodcock Reading Master Test-2nd Edition and story anomaly detection, with poorer readers making greater gains from pre to posttest. Strong correlations existed between all three metalinguistic awareness performance measures, and between the two out of the three comprehension monitoring measures (story anomaly detection and heteronym pronunciation). Strong and significant correlations also existed between reading comprehension, homonyms, and ambiguous sentences detection. Zipke at al. concluded that metalinguistic awareness directly contributes to reading comprehension performance, and that reading comprehension, metalinguistic awareness

58

training can positively impact comprehension in the absence of comprehension strategy training.

Zipke (2012) investigated the impact of metalinguistic training on metalinguistic skills, and the correlation between metalinguistic awareness and language skills of 36 first grade students with typically developing language skills from three different parochial schools in the northeastern United States. She compared the impact of two different conditions; (a) a treatment group focused on metalinguistic training of homonym definitions, ambiguous sentences, and riddles in text, and (b) a control group focused on book reading and discussion. Using random assignment, she assigned participants to one of the two conditions. She conducted two individual sessions for each student for both conditions and analyzed the effects of the treatment condition using a between-groups pretest-posttest design. A prescreening included the PPVT-4, as a measure of receptive vocabulary, and pretesting was completed with the following: expressive vocabulary, sentence structure, phonological awareness, and understanding spoken paragraphs subtests of the CELF-4, and a homonym definition explaining task, an ambiguous sentence explaining task, and a homonym matching task.

Results indicated that students in the treatment group attained significantly more gains than the control group on homonym definitions, with 67% of students n the treatment group making gains. There was not a statistically significant difference in gains in sentence ambiguity between treatment and control groups, with 50% of the control group improving and 67% of the treatment group improving. Correlational analyses indicated that success with homonym knowledge measures related to expressive and receptive vocabulary measures. Listening comprehension scores predicted success

59

with homonyms, and phonological awareness performance predicted success with ambiguous sentences. No correlations between pre to post-test gains were significant, indicating that all students benefitted equally from treatment. In summary, research has indicated that young children are capable of improving their metalinguistic awareness skills given direct strategy instruction (Cain, 2007; Zipke, 2012), and can also successfully utilize strategies to improve independent word learning and reading comprehension (Cain). This indicates that in addition to teaching individual words, explicit strategy instruction is also an important factor in improving students' semantic knowledge and academic performance.

   Use of mnemonic strategies. While direct instruction of metalinguistic strategies can facilitate improved reading and word learning, use of metacognitive strategies incorporating mnemonic devices can facilitate more sophisticated language use during oral and written expression, as well as increased reading comprehension (Graham & Perin, 2007). Graham and Perin conducted a meta-analysis of 123 studies to investigate the impact of various factors in writing instruction on writing performance of adolescents. They included studies based on the following criteria, (a) participants were students in grades four through 12, who (b) attended regular or private schools, (c) researchers used a measure of writing quality, (d) the measure of writing quality was reliable, (e) the design was experimental or quasi-experimental, (f) researchers provided data needed to calculate effect size, and (g) the study itself was relevant to the topic of writing. Graham and Perin indicated a number of factors positively impacted the writing quality of adolescent students. One of these factors was explicit teaching of specific skills, such as grammar, sentence combining, and strategy instruction. Both Self-

60

Regulated Strategies Development (SRSD) and non-SRSD strategy instruction, which is a metacognitive strategy focused on using acronyms as mnemonic devices to recall strategy steps (e.g., revising, editing), had a positive impact on the quality of typically developing students and students with special needs who were struggling with writing skills.

Chalk et al. (2005) conducted research investigating the impact of using mnemonics through SRSD with 15 10th grade students with learning disabilities using a repeated measures design. Researchers drew participants from two resource classes in a school in the southeastern United States who met the following criteria; diagnosis of a learning disability, an IQ between 80 and 115, achievement scores at least two years below grade level, and no presence of other disabilities. Teachers trained by the researchers conducted SRSD with the participants for five 20 to 25 minute sessions during which they taught students to use the DARE (*D*evelop topic sentence, *A*dd supporting details, *R*eject arguments from the other side, *E*nd with a conclusion) mnemonic strategy when writing persuasive essays. Chalk et al. collected data on length (the total number of words) and quality of essays during the following probe conditions: baseline, preskill instruction, modeling, controlled practice, independent practice, post instruction, maintenance, and generalization. Quality was measured using a 6-point scale to rate the following: development, organization, fluency, and conventions. Results indicated significant main effects for condition, with better scores for length and quality in the maintenance and generalization conditions. Linear growth was present across conditions, indicating improvement over time as students received instruction. Thus, Chalk et al.'s study supports using mnemonic strategies for improving writing.

61

Therrien et al. (2009) conducted a study investigating the impact of explicit instruction of a task specific metacognitive strategy on the essay writing skills of 42 7th and 8th grade students with reading and writing disabilities in a rural Ohio school district. The researchers randomly assigned students to either an intervention or control condition. The intervention condition consisted of direct instruction of the 6-step ANSWER (i.e., *A*nalyze the action words, *N*otice the requirements, *S*et up an outline, *W*ork in details, *E*ngineer your answer, *R*eview your answer) mnemonic strategy for essay completion, which included modeling and demonstration, scaffolded practice, and corrective feedback. The control condition consisted of essay writing practice presented with general instructions and examples of the components of a well-written essay. Both conditions lasted for eight sessions over a two-week period during the students 42 minute study hall.

Therrien et al. (2009) administered a pretest one week prior to the eight sessions and a posttest one week after the conclusion of the sessions. For both the pretest and posttest, researchers gave students an essay prompt, and evaluated the students' essays using two different rubrics. One rubric was specific to the ANSWER strategy, for which students received one point for each step in the strategy they completed. The other was a general essay rubric which consisted of a 5-point rating for the following analytical essay traits: ideas and content, organization, voice, word choice, sentence fluency, and conventions. Results indicated that scores for students in the intervention condition were significantly higher GMRT with more pre to post test gains based on the overall scores of both rubrics. Students in the intervention condition also made more gains and had higher scores on analytical essay traits aligned with the ANSWER strategy (i.e., idea and

62

content, organization) compared to the control group. No significant differences existed in scores on analytical traits not aligned with the ANSWER strategy (i.e., voice, word choice, sentence fluency, conventions). When compared to the essays of a randomly selected group of students without disabilities, no significant differences existed between analytical traits aligned with the ANSWER strategy.

However, the group of students without disabilities performed significantly better on traits not aligned with the ANSWER strategy. This indicates that explicit instruction of the 6-step strategy allowed the students with learning disabilities to perform similarly to their peers without disabilities. These findings are consistent with that of Chalk et al., (2005) that suggested that students with learning disabilities can effectively utilize mnemonic strategies to improve written expression. They also suggested that strategy instruction may be an effective means for students with weak language skills to improve writing.

Little et al. (2010) investigated the impact of SRSD on the persuasive writing performance of 13 2nd grade students (ranging from 7 to 8 years old) with emotional disabilities attending school in a rural school district using a multiple-baseline across-participants design. SRSD consisted of explicit instruction on metacognitive strategies focused on goal setting, self-monitoring, self-instruction, and self-reinforcement, and typically involved the use of some type of mnemonic to assist in recall of specific steps in the strategy. The intervention for this study included instruction of the following mnemonics: POW (*P*ick my idea, *O*rganize my notes, *W*rite and say more) and TREE (*T*opic sentence, *R*easons, *E*nding, *E*xamine). The researchers selected participants who scored in the moderate to high risk range on standardized rating scales for antisocial,

63

internalizing, and externalizing behaviors, and below the 37th percentile on the Test of Written Language-3rd Edition (TOWL-3). Seven selected students had externalizing behaviors and six had internalizing behaviors. Intervention consisted of lessons delivered individually by trained graduate students over 7 to 15 30-min individual sessions with the following phases of instruction: developing background knowledge of the strategy, discussing the strategy, modeling, memorizing the strategy, supporting use of the strategy, and independent student performance with the strategy.

Little et al. (2010) conducted three writing probes, (a) baseline, (b) post-intervention, and (c) maintenance. They used a persuasive writing rubric to identify the presence of the following elements: premise, reason, conclusion, and elaboration. They also used a holistic quality 0- to 8-point scale rating based on overall impression of organization, sentence structure, word choice, and grammar. Finally, they documented the number of words in each probe. Social validity was measured through the use of student questionnaires and standardized student and teacher rating scales. Results indicated a strong, positive impact on writing performance and high social validity. Students with internalizing behaviors had 2 to 7 times more persuasive elements, 2.5 to 7 times more words, and increased quality ratings following interventions. Five out of six of the participants maintained these increases on the maintenance probe. Students with externalizing behaviors had three times more persuasive elements, 2.4 to 16 times more words, and quality ratings 3 to 16 times higher following interventions. All students maintained gains on maintenance probes.

Overall, the research on metacognitive strategies has indicated that direct strategy instruction facilitate improved written expression in typical developing students and

64

students with disabilities in early elementary school through high school. Additionally, the use of mnemonic devices in conjunction with these strategies improved retention of specific words or steps associated with the strategy, which in turn improved students' ability to effectively implement them. Direct instruction of metacognitive strategies for memory and written expression are therefore critical elements which can impact students' academic performance.

Summary

Vocabulary skills have a profound impact on academic skills due to the fact that word knowledge can impact reading comprehension skills (Nagy & Scott, 2000). Additionally, performance vocabulary and metalinguistic awareness tasks are correlated with reading comprehension abilities (Dreher & Zenge, 1990; Zipke, 2007). Vocabulary skills not only impact reading comprehension abilities; but also have an effect on one's ability to learn new words while reading. Students who have poor vocabulary and comprehension skills may learn fewer words from texts due to the fact that they struggle to infer meanings of unfamiliar words from text (Cain et al., 2004). However, due to the complexity of text language, students of all ability levels may struggle to infer meanings of unfamiliar words if minimal contextual information is available (Frishkoff et al., 2010).

Students of all ability levels can improve their vocabulary skills given direct instruction of individual words (Apthorp, 2006). Word knowledge learned from explicit explanation of words is even more robust when instruction includes extension activities providing opportunities to use words in generative contexts or analyze semantic features of words (Christ & Wang, 2011; Duff et al., 2008; Walsh & Blewitt, 2006).

65

Metalinguistic and metacognitive strategy instruction, particularly those that utilize mnemonic strategies for memory and that provide multisensory information can also effectively improve the language skills of students both with and without disabilities (Chalk et al., 2005; Delogu et al., 2009; Graham & Perin, 2007; Kibby et al., 2004).

Although there are a number of effective methods for teaching vocabulary, a Matthew Effect often exists in that students entering school with weak background knowledge learn at a rate slower that their typically developing peers (Marulis & Neuman, 2010). It is therefore important to investigate ways to enhance children's vocabulary skills and encourage them to utilize techniques to learn independently. While research has indicated that a number of metalinguistic and metacognitive strategies can effectively improve the expressive language skills of students of all ability levels, a number of questions remain regarding the impact of direct instruction focused on active language use with specific emphasis on semantic features and incorporating the use of a multisensory mnemonic device. Specifically, more information is needed to determine the impact of such a strategy on students' oral and written language skills. Therefore, this study investigated the impact of a multisensory metalinguistic strategy (the EET) focused on students' abilities to describe words in oral and written language and specifically answered the following research questions:

1) What is the impact of whole-class metalinguistic strategy training on the oral language skills of students in elementary school (grade 3)?

2) What is the impact of whole-class metalinguistic strategy training on the written language skills of students' elementary school (grade 3)?

66

3. What is the impact of treatment dosage of metalinguistic strategy instruction on the oral and written language skills of students in elementary school?

CHAPTER III

METHOD

Overview of Design and Method

This chapter provides an explanation of the design and methods of my study. I have organized the chapter to provide descriptions of students, setting, and materials. I have also provided procedures for intervention assignment, descriptions of intervention conditions, training and implementation fidelity procedures, description of data collection procedures, which include the rationale for instruments and interrater analyses. Finally, I have described potential threats to internal and external validity and how I controlled for these threats.

Experimental Design

I used a pretest-posttest-posttest between groups design (Bordens & Abbott, 2005; Weirsma & Jurs, 2005) to evaluate the impact of direct instruction of a metalinguistic strategy on the oral and written expression of the students. The metalinguistic strategy was the "Expanding Expression Tool" (EET; Smith, 2011), which involved teaching students to describe objects or topics using semantic features, and incorporated visual aids and mnemonic devices to recall each feature during oral and written language tasks. The independent variable was the treatment condition. Students assigned to Condition T_1

received instruction on using the EET as a metalinguistic strategy two times a week, while students assigned to Condition T_2 received instruction on using the EET four times a week. Students assigned to the control condition received standard grade level reading enrichment four times per week. The use of a control condition was necessary due to the fact that I was determining the extent to which the EET instruction impacted language skills in school-aged children. Because children in elementary school are constantly learning new vocabulary, it was highly probable that maturation effects could have occurred. Additionally, many teaching strategies implemented in school are likely to improve students' use of vocabulary and metalinguistic strategy use, increasing the chance that improvements in language skills may have been due to other factors than the EET instruction. The use of a control condition was therefore important to ensure that any changes in the dependent variable were more likely due to the independent variable rather than maturation or extraneous variables. I elaborated on this point in my discussions of internal and external validity (Vogt, 2007; Weirsma & Jurz, 2005).

The dependent variables were oral and written language skills. I measured oral expression by tallying "Total Semantic Features Orally" (TFO) and "Total Different Semantic Features Orally" (TDFO) in students' oral descriptions of words. I measured written expression by tallying "Total Semantic Features Written" (TFW) and "Total Different Semantic Features Written" (TFW) in students' written descriptions of words. I measured all four dependent variables at three different times: pretesting (one week prior to the treatment period), posttesting (within one week of completing the intervention period), and delayed posttesting (one month following the intervention period; Biemiller & Boote, 2006; Coyne et al., 2007; Elliot & Gentile, 1986).

69

The use of a pretest established baseline performance for all students prior to treatment. This enabled me to determine not only group differences in quantity at the end of the study, but also determine increases in performance from pretest to posttest comparatively between conditions. The use of a delayed posttest was necessary in measuring generalization of the metalinguistic strategy after the completion of the treatment to determine if students retained the knowledge and skills developed within the treatment period. This is particularly relevant in this study because the purpose of teaching students a metalinguistic strategy is to enable them to utilize strategies independently to learn words and expand expression. Thus, the delayed posttest provided information on the students' ability to retain the ability to use the strategy following intervention. I administered the delayed posttest one month following the study because this was within the range of time utilized in other similar studies (Biemiller & Boote, 2006; Coyne et al., 2007).

Participants

Students in this study included 61 students (36 girls, 25 boys) in third grade who ranged from 7 years, 11 months to 9 years, 2 months at the initiation of the study. All students in the study with IEPs received instruction in the general education classroom for at least 80% of their school day. I have listed participant demographic information in Table 1.

70

Table 1.

Demographic Information on Students.

Condition	Students	Students without identified disabilities	Students with IEPs	Students with 504 plans
T_1	16	11	4	1
T_2	16	15	1	0
Control (C_1+C_2)	29 (14+15)	29 (14+15)	0	0

I used random assignment to assign each of the four classes to one of the three treatment conditions (T_1, T_2, or control), and I have described the demographics of each condition in the following sections. There were a total of 16 students in Condition T_1 (10 girls, 6 boys) ranging from 8 years, 0 months to 9 years, 1 month in age. This class has 19 students on its roster; however six of these students received special education services during the 25-minute enrichment block, and were therefore not included in the study. One student's parent did not return the consent form, so this student participated in the treatment with his class, but was not included in the data collection for the study. This left a total of 12 students in Condition T_1. In order to keep experimental condition sizes consistent and adequate to run necessary data analyses, I randomly selected four students from the control condition (two students from each control classroom) using an online random number generator to add to Condition T_1 so that there were a total of 16 students in this condition. During the block time when the class received the treatment, these four students moved to the classroom that was receiving Condition T_1. Eleven students in this condition did not have any identified disabilities.

One student had a profound bilateral conductive hearing loss. This student had a 504 plan with accommodations relating to a cochlear implant and FM system, but

71

required no other special services. Three students had IEPs for the entire duration of the study. Two of these students had a disability label of "speech and language impaired" and one had a disability label of "autism." One student did not have an IEP during the study, but underwent a case study evaluation and met the criteria for "specific learning disability" with a secondary "speech and language impairment" during the final weeks of the intervention. All students who had IEPs in Condition T_1 were in the general education setting at least 80% of the time.

Condition T_2 had 16 students (10 girls, 6 boys). Students' ages at the initiation of the study ranged from 8 years, 0 months to 9 years, 2 months. The class assigned to Condition T_2 had 19 students on its original roster. Parents of two students did not return the consent forms, so this condition began with the remaining 17 students. The students who did not return consent forms participated in intervention activities during the treatment block but did not participate in any data collection procedures. During the first several weeks of intervention one student withdrew from the study because his individual reading intervention was scheduled during the intervention block time, bringing the number of students in Condition T_2 to 16. Fifteen of these students had no identified disabilities. One of these students had an IEP with a disability label of "speech and language impaired." This student was in the general education setting at least 80% of the school day.

There were two classes selected for the control condition. I will refer to one of these classes as the C_1 class, and the other as the C_2 class; but both of these classes together made up the entire control condition. The C_1 class had a roster of 19 students, with 16 students returning consent forms. From these 16 students, I randomly selected

72

two students to move to Condition T_1. During the time of the day that the teachers were conducting treatments, these two students walked down to the classroom receiving Condition T_1. This left the C_1 class with 14 students who were included in the control condition for the study. The C_2 class had a roster of 19 students, with 17 returning consent forms. Two students from the C_2 class moved to Condition T_1, which left 15 students. The combination of 14 students in the C_1 class plus the 15 from C_2 class resulted in 29 students total in the control condition (16 girls, 13 boys). Ages for these 29 students ranged from 7 years, 11 months to 8 years, 11 months. No students had any identified disabilities.

Although the total participant number for the control condition was higher than Conditions T_1 or T_2, class size across all four classes was similar. Originally, I obtained a large number of students for the control condition to ensure that I would have adequate numbers across all conditions. As I previously explained, I initially had a lower participant number in Condition T_1 so I randomly selected four students from the control condition (C_1 and C_2) and moved them to Condition T_1. I have presented the distribution of students across all conditions in Table 2.

Table 2.

Distribution of Students Across Classrooms.

Class	Students	Additional nonparticipant students	Total students present during intervention block
T_1	16	1	17
T_2	16	2	18
*C_1	14	3	17
*C_2	15	2	17

*Two control classrooms were considered one control condition totaling 29 students. Conditions T_1 and T_2 were separate classes and were separate treatment conditions.

73

I received permission to conduct my study from the principal at the school (see Appendix A) and was granted Institutional Review Board Approval (IRB; see Appendix B). I met with third grade teachers and explained the study, answered their questions, and obtained informed consent to participate in the study (see Appendix C). Teachers completed a demographic form (Appendix D) and demographic information about each teacher is presented in Table 3.

Table 3.

Teacher Demographic Information.

Class	Age	Years teaching	Years teaching in district	Level of education
T_1	44	22	15	Master's
T_2	57	19	19	Bachelor's
C_1	36	3	3	Bachelor's
C_2	32	10	9	Master's

I obtained parental permission during registration at the beginning of the 2013 school year (see Appendix E for Parental Permission Forms). Third grade teachers provided parents with the parental permission forms and I was present during the registration to answer any questions the parents had at that time. Parents of students who registered their student online or after the August registration received permission forms, sent home via their student, the first week of school and returned the forms via their student. I obtained informed student assent by reading an assent to the students and informing them that they could inform their teachers verbally if they did not want to participate (see Appendix F for the assent script).

Setting

I drew my participant sample from an elementary school (Pre-Kindergarten through third grade) in Central Illinois. The school was in a Pre-Kindergarten through

74

12th grade district located in a small rural community with a total student population of 1265 in the 2012-2013 school year. The elementary school had a student population of 328 (89.6% Caucasian) in the 2012-2013 school year. Nineteen percent of the students received a special education service, no students were identified as English Language Learners, 36% of the students were identified as low SES, and the mobility rate was 6%. This school made adequate yearly progress in 2012-2013, and the average class size was 20 students.

There were a total of four third grade classes in the fall of 2013; all of which I assigned to one of the conditions (i.e., Condition T_1, Condition T_2, or control condition). The classroom assigned to Condition T_1 was located on the second floor of the building in the south end. Desks were arranged in pods of three or four students. The two classrooms in the control condition were located across from each other on the first floor in the west end of the building. In the C_1 classroom, desks were arranged in pods of four, while desks in the C_2 classroom were arranged in rows. The Condition T_2 classroom was also located near the west end of the building, down the hall from the control classrooms. Desks were arranged in rows. During the intervention block time, all classrooms had one teacher providing instruction. Students had the same lunch and recess schedules and all children, including those with disabilities, met together for special assemblies or events.

Metalinguistic Strategy Instruction and Materials

As described in Chapter II, EET (Smith, 2011) is a tool used to teach students to use the metalinguistic strategy of describing semantic features of words using a mnemonic device along with visual and tactile cues. The purpose of this strategy is to improve vocabulary and oral and written describing skills. Using the strategy, students

75

are to describe a word by stating the following information: category (group to which an object belongs), function (what the item does, or what you can do with it), physical appearance (descriptions of color, size, and shape), composition (material of which an item is made, or the origin), parts/associated parts (parts of the object or objects associated or used with the item), or location (where an item is found or used). For example, for the word "apple", students may say that it is a kind of fruit (category) that you can eat (function), it can be green, red, or yellow (physical appearance), it is made of fiber (composition), it has a leaf and a stem (parts), and it can be found on a tree (location).

Students were given a visual aid and a chant to help them recall all of these semantic features when they were describing words. Visuals included items such as pictures, a strand of beads with one bead representing each semantic feature, or stickers representing each semantic feature. The strand of beads contained a miscellaneous bead that reminded students to state any other information they knew about a word that did not fit with any of the other semantic features. For example, for the word "apple", described in the above example, the students might say that an apple can taste sweet, because this information is relevant, yet does not fit with any of the other semantic features. This miscellaneous feature/bead is also represented in the other EET visuals.

Aspects of the EET are grounded in research that contends teaching students strategies focused on the features of language can effectively improve academic performance (Smith, 2011). Smith has justified the use of the EET by explaining that teaching students to describe words using semantic features in both oral and written language can improve expressive language skills, and that providing multisensory

76

information during instruction can improve transfer to long-term memory (Sousa, 2011). As I discussed in the literature review, there is strong evidence for the theoretical foundations of the EET development, however there are no published empirical studies that have documented the effectiveness or viability of the EET.

The EET materials that I used in Conditions T_1 and T_2 were: an instruction manual, picture visuals representing semantic features, sticker cues, and the EET mnemonic device. The EET instruction manual (Smith, 2011) contained an introduction and overview of the background, purposes, and intended uses of the EET. It also provided descriptions of the semantic features to target while using this technique. The mnemonic device, also known as "EETchie" (Smith, 2010, p. 8), included a string of beads which provided visual representations of the different semantic features. The pictures and sticker cues were visual representations of the EET semantic features that teachers used as a visual display. The EET materials kit also included several other tools some of which were not used in this study. These were tools that could be used for visual aids (e.g., sticker cues, foam dice representing semantic features, supplemental worksheets); however I limited visual aids to pictures, stickers, and EETchie in order to maintain consistency across Conditions T_1 and T_2.

Description of Experimental Conditions

General education third grade teachers delivered instruction for all experimental groups during whole-class instruction in general education classrooms during a designated reading intervention block time. I randomly assigned each of the four third grade classes to one of three conditions, (a) control condition (n = 29), (b) Condition T_1

77

($n = 16$), and (c) Condition T_2 ($n = 16$). I assigned two classes to the control condition (i.e., C_1 and C_2), one class to Conditon T_1, and one class to Condition T_2.

Control condition. I assigned two classrooms to the control condition. For both control classrooms in the control condition, the teachers conducted the standard third grade enrichment activities aligned with common core standards four days per week (Monday, Tuesday, Thursday, and Friday) in a 25-minute block of time during group instruction. During this time, teachers completed activities focused on skills such as vocabulary development and reading comprehension. Vocabulary development activities included direct explanations of curricular vocabulary from the science, social studies, or reading texts. Reading comprehension activities included reading novels as a class and having class discussions regarding the plot, characters, and setting, as well as answering inferential reasoning questions such as problem solving, comparing or contrasting, or predicting outcomes. Students also worked on providing written responses to reading comprehension questions relating to stories or curricular vocabulary. On Wednesdays, these classes held class meetings during the 25-minute block during which they discussed current issues within the classroom.

Treatment conditions. For Condtions T_1 and T_2, teachers spent 20 to 25 minutes teaching the students to use the EET as a metalinguistic strategy (Smith, 2011) during the reading block time through group instruction in the general education classroom. For Condition T_1, the classroom teacher conducted EET instruction two times per week, and for Condition T_2, the classroom teacher conducted the EET four times per week. In Condition T_1, the teacher held EET instruction on Tuesdays and Thursdays, class meetings (i.e., discussions of issues going on in the classroom and potential solutions) on

78

Wednesdays, and did activities similar to those in the control condition on Mondays and Fridays (e.g., novel reading, independent reading, direct instruction of robust vocabulary). Over the course of the nine-week intervention period, the teacher conducted a total of 18 EET lessons.

In Condition T_2, the classroom teacher conducted EET instruction on Mondays, Tuesdays, Thursdays, and Fridays. She held class meetings on Wednesdays. However, school was not in session the first Monday of intervention, so there were only three weekly sessions during week one. Additionally, during week four there was no EET session on Monday due to a grade level activity. During all other weeks that there were schedule conflicts on one of the scheduled intervention days, the teacher completed an EET lesson on Wednesday to ensure that there were four total weekly sessions. Over the nine-week period the teacher in Condition T_2 conducted a total of 34 EET lessons.

A typical lesson consisted of the teacher introducing the EET strategy and briefly stating its purpose, followed by a demonstration of strategy use relevant to chosen words and their semantic features and opportunities to practice using the strategy to describe words in oral and/or written form. During a portion of the lesson, students had access to some type of visual aid containing the EET semantic features.

While the EET manual (Smith, 2011) contains worksheets, examples of EET use, descriptions of how to use the EET for varying academic activities, and examples of how to use the EET, it does not contain any specific protocol for carrying out EET lessons. Therefore, for purposes of establishing implementation fidelity, I developed specific components to ensure that teachers delivered EET instruction consistently across conditions. These components are as follows: (a) stated name of strategy, (b) explained

79

the rationale, (c) exposed to semantic features, (d) provided models of strategy use, (e) provided opportunities for strategy use, and (f) provided access to mnemonic device cues. I developed these components with feedback from Smith, the author of the EET, and described the components of the EET lessons in the following sections.

Stated name of strategy. To complete this component, teachers told students they would be using the "Expanding Expression Tool,", or EET. Teachers asked students to use the EET to describe some specific topic or item, and told the students to "EET it." (Smith, 2010, p. 8). Teachers could have also told students that they were going to use "EETchie", to describe something. This involved saying, "We are going to use EETchie to talk about ____." The teachers had to directly say "EET it" or "EETchie" to have delivered this component with fidelity.

Explained the rationale. In order to complete this component, the teachers had to directly explain why the students were using the EET strategy. This included any brief statement explaining that students were using this tool to help them come up with ideas, expand their thoughts, organize their ideas, generate more information, recall ideas, or any other relevant reason for using a metalinguistic strategy. Additionally, the teachers also had to remind the students that they could use the EET strategy at any time, regardless of whether or not EET visual cues (e.g., strand of beads, picture cues, stickers) were accessible.

Exposed to semantic features. Teachers were to provide exposure to all semantic features (i.e., category, function, physical appearance, composition, part/associations, location, plus the miscellaneous feature) of the EET in every lesson. Teachers lead students through the EET chant or listed the features verbally during each lesson.

80

Exposure also had to consist of providing visuals of all components. They did this by providing students access to the EET strand of beads, providing students with picture visuals of components to put on their desks, or stickers for written activities.

Provided models of strategy use. For this component to be present, teachers had to ensure students saw an example of someone using the metalinguistic strategy (i.e., using the semantic features taught during EET instruction to describe an object) during the instructional session. This included observing teachers or other students using the EET to describe something verbally. Students could also see written documents containing descriptions of words addressing all semantic features of the EET.

Provided opportunities for strategy use. Teachers provided students with opportunities for active use of the strategy during EET instruction to implement this component with fidelity. This included any type of activity which would require students to engage in active generation of word definitions by explaining the semantic features of the EET. An example would be verbally describing words or concepts, writing lists of semantic features, or writing paragraphs describing words using the EET semantic features.

Access to mnemonic device cues. This component was present if students had access to visual cues to assist them in remembering how to explain each semantic feature. Cues included either pictures of the semantic features or the EET beads. Pictures included a visual display presented up on the board or small pictures of the semantic cues that students could have at their desks.

Instructional Sequence

I trained teachers to implement EET instruction in a training session prior to initiation of the study. The training session lasted 90 minutes and I explained the purpose and theory behind the EET, all of the components and materials and their intended uses. I also completed 20- to 30-minute weekly meetings to discuss lesson plans and to provide additional assistance with implementation. During these weekly sessions the teachers and I met as a group and planned the sequence of instruction and specific content that teachers would discuss while teaching the students to use the EET. In the following sections I have explained the progression of activities and how the teachers implemented all treatment components within the different types of activities.

Across all nine weeks, teachers delivered the first four components of the treatment (i.e., stated name of strategy, explained the rationale, exposed to all semantic features, and modeled strategy use) in a similar fashion. All lessons started with similar statements naming the EET and explaining its purpose (stated name of the strategy and explained rationale), followed by review of the chant or song (exposed to all semantic features). Then, the students all observed an example of someone verbally describing a familiar word using all semantic features of the EET. For example, for the word "cake", the person explaining may state that it is a kind of dessert (category) that you eat (function), that may be round, rectangular or different colors such as brown or white (physical appearance), is made of sugar, flour, milk, and eggs (composition), can have layers, frosting, or candles on it (parts/associated parts), you can find it at a bakery (location), and some people think it tastes delicious and sweet (miscellaneous). The component that varied the most across the nine-week intervention period was

82

"Opportunities for strategy use", because the teachers progressed from easier to more difficult tasks. "Access to mnemonic device cues" varied during the last week, as I explain below.

During the first week, the teachers focused on explaining the EET strategy in detail; explaining how to use the strand of beads or picture cues to remember the semantic features, and explaining the type of information they should say for each semantic feature (e.g., give categorical information such as "fruit" or "animal, describe what an object does or what you do with it). The teachers also verbally modeled descriptions of words using the EET to fulfill the component, "Provided models of strategy use."

During this week, the teachers selected words familiar to the students and had the students describe words as a class. This involved the teacher presenting the EET and asking students to share information for each semantic feature. Each semantic feature was presented in the order it is said during the EET chant. The teacher called on at least one student to share information about each feature before moving on to the next, which fulfilled the component "Provided opportunities for strategy use". During weeks two and three, the teachers continued to allow students to practice describing familiar words verbally. Starting the second week, at least one student volunteer per class described a word using the EET at the beginning of the lesson, in order to fulfill the component, "Provided models of strategy use". In order to implement this component the teacher guided the students through writing information for each semantic feature on paper or dry erase boards; and then asked the students to share information with the class. They first went through this activity as a whole class, and then followed it by having the students

83

work in pairs, small groups, or independently to complete this task of brainstorming information for each feature and writing it down.

During the fourth week, teachers began showing the students how to use the EET to improve their writing. The only component that varied from the first three weeks was "Provided opportunities for strategy use." The teachers allowed students opportunities to use the EET by teaching them to use the EET to write paragraphs. This began with brainstorming information as a whole class and using the worksheet on page 97 of the EET manual (Smith, 2011) as a graphic organizer. During weeks four and five, the teachers first guided the students through the brainstorming process and then guided them on how to transfer information to paragraph form as an entire class. Students followed this by completing the brainstorming as a class and then completed writing paragraphs on their own. During weeks six and seven, students worked on writing paragraphs using the EET stickers as a cue. Each sticker represented one of the semantic features, and the teachers gave the students a strip of stickers in order to help them recall each of the semantic features. The students could start writing their paragraph by sticking the "green-group" sticker on the page and writing information corresponding with that feature next to the sticker. They could then use this process to continue describing the designated word using the remainder of the EET stickers to think of descriptive information. Instruction during week eight continued to focus on writing.

During the treatment period, teachers completed the sixth component, "Provided access to mnemonic device cues" by allowing students to access any of the following visuals during the lesson activity: EET strand of beads, EET stickers, or pictures of the EET. For the first eight weeks of treatment, students had access to one of these visual

84

cues for the entire lesson. However, in order to encourage independence using the EET in settings beyond the treatment sessions, during the ninth week, teachers asked students to write a paragraph without any type of visual aid. The teachers showed the students some type of visual at the beginning of the lesson in order to fulfill the sixth component, and then removed it while the students were given a word and asked to write a paragraph. The teachers encouraged the students to say the chant quietly to help them think of things to write.

Implementation Fidelity

Teachers audiotaped all sessions in Conditions T_1 and T_2 using digital recorders and uploaded all audio files to an online folder. During the teacher training I provided both teachers with the Implementation Fidelity Checklist (Appendix G) and showed them how to use this checklist to ensure that all treatment components were present in each lesson. I trained a second-year graduate assistant in the Department of Communication Sciences and Disorders to complete fidelity checks on every session using the same Implementation Fidelity Checklist the teachers used. The graduate assistant calculated the percentage of metalinguistic strategy components present in the lessons (i.e., stated name of strategy, explanation of rationale, exposure to semantic features, exposure to modeling of strategy use, opportunities for strategy use, and access to mnemonic device cues; see Implementation Fidelity Checklist in Appendix G) and notified me of any missing components weekly. If any component was missing, I notified the teachers and retrained them in the EET instruction.

85

Data Collection Procedures

Establishing equivalency. Prior to the nine-week treatment period, I established equivalency of conditions using the Expressive Vocabulary Test-2nd Edition (EVT-2; Williams, 2007) and Peabody Picture Vocabulary Test-4th Edition (PPVT-4; Dunn & Dunn, 2007), both which are norm-referenced assessments for individuals aged 2 years and 6 months and older. The EVT-2 measured expressive naming abilities. Students either named pictures or stated synonyms of words shown in pictures. The PPVT-4 measured individuals' ability to identify pictures from a group of four. Testing administration ranged from 10 to 20 minutes for both the PPVT-4 and EVT-2 and was conducted in the hallway next to the students' classrooms. I chose the PPVT-4 and EVT-2 as equivalency measures because these two tests provided standardized measures of vocabulary and have established acceptable levels of reliability and validity. The PPVT-4 provided a receptive vocabulary standard score as measured by the students' ability to identify pictures from a group of four. The EVT-2 provided an expressive vocabulary standard score as measured by the students' ability to name items associated with pictures. Receptive and expressive vocabulary were both appropriate measures of equivalency because the metalinguistic strategy I implemented was focused on increasing verbal and written expression; both which are related to expressive and receptive vocabulary skills.

Prior to treatment, students were assessed on the EVT-2 and PPVT-4 either by me or by second-year graduate students from Illinois State University in the Department of Communication Sciences and Disorders. Either I or the graduate students administered the assessments, and I scored all assessments. Graduate students in the Department of

86

Communication Sciences and Disorders receive course instruction in assessments but must be supervised by a speech and language pathologist who is currently certified by the American Speech and Hearing Association; certification is awarded through a Certificate of Clinical Competence. I am a certified speech and language pathologist and qualified to provide supervision.

I provided training in administration of both the PPVT-4 and EVT-2 prior to data collection. This training consisted of describing the administration of each assessment and practice in the administration of the assessment. I provided supervision to the graduate students at least 50% of the administration time. Adequate reliability and validity have been established for the procedures of both the PPVT-4 and the EVT-2. Results of two one-way ANOVAs indicated that there were not significant differences in receptive and expressive vocabulary across any of the three experimental conditions. Descriptive statistics for equivalency testing scores across experimental conditions are presented in Table 4. Results of a one-way ANOVA indicated that there were no significant differences in performance on the EVT-2 across the three experimental conditions, $F(2, 60) = 1.51$, $p = .23$, nor were there any significant differences in scores across conditions on the PPVT-4, $F(2, 60), = .01$, $p = .99$. I therefore drew the conclusion that there were no significant differences between the three experimental conditions in vocabulary skills (T_1, T_2, or control) at the initiation of the study.

Table 4.

Equivalency Testing Descriptive Statistics.

Condition	EVT-2			PPVT-4		
	M	Range	SD	M	Range	SD
T_1 (n=16)	102	77-133	15.66	107.81	81-131	13.92
T_2 (n=16)	102.63	85-122	10.40	107	91-130	10.83
Control (n=29)	106.31	91-126	9.01	107.34	91-129	10.83
TOTAL (N=61)	103.88	77-126	10.85	107.38	81-131	11.52

Pretest, posttest, and delayed posttest measures. I used pretesting (within one week of initiating treatment), posttesting (within one week of the conclusion of the treatment period), and delayed posttesting (one month after the conclusion of the treatment period) to evaluate the effectiveness of the EET instruction, during which I determined the TFO and TDFO in students' oral descriptions of words, and determined the TFW and TDFW in students' written descriptions of words.

Oral descriptions of words. Data collection for students' oral descriptions of words was completed by me and the graduate students in Communication Sciences and Disorders. I trained graduate students in the administration procedures prior to the pretest. For the pretest, posttest, and delayed posttest, we obtained oral descriptions of six different words (see Table 5).

Table 5.

Oral Prompt Words and Corresponding LPT-3 Words.

Oral prompt words	Corresponding LPT-3 words	Semantic category
Apple	Watermelon	Fruit
Television	Computer	Appliance
School	House	Building
Bus	Car	Transportation
Bed	Desk	Furniture
Bumblebee	Butterfly	Insects

I established the protocol for obtaining oral descriptions of definitions based on the data collection procedures described in the Language Processing Test-3rd Edition (LPT-3; Richard & Hanner, 2005) and the EET manual (Smith, 2011). The LPT-3 protocol of the Attributes section involves asking individuals to describe nouns and scoring responses by tallying the types of semantic features listed for a set of 12 words, which can then be converted to a standard score. Smith (2011) also recommended similar procedures in the EET manual, during which she recommended obtaining oral descriptions of nouns and tallying the number of semantic features in responses to measure oral describing abilities.

The directions to the students were developed from the directions included in the LPT-3 and the EET manual (Smith, 2011) and were as follows, "I want you to pretend like I don't know anything about these words. I want you to tell me everything you know about them." For each individual word, we read the following script: 'Tell me everything you know about a ____.' We recorded the students' oral responses verbatim and audiorecorded all testing sessions. Testing administration was administered individually at desks located outside the doors of students' classrooms, and took approximately 15 minutes per student. I supervised 50% of test administrations by graduate students. Additionally, graduate students collected data in pairs, and checked to ensure 100% agreement of recorded responses before continuing to the next participant. I completed administration by myself, but used audiorecordings to check accuracy of recorded responses. All participant responses were recorded on the Oral Description Record form (see Appendix H).

89

I chose six words for the students to describe orally in order to obtain the TFO and TDFO measures. I used the LPT-3 as a reference for choosing these six words because the Attributes section is an established standardized measure used to determine expressive naming skills, and because this section of the test includes nouns commonly familiar to children. In order to ensure that oral language measures included words familiar to children, I selected six words that fell in the same semantic categories as those on the LPT-3. For example, I chose the word "apple" for the EET oral prompt because the LPT-3 also includes the common fruit "watermelon." In Table 2, I have included words from the EET oral prompt and the word from the LPT-3 falling in the corresponding semantic category. TFO and TDFO were all appropriate measures of growth for this study because EET is a metalinguistic strategy focused on teaching children to utilize a mnemonic multisensory strategy to describe words using semantic features.

I determined the TFO and TDFO by coding and tallying verbal descriptions of words using the Semantic Features Rubric (see Appendix I) outlining the seven different semantic features: category, function, physical attributes, composition, parts, location, and associations. To determine the TFO, I tallied the total number of semantic features described across all six words that fit the criteria on the seven different semantic features. I tallied the total number of features named for each of the six words, then added scores for all the words to determine the TFO score. To determine the TDFO, I tallied the total number of different semantic features named. I tallied the total number of different features named for each of the six words and then added responses for each word to get

90

each participant's TDFO score. Any students who were absent on the day of testing completed the procedures in a quiet office the day they returned to school.

The TFO and TDFO measures are different in that the TFO showed the quantity of information, while the TDFO showed the students' ability to state a variety of different semantic features. For example, if a student said "You can eat an apple and you can use it to make pie," both of these responses would be coded as the "function" feature for the TDFO score. The student would get only one point towards their TDFO score for this statement since he/she addressed the "function" feature, and both pieces of information are functions of an apple. On the contrary, for the TFO score, the student would get two points for this statement, because the TFO score was the total number of features named, regardless of the number of statements that addressed the same semantic feature. So, for the above example, he/she would receive credit for both pieces of information (e.g., "You eat it and you use it to make pie,").

Written descriptions of words. I completed all data collection procedures for obtaining students' written descriptions of words. I administered this protocol through group administration. Administration took approximately 15 to 20 minutes per class. I instructed students to write as many semantic features as they could about a word given the prompt, "I want you to pretend like I don't know anything about these words. I want you to write everything you know about them using complete sentences. You will have three minutes to write about each word." For each individual item, I said, "Tell me everything you know about a ____. You have three minutes. You may begin." I established these procedures based on the guidelines in the Attributes section of the LPT-3 (Richard & Hanner, 2005), the recommendations in the EET manual (Smith, 2011), and

91

the writing CBM procedures described in (Weissenburger & Espin, 2005). The task of having students describe everything they know about nouns is consistent with the LPT-3 Attributes section; however the LPT-3 only focuses on having individuals describe nouns orally.

Smith (2011) recommends taking this procedure a step further by asking students to describe nouns in written form. I therefore combined the recommendations from these sources to establish my protocol and written prompt script. Additionally, other established protocols for measuring written language growth, such as the writing CBMs of total words written and correct word sequences commonly include three-minute limits for writing samples (Weissenburger & Espin, 2005). I therefore allowed students a total of three minutes to write about each individual noun I asked them to describe. Any students who were absent on the day of testing completed the procedures in a quiet office when they returned to school.

I completed this procedure with three different nouns per administration (see Appendix J for Written Prompt Script). The pretest, posttest, and delayed posttest included three nouns equivalent in difficulty to the LPT-3 Attributes section (see Table 6). To score the written descriptions of nouns, I tallied the TFW and TDFW using the Semantic Features Rubric outlining the same seven semantic features used for oral responses (i.e., category, function, physical attributes, composition, parts, location, and associations; see Appendix I). To measure TFW, I tallied the total number of semantic features described for each of the three words, and then added these scores to get the TFW score. For the TDFW measure, I tallied the total number of different semantic

92

features named for each of the three words and added these scores to get the TDFW score

for each participant.

Table 6.

Written Prompt Words and Corresponding LPT-3 Words.

Written prompt words	Corresponding LPT-3 words	Semantic Category
Banana	Watermelon	Fruit
Flower	Tree	Plant
Train	Car	Transportation

Semantic features rubric. I referenced both the Attributes section of the LPT-3 as

well as the EET manual to decide which semantic features in the rubric would be used to

code oral and written responses (see Appendix I). My decisions and rationale for

choosing the semantic features in my rubric are explained in this section.

Smith (2011) recommends using category, function, physical attributes,

composition, parts and associations, and location, while the LPT-3 uses function, parts,

color, accessory/necessity, size/shape, category, composition, and location/origin. I made

the decision to use "category", "function", "composition", and "location", because these

were features listed in both the EET manual and the LPT-3. In the EET manual, "origin"

was sometimes included within the "composition" feature. For example, if an individual

were to state that a plant comes from a seed, or that an apple comes from a tree, this

would be considered part of the composition. However, because this type of information

was so similar to the location of an object, I included this type of information within the

"location" feature, which is consistent with the procedures on the LPT-3.

The EET manual included "physical appearance", which included any features

associated with appearance such as size, shape, and color. The LPT-3 included "color"

93

and "size/shape" as two separate features, which were considered one feature in the EET manual. I made the decision to consider "size", "shape", and "color" to be similar types of information, so I recognized this type of information within the same semantic feature, which I called "physical features". The "physical features" group was similar to the "physical appearance" group described in the EET manual, however it included other types of information associated with attributes other than appearance, such as those associated with texture or smell. I made this semantic feature more encompassing in order to give credit for more diverse responses.

In the EET manual, "parts", and "associated parts" were considered one semantic feature. The LPT-3 recognized "parts" and "accessory/necessity" as two distinct semantic features. Both the EET manual and LPT-3 consider the "parts" feature to include information about actual parts of an object; such as the wings on a bird, or the buttons on a pair of pants. The "accessory/necessity" from the LPT-3 included items that are not distinct parts of an object, but rather items associated with or used with an object; such as a worm or a nest for a bird, or a belt for a pair of pants. The EET manual included these "associated parts" in the "parts" feature. Because the parts of an object and those associated with it are, in fact, two distinct features, I made the decision to distinguish the two for the purpose of coding student responses. Therefore, I chose "parts" to include the actual parts of an object, as well as "associations" to include objects associated with the nouns.   For each word in the oral and written tests, I described appropriate responses for each semantic feature prior to scoring student oral or written responses. This rubric also contains information regarding incorrect responses for some of the semantic features.

Internal Validity

Internal validity is the extent to which changes in the dependent variable occur due to the independent variable (Vogt, 2007). Components of my chosen research design enabled me to control for threats to internal validity. One common threat to internal validity is maturation. Maturation effects occur when changes in participant behavior exist due to natural developmental changes rather than the independent variable. I controlled for maturation effects through the use of a control condition to ensure that any improvements in language expression occurred due to metalinguistic strategy instruction.

An additional threat to internal validity is pretesting effects. The potential for pretesting effects existed due to the fact that I administered the same assessments three times throughout the study (pretest, posttest, and delayed posttest). The risk for all data collection measures is that I used the same procedures and words for all administrations. However, keeping the same words for all three administrations of the EET oral and written measures ensured that procedures were consistent across measures. I minimized pretesting effects by ensuring that students did not have any exposure to the test or testing prompts between test administrations. Additionally, the use of a control condition and between groups analysis provided information on the comparison of scores across conditions to determine if significant changes in scores occurred due to pretest effects or treatment effects.

An additional threat to internal validity can occur with inconsistent use of instrumentation (Weirsma & Jurs, 2008). I controlled for effects of instrumentation by conducting training sessions with all individuals administering data collection protocols

95

throughout the study. I also minimized instrumentation effects by calculating interrater reliability coefficients for 30% of all data collection procedures in the experiment.

External Validity

External validity refers to the extent to which one can generalize findings of a study to other individuals or populations beyond the students in the study (Vogt, 2007). One threat to external validity is that of experimental arrangements, which occurs when changes in the dependent variable occur due to artificial arrangements of the study that cannot be generalized to functional situations. I controlled for this threat by implementing treatment in typical general education settings. Other practitioners could feasibly replicate treatment conditions I designed in this study in typical classrooms, which increases the external validity.

Pretesting effects also pose a threat to external validity, because pretesting could potentially clue students to features of the strategy taught during the intervention (Weirsma & Jurs, 2008). I controlled for pretesting by using a control condition to ensure that the effects of the treatment would occur regardless of the presence of a pretest prior to intervention.

An additional threat to external validity could occur due to selection bias. When researchers draw students through nonrandom selection procedures, the chance exists that results may not generalize to other populations. My selection procedures were not random in that I utilized groups which already existed due to classroom assignments. However, I reduced this threat by randomly assigning groups to conditions. Additionally, I used a random number generator to assign extra students to Condition T_1 ensuring random selection, to minimize this threat.

96

CHAPTER IV

RESULTS

Overview of Chapter

In this chapter I have described the methods I have used to for implementation

fidelity, as well as procedures for establishing interrater reliability. I have also described

procedures I used for establishing social validity, and the results of those procedures.

The following sections provide a detailed description of the procedures I used for

analyzing the data, as well as the results of the analyses.

Implementation Fidelity

Implementation fidelity is present when treatments are delivered accurately and

with integrity. Adequate implementation fidelity percentages are necessary to ensure that

any resulting treatment effect occurred due to the intended treatment delivered (Vogt,

2007). Data from the Implementation Fidelity Checklist (see Appendix G) completed by

the graduate student are listed in Table 7. In order to calculate the implementation

fidelity across sessions, I calculated the total number of possible components per session

multiplied by the total number of sessions. The necessary treatment components were:

stated the name of the strategy, explained the rationale, exposed to semantic features,

provided models of strategy use, provided opportunities for strategy use, and provided

97

access to mnemonic device cues. I tallied the total number of components implemented

across all the sessions and divided that by the total number of possible components. I

multiplied that by 100 to get the implementation fidelity percentage. A percentage of at

least 90% was considered adequate (Vogt, 2007). The average implementation fidelity

across a total of 52 interventions sessions was 99% (range = 83-100%), with 98%

(range = 83-100% for Condition T_1 and 99.5% (range = 83-100%) for Condition T_2.

These numbers indicate that implementation fidelity was adequate. I also completed the

Implementation Fidelity Checklist (see Appendix G) on a portion of the audiotaped

sessions by documenting the presence or absence of each treatment component to

establish interrater reliability on fidelity checks.

Table 7.

Implementation Fidelity Percentages for Conditions T_1 and T_2.

Group	Number of sessions	Implementation fidelity percentage	Implementation fidelity percentage range
Condition T_1	18	98	83-100
Condition T_2	34	99.5	83-100
Total	52	99	83-100

I completed duplicate implementation fidelity checks on all sessions during the

first four weeks of treatment in order to ensure that interventions were completed

correctly. For weeks five through eight, I randomly chose two sessions each week from

Condition T_2 and one session from Condition T_1 (i.e., 50% of weekly sessions from each

condition). I did not conduct any duplicate implementation fidelity checks in the final

week. I completed implementation fidelity checks on 12 out of the 18 sessions for

Condition T_1 (i.e., 67% of all Condition T_1 sessions), and 22 out of the 34 sessions in

Condition T_2 (i.e., 65% of all Condition T_2 sessions, for a total of 34 out of the total 52

98

sessions (i.e., 65% of all sessions) across both conditions. I calculated a reliability coefficient by comparing my percentages to the interrater's percentages across the 34 sessions (65% of total sessions conducted) we both checked, dividing the smaller number by the larger number, and multiplying by 100 (Hinkle, Wiersma, & Jurs, 2003). An acceptable reliability percentage was 90% (Vogt, 2007). The reliability percentage was 99.5%, indicating adequate agreement for implementation fidelity.

Interrater Reliability

I established interrater reliability for the TFO, TDFO, TFW, and TDFW measures by having a trained graduate assistant complete duplicate scoring on 30% of all protocols. Training consisted of a one-hour session which included scoring practice and explanation of the Semantic Features Rubric (see Appendix I). This graduate assistant was a second-year student in the Department of Communication Sciences and Disorders at Illinois State University, and had completed coursework in language development and disorders, as well as scoring diagnostic assessments.

For all three test administrations, I randomly selected students for duplicate recording using an online random number generator. I determined the graduate students' point totals for all measures, as well as my point totals for both measures for each individual participant and for the total scores for all measures (i.e., TFO, TDFO, TFW, TDFW) for the entire 30% sample. For each measure, I divided the smaller number by the larger number and multiplied by 100 to determine the percent agreement for both TFO and TDFO (Hinkle et al., 2003). A reliability percentage of 90% was considered acceptable (Vogt, 2007). If interrater agreement fell below 80% for any given participant, I retrained the graduate assistant and rescored that participant.

99

Because agreement for some students fell below 80% after initial scoring of the pretest and posttest, I retrained the graduate assistant following pre and posttest scoring. I conducted one retraining session for the graduate assistant following the scoring of the pretest and the posttest. Because all reliability coefficients for the total scores across the entire 30% of the students was higher than 80% following the retraining, I did not conduct any additional training following the delayed posttest. Reliability coefficients for the pretest, posttest, and delayed posttest are listed in Table 8. The table includes reliability coefficients for the pretest and posttest before and after retraining of the graduate assistant, as well as the reliability percentages for the delayed posttest. These data indicate adequate reliability for all oral and written measures.

Table 8.

Interrater Reliability Percentages.

Measure	Pretest before retraining (range)	Pretest after retraining (range)	Posttest before retraining (range)	Posttest after retraining (range)	Delayed posttest (range)
TFO	92% (75-100%)	93% (75-100%)	89% (78-100%)	90% (78-100%)	94% (76-100%)
TDFO	100% (67-100%)	100% (67-100%)	97% (85-100%)	95% (85-100%)	98% (80-100%)
TFW	87% (67-100%)	95% (83-100%)	91% (62-100%)	98% (73-100%)	94% (63-100%)
TDFW	90% (75-100%)	97% (72-100%)	92% (67-100%)	98% (73-100%)	98% (75-100%)

Social Validity

If teachers feel an intervention has potential for improving behaviors and performance this can increase potential for intervention effectiveness (Gresham, 2004).

100

If teachers do not feel that intervention is appropriate and acceptable, this can negatively impact the fidelity of implementation and intervention. Therefore, I established social validity by having teachers and students complete adapted versions of the Intervention Rating Profile (IRP; Witt & Martens, 1983; see Appendix K for adapted Teacher IRP for Metalinguistic Awareness Instruction and Appendix L for adapted Student IRP for Metalinguistic Awareness Instruction) two times during the course of the study. I received copyright permission to adapt and reprint these scales (see Appendix M for IRP Copyright Permission) and to use three Boardmaker® symbols (Dynavox Mayer-Johnson, 1981-2013) in the adapted student IRP for Metalinguistic Awareness (see Appendix N for Permission to Reprint Boardmaker® Symbols). I also had the teachers complete an open-ended questionnaire one week after starting treatment and one week following the conclusion of treatment (see Appendix O for Teacher Perceptions of Instruction Questionnaire).

Teacher Acceptability Ratings

I assessed the teachers' acceptability ratings using the IRP adapted teacher version (see Appendix K for adapted Teacher IRP for Metalinguistic Awareness Instruction). This was a 12 item 5-point Likert scale which measured teacher perceptions of intervention appropriateness and effectiveness. Teachers assigned to Conditions T_1 and T_2 completed the IRP two times during the study; during the first week of intervention, and the week after the conclusion of the 9-week intervention period. Initial ratings from the teachers in Conditions T_1 and T_2 indicated high levels of acceptability. The teacher in Condition T_2 chose "Strongly agree" to all questions, and the teacher in Condition T_1 chose "Strongly agree" to all items except Item 8, "This method is

101

consistent with those I have used in the classroom setting," to which she chose "Disagree."

Ratings from the teachers one week following the conclusion of treatment also indicated high levels of acceptability. The teacher in Condition T_2 chose "Strongly agree" to all items just as she did during the initial rating. The teacher in Condition T_1 chose "Strongly agree" to all but two items. She chose "Agree" for Item 7, "This method would be appropriate for a variety of students," which was different from the initial rating which was "Strongly agree". She chose "Neutral" for Item 8, "This method is consistent with those I have used in the classroom setting," which was different from her initial rating which was "Disagree."

I also determined social validity by having the teachers assigned to Condtions T_1 and T^2 complete an open-ended written questionnaire during the first week of treatment and one week following the conclusion of interventions. Questions focused on potential for intervention effectiveness, appropriateness of intervention, and any questions or concerns regarding ease of implementation (see Appendix O for Teacher Perceptions Questionnaire). Responses from the initial open-ended questionnaire completed after the first week of intervention indicated positive feedback from the teachers regarding the perceived potential of the intervention and its appropriateness and feasibility.

For Item 1, "How did you feel this intervention addressed students' oral and written language expression skills?", responses from both teachers indicated that they felt that the intervention was appropriate and relevant to the curriculum, although the teacher in Condition T_1 indicated that she had not yet had the opportunity to see significant student growth after only one week. For Item 2, "What feedback do you have regarding

102

the appropriateness of this intervention in addressing language expression in the classroom?" both teachers indicated they felt it was appropriate, but the teacher in Condition T_1 expressed concerns regarding ways to meet the needs of all students' ability levels using this intervention. This teacher (Condition T_1) also expressed this same concern regarding differentiation for Item 3, "What questions/concerns do you have regarding the implementation of this intervention?" She stated that using enrichment time to complete this intervention may limit the time available for other enrichment activities.

The teacher in Condition T_2 stated that she felt the intervention would be beneficial in helping the students provide more information when describing objects. For Items 4, "Do you think that the goals of this intervention will be met?", and 5 "Do you think the students will benefit from this intervention?", both teachers indicated that they felt the intervention would improve students' describing abilities. For Item 6, "How easy were the procedures for implementation?", both teachers reported that the intervention was easy to implement, but the teacher in Condition T_1 stated that it may become more difficult if teachers needed to choose their own vocabulary to target during interventions.

The teachers also completed the open-ended questionnaire one week following the completion of intervention. For Item 1, "How did you feel this intervention addressed students oral and written language expression skills?", both teachers responded positively, stating that the intervention helped students provide more information when describing words. For Item 2, "What feedback do you have regarding the appropriateness of this intervention in addressing language expression in the

103

classroom?", the teacher in Condition T_2 responded similarly to her initial statement that the intervention would be easy to learn with the help of visuals, while the teacher in Condition T_1 stated the intervention seemed more appropriate for students with low or average abilities than students with higher abilities. For Item 3 "What questions/concerns do you have regarding the implementation of the intervention?", the teacher in Condition T_1 indicated that a challenge to implementing the EET would be finding time to schedule it during the day, while the teacher in Condition T_2 stated that students with less semantic knowledge still struggled to describe words using the EET. For Items 4, "Do you think that the goals of this intervention will be met?", and 5 "Do you think the students will benefit from this intervention?", both teachers agreed with these statements. They also stated on Item 6, "How easy were the procedures for implementation?", that the procedures were easy to implement.

Student Acceptability Ratings

I assessed the student's acceptability ratings with a survey I adapted from the IRP (Witt & Martens, 1983; see Appendix L), which was a 10 item 3-point Likert scale which measured the students' perceptions of intervention acceptability and effectiveness. Students in Condition T_1 and T_2 completed the IRP at two times during the study; after the first week of intervention, and within one week of the conclusion of the nine-week intervention period.

I assigned each item a point value to calculate the mean for each item. "Disagree" responses received a point value of 1, "Not Sure" responses received a point value of 2, and "Agree" responses received a point value of 3. For each item, I tallied the total number of students who indicated each response. I then multiplied that number by the

104

point value for that response. For example, if there were 13 students indicating "Agree",

I multiplied 13 by 3 to get a point total of 39 for that response. I then added the point

totals for each response, and divided by the total number of students to get a mean

response score for each item. I added up point totals for each individual item and divided

that by the total responses across the entire survey for all students to get a total mean for

the survey. The highest possible mean for each item was 3.00, so I determined that a

mean of 2.5 would indicate high acceptability for individual items, and for total means. I

also used the point values to calculate the standard deviation for each item. I have

reported means and standard deviations for Conditions T_1 and T_2 survey responses at the

beginning and following intervention in Tables 9 and 10.

Table 9.

Student IRP Ratings for Condition T₁.

Item	One week into treatment		One week following treatment	
	M	SD	M	SD
1. Using the EET is a good way for me to learn and explain new words.	2.81	.40	2.63	.62
2. Most kids would like to use the EET to help them learn and explain new words.	2.50	.52	2.50	.63
3. Using the EET can help me to write and speak more easily.	2.81	.54	2.38	.81
4. I would tell my friends to use the EET to help them learn new words.	2.18	.75	2.31	.79
5. I would tell my friends to use the EET to help them think of things to write.	2.18	.98	2.50	.82
6. I would like to use the EET to learn new words.	2.50	.73	2.31	.87
7. I would like to use the EET to explain what words mean.	2.75	.58	2.43	.81
8. I would like to use the EET to think of things to write.	2.31	.79	2.50	.81
9. Using the EET would be a good way to help me lean and explain difficult words.	3.00	0	2.63	.62
10. Using the EET will help me perform better in school.	2.43	.73	2.50	.82

106

Table 10.

Student IRP Ratings for Condition T₂.

Item	One week into treatment		One week following treatment	
	M	SD	M	SD
1. Using the EET is a good way for me to learn and explain new words.	2.94	.24	2.69	.70
2. Most kids would like to use the EET to help them learn and explain new words.	2.47	.71	2.50	.52
3. Using the EET can help me to write and speak more easily.	2.29	.59	2.81	.40
4. I would tell my friends to use the EET to help them learn new words.	2.35	.61	2.38	.72
5. I would tell my friends to use the EET to help them think of things to write.	2.47	.72	2.81	.40
6. I would like to use the EET to learn new words.	2.24	.97	2.63	.72
7. I would like to use the EET to explain what words mean.	2.52	.62	2.75	.58
8. I would like to use the EET to think of things to write.	2.35	.79	2.81	.54
9. Using the EET would be a good way to help me lean and explain difficult words.	2.71	.47	2.75	.58
10. Using the EET will help me perform better in school.	2.47	.72	2.69	.60

*There were 17 students during the first survey administration. Due to one student dropping from the study, there were 16 students during the second survey administration.

For the survey conducted at the beginning of treatment for Condition T_1, Items 1 (2.81), 2 (2.50), 3 (2.81), 6 (2.50), 7 (2.75), and 9 (3.00) had means above 2.50. High means on Items 1, 2, and 3 indicated that the students had high ratings for items stating

that the EET would help them learn and explain new words, that most students would like to use this strategy, and that the EET could help them write and speak more easily. Additionally, high ratings on Items 6, 7, and 9 showed that students would like to use the EET to learn and explain words. Means for Items 4 (2.18) and 5 (2.18) were less than 2.50, suggesting that students were less likely to tell their friends to use the EET to explain new words or think of things to write. Mean ratings were also below the criterion for Item 8 (2.31), "I would like to use the EET to think of things to write," indicating that all though most students felt that the EET would help them, they were less likely to use it in writing. Item 10 (2.43) also fell just below 2.50, indicating that students were not sure if the EET would help them overall in school.

For the survey administered following treatment for Condition T_1, mean responses for Item 4, "I would tell my friends to use the EET to help them learn new words," were still below the criterion, but increased from 2.18 during the first survey to 2.31 in the second survey. Item 3 fell below the criterion, falling from 2.81 in the first survey to 2.38 in the second survey, indicating that students did not feel strongly that the EET would help them write and speak more easily. Responses fell from 2.50 to 2.31 for Item 6, and from 2.75 to 2.43 for Item 7, indicating that not as many students indicated that they would like to use the EET to help them learn new words or explain what words mean. Responses slightly increased for Item 5, from 2.18 to 2.50, revealing that students were more likely to tell their friends to use the EET in writing following treatment. Responses for Items 1 and 9 decreased slightly, but were still above the criterion for high acceptability, showing that students still felt that the EET could help them learn and explain new and difficult words. Responses for Item 8 increased from 2.31 to 2.50 and

108

for Item 10 from 2.43 to 2.50, indicating that more students felt that they would like to use the EET to think of things to write and that more students felt that the EET would help them perform better in school.

It is important to note that several of the items did not meet the high acceptability criterion in the initial survey but did during the second administration (i.e., 5, 8, and 10); and several of the items that met the criterion during the initial survey did not meet criterion for the second survey administration (i.e., 3, 6, and 7). This indicated that the students' changed their opinions over the course of the intervention period regarding specific aspects and uses of the EET. Additionally, although response means for Items 3, 4, and 7 were below the criterion, they were all less than .2 below, indicating that they were close to meeting the acceptability criterion. The fact that six out of ten items received high acceptability ratings indicated that overall perceptions of the EET were positive among students in Condition T_1.

In Condition T_2, for the survey conducted at the beginning of treatment the only items that met the 2.50 criterion for mean responses were Item 1 (2.94), "Using the EET is a good way for me to learn and explain new words", Item 7 (2.52), "I would like to use the EET to explain what words mean", and Item 9 (2.71), "Using the EET would be a good way to help me learn and explain difficult words." All other means fell below the criterion. This indicated that not as many students felt they would like to use the EET to help them learn and explain words, that they would tell their friends to use the EET to learn words or think of things to write, that they would like to use the EET to learn and explain words or think of things to write, or that they felt the EET would help them to perform better in school. However, Item 2 (2.47), "Most kids would like to use the EET

109

to help them learn and explain new words", Item 5 (2.47), "I would tell my friends to use the EET to help them think of things to write," and Item 10 (2.47), "Using the EET will help me perform better in school", were close to meeting the criterion. For the survey conducted after treatment for Condition T_2, the only item with a response mean falling below the criterion was Item 4 (2.38), "I would tell my friends to use the EET to help them learn new words."

For Condition T_2, many of the mean responses for survey items did not meet the high acceptability criterion for the first survey administration (i.e., 2, 3, 4, 5, 6, 7, 8, and 10), indicating that some of the students were unsure regarding the usefulness of the EET. However, three items (i.e., 2, 5, and 10) had response means (2.47) that did not fall far below the mean, indicating that many of the students reacted positively to the EET after the first week of intervention. Mean responses from the second survey indicated that acceptability increased after the intervention period, indicating that students reacted positively towards the EET and they were more likely to use the EET following the intervention. These results showed that exposure to the EET throughout the course of intervention may have improved their perceptions of this strategy. Students in Condition T_2, had received the higher intervention dose (four times per week), and they rated the EET higher than those in Condition T_1 who received the lower intervention dose (two times per week). There is a chance that the higher dose of exposure resulted in improved perceptions of the EET. It is also interesting to note that Item 4, "I would tell my friends to use the EET to help them learn new words," did not meet the 2.50 criterion for Condition T_1 and Condition T_2 at the beginning or following treatment. This suggested that treatment activities or the rationale the teachers provided did not adequately

110

emphasize using the EET to learn words, rather than just using it to describe words in oral and written language. Overall, survey results indicated that students had positive perceptions of the EET.

Data Analysis

In the following sections, I have described the methods of analysis I used to determine and report the results of the study. I conducted analyses of between-subjects effects and within-subject contrasts using the Statistical Package for Social Sciences (SPSS) version 20. For both analyses I completed a repeated measures multivariate analysis of variance (MANOVA; Green & Salkind, 2005). A MANOVA is a multivariate analysis containing at least one independent variable with multiple levels, and multiple dependent variables (Green & Salkind, 2005).

I began this section by explaining the assumptions of a MANOVA, and describing the measures I took to determine if my data met these assumptions, as well as considerations I made to address any violations of these assumptions. I followed by reporting descriptive statistics. After presenting descriptive statistics, I reported Wilk's Λ for between-subjects effects and within-subjects effects. Wilk's Λ is a conservative measure of effect size; and is robust enough to detect interactions and main effects when assumptions have been violated (Cohen, 1988). I also reported interactions and main effects from between-subjects and within-subjects MANOVAs, as well as η_p^2. The η_p^2 statistic explains the amount of variance in the dependent variables explained by the independent variables. Values of at least .01 are considered small effect sizes, values of at least .06 are considered medium effect sizes, and values of at least .14 are considered large effects sizes (Cohen, 1988).

111

Independent variables for repeated measures MANOVAs were the treatment condition, which had three levels (i.e., T_1, T_2, and control) and testing time which also had three levels (i.e., pretest, posttest, and delayed posttest). The four dependent variables were: Total Semantic Features Orally (TFO), Total Different Semantic Features Orally (TDFO), Total Semantic Features Written (TFW), and Total Different Semantic Features Written (TDFW).

Assumptions of Multivariate Analysis of Variance (MANOVA)

Descriptive statistics and tests of homogeneity were used to determine if assumptions for a repeated measures MANOVA were met. The first assumption of a MANOVA was that the dependent variables were normally distributed for each level of the independent variables. I first examined kurtosis and skewness statistics to determine the distribution of scores for dependent variables. A kurtosis statistic of three indicates a normal distribution (Cohen, 1988). Kurtosis statistics of less than three indicate platykurtic distribution, with a flatter than normal distribution that has widespread values around the mean. Kurtosis statistics of more than three indicate a leptokurtic distribution, with a sharper than normal distribution that has fewer scores than normal around the mean (Cohen, 1988). Skewness statistics were also calculated using SPSS to determine the normality of the distribution for dependent variables across testing times and conditions. A skewness statistic of more than zero indicates a right-skewed distribution, with the majority of values on the left of the mean with extreme values to the right, while a skewness statistic of less than zero indicates a left-skewed distribution, with the majority of values on the right with outliers on the left. A skewness statistic of zero indicates a normal distribution (Cohen, 1988).

112

All kurtosis statistics for all dependent variables indicated a platykurtic distribution, demonstrating that dependent variables overall had a flatter than normal distribution (see Table 11). The kurtosis statistics for all dependent variables across conditions at the pretest were less than three, indicating a playtkurtic distribution (see Table 12). At posttest, kurtosis for the TFO scores for Condition T_2 were greater than three, which indicated a leptokurtic distribution; while all other dependent variables across conditions indicated a platykurtic distribution at posttest (see Table 13). At delayed posttest, kurtosis statistics for all dependent variables across conditions were less than three, indicating a platykurtic distribution (see Table 14).

Table 11.

Descriptive Statistics for Dependent Variables Across Testing Times (N=61).

Measure	Time	Min.	Max.	M	SD	Kurtosis
TFO	Pretest	10	77	28.49	13.81	2.75
	Posttest	14	82	38.02	15.10	.09
	Delayed Posttest	16	74	38.97	14.11	-.08
TDFO	Pretest	6	24	14.75	4.12	-.40
	Posttest	8	29	17.25	4.68	-.30
	Delayed Posttest	10	29	18.30	4.41	-.61
TFW	Pretest	2	17	8.56	3.57	-.57
	Posttest	5	25	12.66	4.76	-.65
	Delayed Posttest	5	24	13.38	4.67	-.36
TDFW	Pretest	2	11	6.15	2.17	-.52
	Posttest	3	14	8.18	2.86	-.93
	Delayed Posttest	4	16	8.46	2.46	.43

113

Table 12.

Descriptive Statistics for Dependent Variables at Pretest Across Conditions.

Measure	Condition	N	Min.	Max.	M	SD	Std. Error Mean	Kurtosis	Skewness
TFO	T_1	16	10	54	28.38	11.53	2.88	.39	.52
	T_2	16	13	51	27.88	10.98	2.75	-.41	.57
	Control	29	12	77	28.90	16.50	3.06	2.7	1.67
TDFO	T_1	16	7	21	14.44	4.08	1.02	-.79	-.18
	T_2	16	9	23	15.13	4.37	1.09	.90	.27
	Control	29	6	24	14.72	4.13	.77	.24	.25
TFW	T_1	16	2	15	9.00	3.72	.93	-.73	-.28
	T_2	16	2	12	6.38	2.92	.73	-.53	.27
	Control	29	4	17	9.52	3.40	.63	-.61	.41
TDFW	T_1	16	2	9	5.75	1.98	.50	-.77	-.25
	T_2	16	2	8	5.06	1.84	.46	-.91	.12
	Control	29	2	11	6.97	2.18	.41	-.43	-.06

Table 13.

Descriptive Statistics for Dependent Variables at Posttest Across Conditions.

Measure	Condition	N	Min.	Max.	M	SD	Std. Error Mean	Kurtosis	Skewness
TFO	T_1	16	14	58	39.00	14.44	3.61	-1.28	-.23
	T_2	16	26	82	40.63	13.94	3.48	4.40	1.99
	Control	29	17	74	36.03	16.25	3.02	-.19	1.08
TDFO	T_1	16	8	27	18.50	5.11	1.28	-.29	-.29
	T_2	16	11	29	17.94	4.22	1.06	2.26	1.04
	Control	29	9	27	16.17	4.58	.85	-.45	.54
TFW	T_1	16	6	20	14.31	4.53	1.13	-1.24	-.18
	T_2	16	7	25	12.88	4.54	1.14	2.08	1.28
	Control	29	5	21	11.62	4.68	.87	-1.00	.41
TDFW	T_1	16	3	13	9.25	2.98	.74	-.47	-.62
	T_2	16	5	14	8.38	2.50	.63	.27	.68
	Control	29	3	14	7.48	1.97	.37	-.92	.34

115

Table 14.

Descriptive Statistics for Dependent Variables at Delayed Posttest Across Conditions.

Measure	Condition	N	Min.	Max.	M	SD	Std. Error Mean	Kurtosis	Skewness
TFO	T_1	16	21	69	40.00	13.52	3.38	-.25	.36
	T_2	16	27	72	43.69	13.04	3.26	-.28	.79
	Control	29	16	74	35.79	14.62	2.71	.66	.87
TDFO	T_1	16	11	29	19.56	4.99	1.25	-.17	-.20
	T_2	16	12	26	19.00	3.95	.99	-.27	-.18
	Control	29	10	26	17.21	4.20	.78	-.70	.30
TFW	T_1	16	6	22	13.63	4.76	1.20	-.32	.42
	T_2	16	9	23	15.38	4.06	1.02	-.47	.36
	Control	29	5	24	12.14	4.68	.87	.87	1.13
TDFW	T_1	16	4	16	9.00	2.92	.73	1.10	.77
	T_2	16	5	14	9.06	2.65	.66	-.79	.17
	Control	29	4	11	7.83	1.97	.37	-.80	.11

At pretest, skewness statistics for TFO means for all conditions were greater than zero, indicating that the distribution was right-skewed (see Table 12). Skewness statistics for TDFO Condition T_2 and the control condition also indicated right-skewed distributions at the pretest, while for T_1 the skewness statistic was less than zero, indicating a left-skewed distribution. Distribution for TFW was right-skewed for Condition T_2 and the control condition, and left-skewed for Condition T_1. TDFW scores were right-skewed for Condition T_2 and left-skewed for Condition T_1 and the control

116

condition. Skewness statistics on the posttest indicated that all dependent variables were right-skewed for Condition T_2 and the control condition, and all dependent variables for Condition T_1 were left-skewed (see Table 13). At delayed posttest, skewness statistics for TDFO for Conditions T_2 and T_1 showed left-skewed distributions, while all other scores for dependent variables across conditions showed right-skewed distributions (see Table 14).

Because skewed distributions occur due to outliers, I manually examined the data to determine the presence of extreme scores. This examination revealed that two students were significantly higher on the TFO measure on the pretest and posttest compared to the rest of the participant sample. However, I kept these students in the data sample because they were not outliers on other measures, indicating that their skills overall were not higher than the rest of the sample.

In addition to kurtosis and skewness statistics, I conducted a one-sample Kolmogorov-Smirnov test (see Table 15) to determine normality of the distribution. This test indicated scores were not significant, indicating that the distribution of scores across all dependent variables for all testing times did not vary significantly from a normal distribution.

Table 15.

One-Sample Kolmogorov-Smirnov Test.

	Asymp. Sig. (2-tailed)			
Testing Time	TFO	TDFO	TFW	TDFW
Pretest	.96	.99	.78	.72
Posttest	.88	.96	.80	.80
Delayed Posttest	.98	.90	.83	.69

117

Because the results of the Kolmorogorov-Smirnov test did not indicate significant scores for any of the dependent variables, I determined that the scores on dependent variables fell close enough to a normal distribution to conduct a repeated measures MANOVA. Although there were two outliers on the TFO measures, I determined that, the *F* statistic was robust enough to indicate significant differences in dependent variables across testing time and conditions because no other extreme scores were present. Because kurtosis and skewness statistics did not fall in the range indicated for a normal distribution, I chose to include the Wilk's Λ statistic in my repeated measures MANOVA because it is effective in detecting interactions and main effects if some assumptions have been violated (Cohen, 1988).

The second assumption of a MANOVA was that scores for dependent variables for each participant were independent from one another (Green & Salkind, 2005). This assumption was met through my research design and method. The third assumption of a MANOVA is that covariance and variance of scores across dependent variables is homogenous (Green & Salkind, 2005). There were no known covariates for this study. To determine homogeneity of variance, I ran Levene's test of equality of error variances (see Table 16).   Results indicated no significant scores for any dependent variable, indicating that the variances were homogeneous and the third assumption was met.

118

Table 16.

Levene's Test of Equality of Error Variances.

Testing Time	Measure	F	df1	df2	Sig.
Pretest	TFO	1.28	2	58	.29
	TDFO	.23	2	58	.80
	TFW	.54	2	58	.59
	TDFW	.38	2	58	.69
Posttest	TFO	.54	2	58	.59
	TDFO	.95	2	58	.39
	TFW	.53	2	58	.59
	TDFW	.91	2	58	.41
Delayed Posttest	TFO	.01	2	58	1.00
	TDFO	.32	2	58	.73
	TFW	.15	2	58	.86
	TDFW	1.08	2	58	.35

Results

Descriptive Statistics

Descriptive statistics across all experimental conditions are presented in Table 11, and descriptive statistics for individual experimental conditions are presented in Tables 12, 13, and 14. Across all students (N=61) the means and standard deviations for the dependent variables were as follows: for TFO at pretest $M = 28.49$, $SD = 13.81$, at posttest $M = 38.02$, $SD = 15.10$, and at delayed posttest $M = 38.97$, $SD = 14.11$. For TDFO at pretest $M = 14.75$, $SD = 4.12$, at posttest $M = 17.25$, $SD = 4.68$, and at delayed posttest $M = 18.30$, $SD = 4.41$. TFW at pretest was $M = 8.56$, $SD = 3.57$, at posttest $M = 12.66$, $SD = 4.76$, and delayed posttest $M = 13.38$, $SD = 4.67$. TDFW at pretest was $M = 6.15$, $SD = 2.17$, at posttest $M = 8.18$, $SD = 2.86$, and delayed posttest $M = 8.46$, $SD = 2.46$.

Pretest statistics for dependent variables across conditions are reported in Table 12. For the TFO measure at pretest, Condition T_1 $M = 28.38$, $SD = 11.53$, Condition T_2 $M = 27.88$, $SD = 10.98$, and for control $M = 28.90$, $SD = 16.50$, with the control condition being the highest and Condition T_2 the lowest. For TDFO at pretest, Condition T_1 $M = 14.44$, $SD = 4.08$, Condition T_2 $M = 15.13$, $SD = 4.37$, and control $M = 14.72$, $SD = 4.13$, with Condition T_2 being the highest and Condition T_1 being the lowest. For TFW at pretest, Condition T_1 $M = 9.00$, $SD = 3.72$, Condition T_2 $M = 6.38$, $SD = 2.92$, and control condition $M = 9.52$, $SD = 3.40$, while for TDFW at pretest Condition T_1 $M = 5.75$, $SD = 1.98$, Condition T_2 $M = 5.06$, $SD = 1.84$, and for control $M = 6.97$, $SD = 2.18$. For both written language measures (i.e., TFW, TDFW), the control condition was the highest and Condition T_2 was the lowest.

Descriptive statistics for the posttest are presented in Table 13. Statistics for TFO at posttest were: Condition T_1 $M = 39.00$, $SD = 14.44$, Condition T_2 $M = 40.63$, $SD = 13.94$, and control condition $M = 36.03$, $SD = 16.25$, with Condition T_2 as the highest and the control condition as the lowest. For TDFO at posttest, Condition T_1 $M = 18.50$, $SD = 5.11$, for Condition T_2 $M = 17.94$, $SD = 4.22$, and control condition $M = 16.17$, $SD = 4.58$, with Condition T_1 as the highest and the control condition as the lowest. Statistics for TFW at posttest were: Condition T_1 $M = 14.31$, $SD = 4.53$, Condition T_2 $M = 12.88$, $SD = 4.54$, and control condition $M = 11.62$, $SD = 4.68$, with Condition T_1 as the highest and the control condition as the lowest. Statistics for TDFW at posttest were: Condition T_1 $M = 9.25$, $SD = 2.98$, Condition T_2 $M = 8.38$, $SD = 2.50$, and control condition $M = 7.48$, $SD = 1.97$ with Condition T_1 as the highest and the control condition as the lowest.

120

Delayed posttest statistics across conditions are presented in Table 14. The TFO statistics at delayed posttest were as follows: Condition T_1 $M = 40.00$, $SD = 13.52$, Condition T_2 $M = 43.69$, $SD = 13.04$, and control condition $M = 35.79$, $SD = 14.62$, with Condition T_2 as the highest and the control condition as the lowest. For TDFO at delayed posttest, Condition T_1 $M = 19.56$, $SD = 4.99$, Condition T_2 $M = 19.00$, $SD = 3.95$, and control condition $M = 17.21$, $SD = 4.20$, with Condition T_1 as the highest and the control condition as the lowest. For TFW at delayed posttest, Condition T_1 $M = 13.68$, $SD = 4.76$, Condition T_2 $M = 15.38$, $SD = 4.06$, and control condition $M = 12.14$, $SD = 4.68$. Statistics for TDFW delayed posttest were: Condition T_1 $M = 9.00$, $SD = 2.92$, Condition T_2 $M = 9.06$, $SD = 2.65$, and control condition $M = 7.83$, $SD = 1.97$. For both of the written language measures (i.e., TFW, TDFW), Condition T_2 was the highest and the control condition was the lowest.

Interactions of Testing Time and Condition

Between-subjects analysis indicated that a significant interaction was present for Wilk's Λ across testing time and condition, with $F(16, 102) = 2.70$, $p < .01$, $\eta_p^2 = .30$, which indicated that significant differences existed somewhere between conditions across the testing times (see Table 17).

121

Table 17.

Wilk's Λ for Multivariate Tests of Between-Subjects Effects.

	Effect	F	Hypothesis df	Error df	Sig.	η_p^2	Observed Power
Between-Subjects	Condition	398.55	.42	110.00	.91	.03	.19
Within-Subjects	Time	20.06	8.00	51.00	.00*	.76	1.00
	Time * Condition	2.70	16.00	102.00	.00*	.30	.99

*p < .01

Within-subjects effects (see Table 18) also indicated there was a significant

interaction between time and condition for Wilk's Λ, with F (16, 345.86) = 2.90, p < .01,

η_p^2 = .09, indicating that differences in dependent variables existed when comparing

differences across testing times for each condition.

Table 18.

Wilk's Λ for Multivariate Tests of Within-Subjects Effects.

	Effect	F	Hypothesis df	Error df	Sig.	η_p^2	Observed Power
Within-Subjects	Time	14.88	8.00	226.00	.00*	.35	1.00
	Time * Condition	2.90	16.00	345.86	.00*	.09	.98

*p < .01

Results of within-subjects contrasts (see Table 19) indicated that no significant

interaction was present across time and condition for the dependent variables measuring

oral language, with TFO at F (2, 58) = 2.04, p=.14, η_p^2 = .67 and with TDFO at

F (2, 58) = 2.18, p=.12, η_p^2 = .07, which indicated that no significant differences existed

in the oral measures when comparing each condition across each testing time. A

significant interaction was present across time and condition (see Table 19) for the

122

dependent variables measuring written language, with TFW at $F(2, 58) = 13.16$, $p < .01$, $\eta_p^2 = .31$, and TDFW at $F(2, 58) = 9.50$, $p < .01$, $\eta_p^2 = .25$. These results showed differences existed in the writing scores when comparing each condition across each testing time.

Table 19.

MANOVA Tests of Within-Subjects Contrasts.

Within-Subjects Effect	Measure	Mean	Standard Error	df	F	Sig.	η_p^2	Observed Power
Time	TFO	35.59	1.67	1	35.53	.00*	.38	1.00
	TDFO	16.96	.50	1	48.50	.00*	.46	1.00
	TFW	11.65	.47	1	103.59	.00*	.64	1.00
	TDFW	7.6	.26	1	64.95	.00*	.53	1.00
Time * Condition	TFO	35.59	1.67	2	2.04	.14	.67	.40
	TDFO	16.96	.50	2	2.18	.12	.07	.43
	TFW	11.65	.47	2	13.16	.00*	.31	1.00
	TDFW	7.6	.26	2	9.50	.00*	.25	.98

* $p < .01$

Impact of Condition

Wilk's Λ across conditions was not significant at $F(8, 110) = .42$, $p = .91$, $\eta_p^2 = .03$, indicating that there were not significant differences across any dependent variable with scores for pretest, posttest, and delayed posttest combined for dependent variables (see Table 17). Results of the repeated measures MANOVA of between-subjects effects revealed no significant main effects across conditions for each of the four dependent variables (see Table 20). For TFO, $F(2, 58) = .51$, $p = .60$, $\eta_p^2 = .02$, for TDFO, $F(2, 58), = 1.04$, $p = .36$, $\eta_p^2 = .04$, for TFW, $F(2, 58), = .62$, $p = .54$, $\eta_p^2 = .02$, and for TDFW, $F(2, 58) = .49$, $p = .62$, $\eta_p^2 = .02$. This indicated that when all scores from each dependent variable from each testing time (i.e., pretest, posttest, and delayed

123

posttest) were combined together, no significant differences existed between conditions

for any dependent variable.

Table 20.

MANOVA Tests of Between-Subjects Effects.

Between-Subjects Effect	Measure	Mean	Standard Error	df	F	Sig.	η_p^2	Observed Power
Condition	TFO	35.59	1.67	2	.51	.60	.02	1.30
	TDFO	16.96	.50	2	1.04	.36	.04	.22
	TFW	11.65	.47	2	.62	.54	.02	.15
	TDFW	7.6	.26	2	.49	.62	.02	.13

*p < .01

Impact of Testing Time

Wilk's Λ was significant across testing time for between-subjects effects, with

$F (8, 51) = 20.06$, p < .01, $\eta_p^2 = .76$, showing that significant differences existed in

dependent variables across testing times (see Table 17). Wilk's Λ across time for

within-subjects effects indicated a significant main effect, with $F (8, 226) = 14.88$,

p < .01, $\eta_p^2 = .35$ indicating significant differences in dependent variables across testing

time. Results of a repeated measures MANOVA of within subjects contrasts revealed

significant main effects were present for all dependent variables across time, with TFO at

$F (1, 58) = 35.53$, p < .01, $\eta_p^2 = .38$, TDFO at $F (1, 58) = 48.50$, p < .01, $\eta_p^2 = .46$, TFW

at $F (1, 58) = 103.59$, p < .01, $\eta_p^2 = .64$ and with TDFW at $F (1, 58) = 64.95$, p < .01,

$\eta_p^2 = .53$. This indicated that when scores for all conditions were combined for all

dependent variables, significant differences in means occurred somewhere when

comparing the pretest, posttest, and delayed posttest scores (see Table 18).

124

Between-Subjects Comparisons Across Condition and Time

Because the repeated measures MANOVA of between-subjects effects indicated that a significant interaction existed between condition and testing time, I conducted post hoc Bonferroni multiple comparisons across condition and time. This analysis compared the group means for dependent variables across conditions at each testing time. These results are displayed in Tables 21 and 22.

Table 21.

Bonferroni Multiple Comparisons Across Condition and Time for TFO and TDFO.

Measure	Testing Time	Condition Comparison	Mean Difference	Standard Error	Sig.
TFO	Pretest	T_1-T_2	.50	4.96	1.00
		T_1-Control	-.52	4.37	1.00
		T_2-Control	-1.02	4.37	1.00
	Posttest	T_1-T_2	-1.63	5.38	1.00
		T_1-Control	2.97	4.74	1.00
		T_2-Control	4.59	4.74	1.00
	Delayed Posttest	T_1-T_2	-3.69	4.93	1.00
		T_1-Control	4.21	4.34	1.00
		T_2-Control	7.89	4.34	.22
TDFO	Pretest	T_1-T_2	-.69	1.48	1.00
		T_1-Control	-.29	1.30	1.00
		T_2-Control	.40	1.30	1.00
	Posttest	T_1-T_2	.56	1.64	1.00
		T_1-Control	2.33	1.44	.34
		T_2-Control	1.77	1.44	.68
	Delayed Posttest	T_1-T_2	.56	1.64	1.00
		T_1-Control	2.36	1.36	.26
		T_2-Control	1.79	1.36	.57

125

Table 22.

Bonferroni Multiple Comparisons Across Condition and Time for TFW and TDFW.

Measure	Testing Time	Condition Comparison	Mean Difference	Standard Error	Sig.
TFW	Pretest	T_1-T_2	2.63	1.19	.09
		T_1-Control	-.52	1.05	1.00
		T_2-Control	-3.14	1.05	.01*
	Posttest	T_1-T_2	1.44	1.66	1.00
		T_1-Control	2.69	1.47	.21
		T_2-Control	1.25	1.47	1.00
	Delayed Posttest	T_1-T_2	-1.75	1.61	.84
		T_1-Control	1.49	1.42	.90
		T_2-Control	3.24	1.42	.08
TDFW	Pretest	T_1-T_2	.69	.72	1.00
		T_1-Control	-1.22	.64	.18
		T_2-Control	-1.90	.64	.01*
	Posttest	T_1-T_2	.88	.99	1.00
		T_1-Control	1.77	.88	.14
		T_2-Control	.89	.88	.94
	Delayed Posttest	T_1-T_2	-.06	.86	1.00
		T_1-Control	1.17	.76	.38
		T_2-Control	1.23	.76	.32

*p < .05

Between-subjects comparisons across conditions and time for TFO and TDFO. No significant differences existed between conditions for the oral language measures (i.e., TFO, TDFO) at the pretest. For TFO, the T_1 to T_2 comparison indicated a mean for Condition T_1 ($M = 28.38$) that was higher than the mean for Condition T_2 ($M = 27.88$), p = 1.00. The T_1 to control comparison showed that Condition T_1 ($M = 28.38$) had a mean TFO lower than for the control condition ($M = 28.90$), p = 1.00, while the T_2 to control comparison showed a mean TFO for Condition T_2 ($M = 27.88$) that was lower than for the control condition ($M = 28.90$), p =1.00. For TDFO, the T_1 to

126

T_2 comparison indicated that the mean TDFO for Condition T_1 ($M = 14.44$) was lower than for Condition T_2 ($M = 15.13$), p = 1.00, and the T_1 to control comparison showed that the mean TDFO for Condition T_1 ($M = 14.44$) was lower than for the control condition ($M = 14.72$), p = 1.00. The T_2 to control comparison showed that Condition T_2 ($M = 15.13$) had a mean higher than for the control condition ($M = 14.72$), p = 1.00.

No significant differences occurred between conditions for the oral measures at the posttest. For TFO, the T_1 to T_2 comparison indicated a mean for Condition T_1 ($M = 39.00$) that was lower than for Condition T_2 ($M = 40.63$), p = 1.00. The T_1 to control comparison indicated that the mean TFO for Condition T_1 ($M = 39.00$) was greater than for the control condition ($M = 36.03$), p =1.00, and the T_2 to control comparison showed that the mean TFO for Condition T_2 ($M = 40.63$) was greater than for the control condition ($M = 36.03$), p =1.00. For TDFO, the T_1 to T_2 comparison showed that the mean TDFO for Condition T_1 ($M = 18.50$) was higher than Condition T_2 ($M = 17.94$), p = 1.00, and the T_1 to control comparison indicated that the mean TDFO for Condition T_1 ($M = 18.50$) higher than for the control condition ($M = 16.17$), p = .34. The T_2 to control comparison indicated that the mean TDFO for Condition T_2 ($M = 17.94$) higher than the control condition ($M = 16.17$), p = .68.

There were no significant differences between conditions for TFO at the delayed posttest. The T_1 to T_2 comparison showed that the mean TFO Condition T_1 ($M = 40.00$) lower than for the Condition T_2 ($M = 43.69$), p =1.00, and the T_1 to control comparison showed that the mean TFO for Condition T_1 ($M = 35.79$) higher than for the control condition, p = 1.00. The T_2 to control comparison showed that the mean TFO for Condition T_2 ($M = 43.69$) higher than for the control condition ($M = 35.79$), p = .22. The

127

T_1 to T_2 comparison showed a TDFO mean for Condition T_1 ($M = 19.56$) that was higher than the mean TDFO for Condition T_2 ($M = 19.00$), p = 1.00. The T_1 to control comparison showed a mean TDFO higher for Condition T_1 ($M = 19.56$) compared to the control condition ($M = 17.21$), p = .26, and the T_2 to control comparison showed a mean TDFO higher for Condition T_2 ($M = 19.00$) compared to the control condition ($M = 17.21$), p = .57. These results indicated that for the oral language measures of TFO and TDFO, no significant differences existed between conditions prior to treatment, following treatment, and one month following the cessation of treatment in Conditions T_1 and T_2.

Between-subjects comparisons across conditions and time for TFW and TDFW. The results of the Bonferroni multiple comparisons indicated that there were significant differences between conditions at the pretest for the written language measures (i.e., TFW, TDFW). The T_1 to T_2 comparison for TFW at the pretest revealed a mean TFW higher for Condition T_1 ($M = 9.00$) compared to Condition T_2 ($M = 6.38$), p = .09, and the T_1 to control comparison showed a mean TFW lower for Condition T_1 ($M = 9.00$) compared to the control condition ($M = 9.52$), p = 1.00. The T_2 to control comparison was significant; with the mean TFW for Condition T_2 ($M = 6.38$) lower than for the control condition ($M = 9.52$), p < .05. The same pattern was present for TDFW means at the pretest. The T_1 to T_2 comparison showed a mean TDFW higher for Condition T_1 ($M = 5.75$) compared to Condition T_2 ($M = 5.06$), p = 1.00, and the T_1 to control comparison showed a mean TDFW lower for Condition T_1 ($M = 5.75$) than the control condition ($M = 6.97$), p = .18. The T_2 to control comparison was significant, with

128

a mean TDFW for Condition T_2 (M = 5.06) that was lower than for the control condition (M = 6.97), p < .05.

No significant differences existed between conditions for the written language measures at the posttest and the delayed posttest. For the posttest, the T_1 to T_2 comparison indicated the TFW mean for Condition T_1 (M = 14.31) higher compared to Condition T_2 (M = 12.88), p = 1.00, and the T_1 to control comparison showed the mean for Condition T_1 (M = 14.31) was higher than for the control condition (M =11.62), p = .21. The T_2 to control comparison was no longer significant at the posttest, with the mean TFW for Condition T_2 (M = 14.31) higher than for the control condition (M = 11.62), p = 1.00. For TDFW, the T_1 to T_2 comparison showed a mean TDFW for Condition T_1 (M = 9.25) higher than for Condition T_2 (M = 8.38), p = 1.00, and the T_1 to control comparison showed a mean TDFW higher for Condition T_1 (M = 9.25) compared to the control condition (M = 7.48), p = .14. The T_2 to control comparison was also no longer significant at posttest, with a mean TDFW higher for Condition T_2 (M = 8.38) compared to the control condition (M = 7.48), p =. 04.

For the delayed posttest, the T_1 to T_2 comparison revealed that the mean TFW for Condition T_1 (M =13.63) was lower than for Condition T_2 (M = 15.38), p = .84, the T_1 to control comparison showed a mean TFW for Condition T_1 (M = 13.63) that was higher than the control condition (M = 12.14), p = .90, and the T_2 to control comparison showed a mean TFW for Condition T_2 (M = 15.38) that was higher than for the control condition (M = 12.14), p = .08. Comparisons for TDFW followed a similar pattern; the T_1 to T_2 comparison showed a mean TDFW lower for Condition T_1 (M = 9.00) compared to Condition T_2 (M = 9.06), p = 1.00, a T_1 to control comparison with a mean TDFW higher

129

for Condition T_1 ($M = 9.00$) compared to the control condition ($M = 7.83$), p = .38, and T_2 to control comparison which showed a mean TDFW higher for Condition T_2 ($M = 9.06$) than for the control condition ($M = 7.83$), p = .32. The results of this analysis showed that TFW and TDFW means were significantly lower for Condition T_2 compared to the control condition at the pretest, although no significant differences existed between TFW and TDFW scores between Conditions T_1 and T_2, and between Condition T_1 condition and the control condition. Following treatment, the posttest indicated that there were no longer significant differences between conditions on the written language measures, indicating that students in Condition T_2 increased their scores to that of the control condition. There were no significant differences between conditions at the delayed posttest.

Pairwise Comparisons Across Testing Time

Because there was a significant main effect across testing time for all dependent variables, I conducted post hoc Bonferroni pairwise comparisons across testing times to determine where differences across testing times existed (see Table 23). Results of the analysis indicated significant differences for all dependent variables when comparing posttest means to pretest means for the entire participant sample ($N = 61$). There were significant differences in the posttest-pretest comparison for the TFO measure, p < .01, with the posttest ($M = 38.02$) higher than the pretest ($M = 28.49$), as well as for the TDFO measure, p < .01, with the posttest ($M = 17.25$) higher than the pretest ($M = 14.75$). There were also significant differences in the posttest-pretest comparison for the written language measures. For TFW, p < .01, with the posttest ($M = 12.66$) higher than the pretest ($M = 8.56$), and p < .01 for TDFW with the posttest ($M = 8.18$) higher than the

130

pretest (M = 6.15). These results indicated significant increases from pretest to posttest

for the oral language measures (i.e., TFO, TDFO), as well as the written language

measures (i.e., TFW, TDFW).

Table 23.

Bonferroni Pairwise Comparisons Across Testing Times.

Measure	Time Comparison	Mean Difference	Standard Error	Sig.
TFO	Posttest-Pretest	10.17	1.59	.00*
	Delayed Posttest-Posttest	1.27	1.41	1.00
	Delayed Posttest-Pretest	11.45	1.92	.00*
TDFO	Posttest-Pretest	2.77	.50	.00*
	Delayed Posttest-Posttest	1.05	.49	.11
	Delayed Posttest-Pretest	3.83	.55	.00*
TFW	Posttest-Pretest	4.64	.57	.00*
	Delayed Posttest-Posttest	.78	.53	.44
	Delayed Posttest-Pretest	5.42	.53	.00*
TDFW	Posttest-Pretest	2.44	.38	.00*
	Delayed Posttest-Posttest	.26	.34	1.00
	Delayed Posttest-Pretest	2.70	.34	.00*

* $p < .01$

The Bonferroni pairwise comparisons did not indicate significant differences in

the delayed posttest-posttest comparison for any of the dependent variables. For the TFO

measure, p = 1.00, with the delayed posttest (M = 38.97) higher than the posttest

(M = 38.02). For TDFO, p = 1.00, with the delayed posttest higher (M = 18.30) than the

posttest (M = 17.25). For the written measures, p = 1.00 for TFW, with the delayed

posttest (M = 13.38) higher than the posttest (M = 12.66), and p = 1.00 for TDFW, with

the delayed posttest (M = 8.46) higher than the posttest (M = 8.18). These results

indicated there were not significant changes in the dependent variables when comparing

the mean scores of the entire participant sample (N = 61).

131

Comparisons from the delayed posttest-pretest comparison revealed significant differences for all dependent variables. For the oral measures, p < .01 for TFO, with the delayed posttest (M = 38.97) higher than the pretest (M = 28.49), and p < .01 for TDFO, with the delayed posttest (M = 18.30) higher than the pretest (M = 14.75). For the written measures, p <. 01 for TFW, with the delayed posttest (M = 13.38) higher than the pretest (M = 8.56), and p <. 01 for TDFW, with the delayed posttest (M = 8.46) higher than the pretest (M = 6.15). The results of the Bonferroni pairwise comparisons indicated that group means for the entire participant sample (N=61) significantly increased from the pretest to the posttest, and were maintained from posttest to delayed posttest.

Paired Samples Comparisons Across Time

Because results of the repeated measures MANOVA of within-subjects contrasts indicated a significant interaction of time and condition, and because the Bonferroni pairwise comparison did not show the differences in means of each separate condition across testing time, I ran paired samples t-tests across time for each separate condition (i.e., T_1, T_2, Control).

Paired samples comparisons across time for Condition T_1. Results for the paired samples analyses of Condition T_1 across time are displayed in Table 24. For the Condition T_1 (n=16), results for pretest-posttest measures comparison revealed significant difference for TFO, t (15) = 4.60, p < .01, with the posttest (M = 39.00) higher than the pretest (M = 28.38), as well as for TDFO, t (15) = 3.96, p < .01, with the posttest (M = 18.50) higher than the pretest (M = 14.44). For TFW, t (15) = 5.63, p < .01, with the posttest (M =14.31) higher than the pretest (M= 9.00), and for TDFW, t (15) = 5.12,

132

p < .01, with the posttest ($M = 9.25$) higher than the pretest ($M = 5.75$). These results

indicated significant increases from pretest to posttest for all dependent variables.

Table 24.

Paired Samples t-tests Across Time for Condition T₁ (n=16).

Measure	Time Comparison	Mean Difference	t	df	Sig.
TFO	Posttest-Pretest	10.63	4.60	15	.00*
	Delayed Posttest-Posttest	1.00	.32	15	.75
	Delayed Posttest-Pretest	11.63	3.28	15	.01**
TDFO	Posttest-Pretest	4.06	3.96	15	.00*
	Delayed Posttest-Posttest	1.06	1.02	15	.32
	Delayed Posttest-Pretest	5.13	5.09	15	.00*
TFW	Posttest-Pretest	5.31	5.63	15	.00*
	Delayed Posttest-Posttest	-.69	-.83	15	.42
	Delayed Posttest-Pretest	4.63	5.29	15	.00*
TDFW	Posttest-Pretest	3.50	5.12	15	.00*
	Delayed Posttest-Posttest	-.25	-.38	15	.71
	Delayed Posttest-Pretest	3.25	4.43	15	.00*

* p < .01
**p < .05

There were no significant differences in the delayed posttest to posttest

comparison for any of the dependent variables. For TFO, t (15) = .32, p = .75, with a

delayed posttest (M =40.00) higher than the posttest (M =39.00), and for TDFO

t (15) = 1.02, p = .32, with the delayed posttest (M =19.56) higher than the posttest mean.

For TFW, t (15) = -.83, p = 42, with the delayed posttest (M =13.63) less than the posttest

(M = 14.31), and for TDFW, t (15) = -.38, p = .71, with the delayed posttest (M =9.00)

less than the posttest (M = 9.25). For the oral measures, the scores increased from

posttest to delayed posttest, and for the written measures decreased from posttest to

delayed posttest; however none of these changes were statistically significant.

133

Delayed posttest to pretest comparisons were significant for all dependent variables. For the TFO measure, $t(15) = 3.28$, p < .05, with the delayed posttest ($M =40.00$) higher than the pretest ($M =28.38$), and for TDFO, $t(15) = 5.09$, p <.01, with the delayed posttest ($M =19.56$) higher than the pretest ($M = 14.44$). For TFW, $t(15) = 5.29$, p < .01, with the delayed posttest ($M = 13.63$) higher than the pretest ($M =9.00$), and for TDFW $t(15) = 4.43$, p < .01, with a delayed posttest ($M = 9.00$) higher that the pretest ($M = 5.75$).

Paired samples comparisons across time for Condition T$_2$. Results of the t-tests across time for Condition T$_2$ (n=16) are listed in Table 25. Comparisons of posttest to pretest scores were significant for all dependent variables. For TFO, $t(15) = 4.08$, p < .01, with the posttest ($M = 40.63$) higher than the pretest ($M = 27.88$), and for TDFO, $t(15) = 2.50$, p < .05, with the posttest ($M = 17.94$) higher than the pretest ($M = 15.13$). For TFW, $t(15) = 5.29$, p < .01, with the posttest ($M = 12.88$) higher than the pretest ($M = 6.38$), and for TDFW, $t(15) = 6.55$, p < .01, with the posttest ($M = 8.38$) higher than the pretest ($M = 5.06$). These results indicate significant increases in all dependent variables from pretest to posttest.

Table 25.

Paired Samples t-tests Across Time for Condition T$_2$ (n=16).

Measure	Time Comparison	Mean Difference	t	df	Sig.
TFO	Posttest-Pretest	12.75	4.08	15	.00*
	Delayed Posttest-Posttest	3.06	1.05	15	.31
	Delayed Posttest-Pretest	15.81	3.90	15	.00*
TDFO	Posttest-Pretest	2.81	2.50	15	.03**
	Delayed Posttest-Posttest	1.06	1.03	15	.32
	Delayed Posttest-Pretest	3.88	2.81	15	.01**
TFW	Posttest-Pretest	6.50	5.29	15	.00*
	Delayed Posttest-Posttest	2.50	2.71	15	.02**
	Delayed Posttest-Pretest	9.00	9.44	15	.00*
TDFW	Posttest-Pretest	3.31	6.55	15	.00*
	Delayed Posttest-Posttest	.69	1.26	15	.23
	Delayed Posttest-Pretest	4.00	6.46	15	.00*

* p < .01
**p < .05

No significant differences occurred in the delayed posttest to posttest comparison for both of the oral measures, TFO and TDFO, and for one of the written measures, TDFW. For TFO, t (15) = 3.90, p = .31, with the delayed posttest (M = 43.69) higher than the posttest (M = 40.63), and for TDFO, t (15) = 1.03, p = .32, with the delayed posttest (M = 19.00) higher than the posttest (M =18.50), and for TDFW, t (15) = 1.26, p = .23, with the delayed posttest (M = 9.06) higher than the posttest (M = 8.38). For TFW, t (15) = 2.71, p < .05, with the delayed posttest (M = 15.38) higher than the posttest (M =12.88). These results indicated slight increases in the TFO, TDFO, and TDFW measures from posttest to delayed posttest; although none of these increases were significant. The TFW measure significantly increased from posttest to delayed posttest, which was not apparent on the Bonferroni pairwise comparisons of scores for all students combined.

135

Results of the delayed posttest to pretest comparisons revealed that all measures were higher on the delayed posttest than the pretest. For TFO, t (15) = 3.90, p < .01, with the delayed posttest (M = 43.69), higher than the pretest (M = 27.88), and for TDFO t (15) = 2.81, p < .05, with the delayed posttest (M = 19.00) higher than the pretest (M = 15.13). For TFW t (15) = 9.44, p < .01, with the delayed posttest (M = 15.38) higher than the pretest (M = 6.38), and TDFW t (15) = 6.46, p < .01, with the delayed posttest (M = 9.06) higher than the pretest (M = 5.06).

Paired samples comparisons across time for control condition. Results of paired samples t-tests across time for the control condition (n = 29) are presented in Table 26. The posttest to pretest comparisons were significant for the TFO, TDFO, and TFW measures, but not for the TDFW measure. For TFO, t (28) = 3.00, p <.05, with the posttest (M = 36.03) higher than the pretest (M = 28.90), for TDFO, t (28) = 2.52, p < .05, with the posttest (M = 16.17) higher than the pretest (M =14.72), and for TFW, t (28) = 2.69, p < .05, with the posttest (M = 11.62) higher than the pretest (M = 9.52). For TDFW t (28) = .85, p = .40, with the posttest (M = 7.48) higher than the pretest (M = 6.97). These results indicate that significant increases occurred for TFO, TDFO, and TFW from pretest to posttest for the control condition, but no significant increases occurred for the TDFW measure.

136

Table 26.

Paired Samples t-tests Across Time for the Control Condition (n=29).

Measure	Time Comparison	Mean Difference	t	df	Sig.
TFO	Posttest-Pretest	7.34	3.00	28	.01**
	Delayed Posttest-Posttest	-.24	-.15	28	.88
	Delayed Posttest-Pretest	6.90	2.76	28	.01**
TDFO	Posttest-Pretest	1.45	2.52	28	.02**
	Delayed Posttest-Posttest	1.03	1.75	28	.09
	Delayed Posttest-Pretest	2.48	4.17	28	.00*
TFW	Posttest-Pretest	2.10	2.69	28	.01**
	Delayed Posttest-Posttest	.52	.63	28	.53
	Delayed Posttest-Pretest	2.62	3.26	28	.00*
TDFW	Posttest-Pretest	.52	.85	28	.40
	Delayed Posttest-Posttest	.35	.71	28	.49
	Delayed Posttest-Pretest	.86	2.02	28	.05

* p < .01
**p < .05

Results of the delayed posttest to posttest comparisons were not significant for any of the dependent variables. For TFO, t (28) = -.15, p = .88, with the delayed posttest (M = 35.79) lower that the posttest (M = 36.03), for TDFO t (28) = 1.75, p = .09, with the delayed posttest (M = 17.21) more than the posttest (M = 16.17), for TFW, t (28) = .63, p = .53, with the delayed posttest (M = 12.14) higher than the posttest (M = 11.62), and for TDFW t (28) = .71, p = .49, with a delayed posttest (M = 7.83) more than the posttest (M = 7.48). These results indicate that for TDFO, TFW, and TDFW, there were increases from posttest to delayed posttest that were not significant, and there was a decrease from posttest to delayed posttest for TFW that was not significant.

The delayed posttest to pretest comparisons indicated significant differences for TFO, TDFO, and TFW, but not TDFW. For TFO, t (28) = 2.76, p < .05, with the delayed

137

posttest ($M = 35.79$) higher than the pretest ($M = 28.90$), for TDFO, t (28) = 4.17, p < .01, with the delayed posttest ($M = 17.21$) more that the pretest ($M = 14.72$), and for TFW, t (28) = 3.26, p < .01, with the delayed posttest ($M = 12.14$) higher than the pretest ($M = 9.52$). For TDFW, t (28) = 2.02, p = .05, with the delayed posttest ($M = 7.83$) higher than the pretest ($M = 6.97$). Increases from pretest to delayed posttest were significant for TFO, TDFO, and TFW, but not for TDFW.

Summary and Conclusions

The results of the data analysis indicated that there was a significant interaction of time and condition between-subjects, which showed that the testing time had a differential impact on treatment condition across mean scores on oral and written measures. There was also a significant interaction of testing time and treatment condition within-subjects, indicating that the condition impacted the oral and written scores across testing times. There were no significant differences across experimental conditions on between-subject analyses, indicating that there were no significant differences in mean scores in oral and written language when means were not separated by their individual testing times. There was a significant main effect across testing time, indicating that mean scores for oral and written describing abilities changed somewhere across the pretest, posttest, and delayed posttest.

Post hoc analyses indicated that students in all three treatment conditions increased their oral semantic describing measures from pretest to posttest, and did not incur significant changes in these scores from posttest to delayed posttest. This suggested that oral describing abilities improved for all conditions over the course of the study, and were maintained one month following the end of the treatment period. On the writing

138

measures, Condition T_2 was significantly lower than the control condition at the pretest. No other significant differences existed between any other conditions at the pretest. There were no significant differences across conditions on the writing measures at posttest and delayed posttest, indicating that the students in Condition T_2 improved more than those in the control condition. Comparisons of each condition across testing times for the writing measures showed that all conditions improved on the TFW measure from pretest to posttest, showing that all conditions improved the number of semantic features they used to describe words in writing. Conditions T_1 and T_2 significantly improved their TDFW scores from pretest to posttest, while the pretest to posttest comparison was not significant for the control condition. This suggested that students in Conditions T_1 and T_2 improved their ability to use different types of semantic information when writing, but the control condition did not. Condition T_1 and T_2 maintained TDFW scores on the delayed posttest, suggesting that they not only improved these scores, but maintained them after treatment concluded. Condition T_2 significantly increased TFW scores from posttest to delayed posttest, indicating that they continued to improve in this area following the end of treatment; while Condition T_1 and the control condition did not have significant changes in TFW scores. This suggested that Condition T_2 resulted in more increases in the number of semantic features used to describe words as compared to Condition T_1 and the control condition. These findings present useful information to educators with regards to effective instruction that can improve oral and written semantic describing abilities of children in elementary school.

139

CHAPTER V

DISCUSSION

Introduction

In the following sections, I have discussed the results of the study and my

conclusions based on these findings in relation to each of my three research questions:

1) What is the impact of whole-class metalinguistic strategy training on the oral

language skills of students in elementary school (grade 3)?

2) What is the impact of whole-class metalinguistic strategy training on the

written language skills of students' elementary school (grade 3)?

3) What is the impact of treatment dosage of metalinguistic strategy instruction on

the oral and written language skills of students in elementary school?

Additionally, I have disclosed the limitations to this study, and have discussed

implications for future practice. I have concluded this section with a discussion of future

research directions in relation to metalinguistic strategy instruction and semantic

describing abilities.

Research Question 1: What is the impact of whole-class metalinguistic strategy

instruction on the oral language skills of students in elementary school (grade 3)?

For the oral language measures (i.e, TFO, TDFO), the variables of testing time

did not significantly interact with treatment condition. There were no significant

140

differences across conditions, indicating that no significant differences existed between Condition T_1, T_2, or control at any of the testing times. This meant that prior to the treatment period; no significant differences existed across the three experimental conditions in the students' ability to describe common objects orally using semantic features. This was true with regards to the total number of semantic features the students stated, as measured by the TFO measure, as well as the different types of features they said, as measured by the TDFO measure. There were no significant differences in oral language following the treatment period, nor were there any significant differences between groups on oral language measures across delayed posttest scores taken one month following the cessation of treatment in Conditions T_1 and T_2.

Across all three conditions, TFO and TDFO means significantly improved from pretest to posttest, showing that all experimental conditions resulted in significant increases in oral language. No significant changes occurred in TFO or TDFO means in any condition from posttest to delayed posttest. This indicated that students in both Condition T_1 and T_2 maintained their gains after a brief cessation in treatment, and students in the control group did not make significant changes in their oral language skills during the time period between the posttest and delayed posttest.

The fact TFO and TDFO means improved for all conditions following the treatment period indicated that all three conditions were effective in improving the students' oral language skills. This is encouraging, because these results suggest that the activities done in the control condition were just as effective at improving students' expressive language skills as the metalinguistic strategy instruction conducted in Conditions T_1 and T_2. I can also conclude that no condition was significantly more

141

effective than another in this particular study in improving students' ability to describe words using semantic features.

While teachers withheld treatment for Conditions T_1 and T_2 following the posttest, the teachers in the control condition did not discontinue activities and students received the regular curricular instruction throughout the entire duration of the study. Because the students in the control condition continued to receive their assigned instruction between the posttest and the delayed posttest, it is difficult to tell whether they would have retained skills and strategies after a cessation of these activities. I can conclude that the treatment students received in Conditions T_1 and T_2 was effective and that students maintained their oral language skills after a month when they did not receive the treatment. I cannot determine if this holds true for the control condition. Withholding instruction, however, would not have been desirable or ethical.

There may be two possible reasons why students in Conditions T_1 and T_2 maintained their oral language skills after the EET was discontinued. First, all students in Conditions T_1 and T_2 returned to their regular curricular instruction which was similar to the instruction the control condition had received during the treatment conditions. This may have helped facilitate the maintenance of their oral language skills gained during the treatment period. Second, students in Conditions T_1 and T_2 may have continued to apply the EET strategy they learned during the course of the treatment period. In a strict experimental design, instruction would have stopped between posttest and delayed posttest, however I conducted this study in classrooms and could not ethically ask teachers to stop delivering instruction.

142

Although no statistically significant differences existed between group means for any testing time for the TFO measure, it is interesting to note the pattern of improvement across the conditions. After the pretest, differences between groups were less than 1.5 points (T_1 M = 28.38, T_2 M = 27.88, control M = 28.90; see Table 12). The control condition had the highest mean (28.90), Condition T_1 had the second highest mean (28.38), and Condition T_2 had the lowest mean (27.88). After the posttest, Condition T_2 had the highest mean (40.63), followed by Condition T_1 (39.00), and then followed by the control condition (36.03; see Table 13). Although these differences were not statistically significant it is important to note that the two treatment classrooms had the largest increases in scores; Condition T_1 increased by 10.68 (see Table 23) and Condition T_2 increased by 12.75 (see Table 24).

Mean differences from posttest to delayed posttest, while not statistically significant, did show increases in both Condition T_1 (1.00; see Table 23) and T_2 (3.06; see Table 24). The control condition had a slight decrease (.24; see Table 25). Consequently, even though the differences between groups across testing times for TFO were not statistically significant, it is interesting to note that both of the groups that received metalinguistic strategy instruction increased more than the control condition.

A similar pattern was present for TDFO scores. There was less than a one point difference in the means across all conditions (T_1=14.44, T_2=15.13, control=14.72; see Table 12) at the pretest. Both treatment conditions had an increase at the posttest; specifically Condition T_1 had an increase of 4.06 (see Table 23), Condition T_2 had an increase of 2.81 (see Table 24), and the control condition had an increase of 1.45 (see Table 25). Consequently from pretest to posttest, the TDFO for the two treatment

143

conditions improved more than the control condition. Increases in TDFO means from posttest to delayed posttest were not significant; as Conditions T_1 and T_2 both had increases of 1.06 (see Tables 23 and 24) and the control condition had an increase of 1.03 (see Table 25).

I recognize that a longer treatment period might yield greater differences between the conditions and that the length of the treatment period in this study may not have been sufficient to yield significant differences between pretest and posttest means. There is another possible explanation for the lack of significance between the increases from pretest to posttest in the treatment conditions. Some students may not need explicit metalinguistic training to improve their ability to describe words using semantic features. Previous research has indicated that students of all ability levels benefit from direct metalinguistic strategy instruction (Beck & Mckeown, 2007; Cain, 2007; Zipke, 2012). However, some of the students in this study may not have needed this type of training, as the instruction given in the control condition was effective in improving oral describing abilities. Therefore, there may have been no significant impact of treatment condition because some students may have had schematic representations of words vivid enough that explicit instruction was not needed.

Research Question 2: What is the impact of whole-class metalinguistic strategy instruction on the written language skills of students in elementary school (grade 3)?

For both the writing measures (i.e., TFW, TDFW), there was a significant interaction of time and condition, indicating that there was a differential impact of testing time on the participants' writing performance. Post hoc analyses revealed that there were significant differences between students in the Condition T_2 and the control condition in

144

both writing measures at the pretest, which suggested that the students' written language skills across experimental conditions were not equivalent prior to the treatment period with regards to the total number of semantic features the participants could generate in written form (TFW), as well as the different types of semantic information they could produce (TDFW). My intention in using the PPVT-4 and EVT-2 to determine vocabulary skills prior to intervention was to establish that students in all experimental conditions were equivalent prior to beginning intervention. Although there were not statistically significant differences between groups in receptive and expressive vocabulary as measured by the PPVT-4 and EVT-2 at the beginning of the study, it is possible that these measures were not sensitive enough to detect subtle differences in vocabulary that could impact written language. Neither test requires individuals to produce written descriptions of words, which could have explained why the conditions were equivalent on the PPVT-4 and EVT-2 scores but not the TFW and TDFW measures.

Further analysis revealed that the control group had the highest pretest mean scores (TFW $M = 9.52$, TDFW $M = 6.97$), followed by Condition T_1 (TFW $M = 9.00$, TDFW $M = 5.75$), and then Condition T_2 (TFW $M = 6.38$, TDFW $M = 5.06$; see Table 12). There was a .52 difference between means of the Condition T_1 and the control condition and a 3.14 difference between the Condition T_2 and the control condition means (see Table 21). I therefore conclude that although the groups were similar in oral language skills at the beginning of the study (as measured by the PPVT-4, EVT-2, TFO, and TDFO measures), written language scores were not equivalent across all conditions prior to treatment.

145

Within-subjects analyses indicated that there were significant increases in all groups from pretest to posttest for the TFW, indicating that all treatment conditions were effective in improving the number of semantic features students used to describe words. For the TDFW, posttest means of Conditions T_1 and T_2 significantly increased following the treatment period, but the control condition did not. This suggested the treatment conditions were more effective than the control condition in improving the different types of semantic information produced in written form (TDFW). This is contrary to the results regarding the TFO and the TDFO, which indicated no significant differences between pre and posttest means.

Therefore, regarding the TFW and the TDFW, the students in the control condition may have learned to describe words orally using diverse semantic information without explicit metalinguistic strategy instruction, but they did not use this skill in writing. This result is consistent with previous research that has indicated that students of varying populations and ability levels may require direct instruction in metacognitive strategies to improve writing; such as those used in the EET that enable students to use mnemonic devices as memory aides to organize and expand ideas in written form (Chalk et al., 2004; Graham & Perin, 2007; Little et al., 2010; Therrien et al., 2009).

Comparison across conditions indicated that there were no significant differences across conditions at the posttest, indicating that Condition T_2 was no longer significantly lower than the control group following treatment for both TFW and TDFW. At posttest, the ranking of means changed; Condition T_1 had the highest mean (TFW $M = 14.31$, TDFW $M = 9.25$), followed by Condition T_2 (TFW $M = 12.88$, TDFW $M = 8.38$), and then the control condition (TFW $M = 12.14$, TDFW $M = 7.83$; see Table 13). In other

146

words, students in the condition that was initially the lowest, Condition T_2, closed the gap in their skills compared to the control condition, which was initially the highest scoring condition.

This finding is important because it varies from previous research that has indicated that high performing students increase their vocabulary skills at a greater rate than low performing students (i.e., the Matthew Effect; Cain et al., 2004; Christ & Wang, 2011). The Matthew Effect has often been present in previous studies; as student with low vocabularies have made fewer gains even with effective direct vocabulary instruction (Christ & Wang, 2011; Ewers & Brownson, 1999; Marulis & Neuman, 2010). The Matthew Effect is often a concern because the students who are behind often stay behind. Therefore, is it is encouraging that Condition T_2 posttest means were similar and even slightly higher than the control condition. This result implies that teaching metalinguistic strategies to students can help bridge gaps in written language.

The results indicated that from posttest to delayed posttest, no significant changes in TFW or TDFW scores occurred for Condition T_1. I believe this indicates that the students maintained their writing skills gained during the treatment period. Similarly, Condition T_2 also maintained their TDFW scores, and even made a significant increase in the mean TFW score. This indicates that the students continued to improve in the number of semantic features they wrote for words. The TFW and TDFW scores of students in the control condition did not significantly change from posttest to delayed posttest. This result indicates that students in the control condition maintained skills they had at the posttest with regards to the number of semantic features they used to describe words (as measured by the TFW score). However, students in the control condition did

147

not change their ability to describe words with different semantic features (as measured by the TDFW score) throughout the course of the study across any of the testing times. I believe this is an important result since it clearly indicates that students who received the EET instruction were positively affected in their ability to write word descriptions when compared to students in the control condition.

Although there were no statistically significant differences across conditions at the posttest in the TFW and TDFW measures, the ranking of group means changed once again at the delayed posttest. For the delayed posttest, Condition T_2 had the highest mean (TFW $M = 15.38$, TDFW $M = 9.06$; see Table 14), followed by Condition T_1 (TFW $M = 13.38$, TDFW $M = 12.14$), with the control condition with the lowest mean (TFW $M = 12.14$, TDFW $M = 7.83$). Therefore, while there were no statistically significant differences in the conditions at the posttest or delayed posttest, the conditions that received metalinguistic strategy instruction improved at a greater rate for both TFW and TDFW compared to the control condition that did not receive the treatment. I believe that longer treatment period would have resulted in greater differences between conditions at posttest.

Because students in Condition T_2 began the treatment period with significantly lower scores than the control condition, they had to increase their writing scores at a greater rate in order to achieve equivalent scores following treatment. The fact that there were no significant differences between Condition T_2 and control at the posttest and delayed posttest following treatment suggested that students in Condition T_2 achieved more gains than the control condition. I can therefore conclude that Condition T_2, which received metalinguistic strategy instruction four times per week, was more effective in

148

improving writing performance compared to the control condition. If students in Condition T_2 had started the study with equivalent writing scores and increased at a greater rate than the control condition, it is likely that they would have had significantly higher scores following treatment.

Research Question 3: What is the impact of treatment dosage of metalinguistic strategy instruction on the oral and written language skills of students in elementary school?

Treatment Dose and Oral Language

While I aimed to investigate the impact of the presence of metalinguistic strategy instruction with the first two research questions; I aimed to determine whether the frequency of treatment impacted oral and written language measures with the third research question. Pertaining to the oral measures, there were no significant interactions of condition and time; and there were no significant differences across conditions at any of the testing times. Students in Condition T_1, who received treatment twice per week, and Condition T_2, who received treatment four times per week, both made significant differences in oral language measures (i.e., TFO, TDFO) following the treatment period. For TFO, students in Condition T_2 made greater increases from pretest to posttest (12.75 increase; see Table 24), when compared to students in Condition T_1 (10.63 increase; see Table 23). This indicated that the students who received treatment four times per week increased at a higher rate than the students that received treatment two times per week.

Interestingly, this pattern was reversed for TDFO scores, as students in Condition T_1 increased 4.06 (see Table 23) and Condition T_2 increased 2.81 (see Table 24). However, because differences in means for TFO and TDFO across conditions were not

149

significant from pretest and the posttest, I conclude that both doses were equally effective in improving the students' ability to describe words orally. When comparing posttest to delayed posttest; means for either condition did not change significantly. This indicated that not only were both doses equally effective in improving skills following treatment; but they were also equally effective in maintaining skills students gained in treatment.

Treatment Dose and Written Language

When comparing the written language measures (i.e., TFW, TDFW), there was a differential impact of testing time on condition. Further analysis revealed that there were significant differences between the mean scores of students in the control condition and in Condition T_2 at pretest in both of the writing measures. However there were not significant differences in mean pretest scores between Condition T_1 and Condition T_2. Condition T_1 had higher scores at pretest, while Condition T_2 had higher scores at posttest; indicating that mean scores for students in Condition T_2 increased more than the mean scores of students in Condition T_1 on both measures.

While no statistically significant increases were present between mean scores of students in Conditions T_1 and T_2 on delayed posttest, different patterns were present when I examined each condition separately across testing times. Students in Condition T_1 and Condition T_2 did not have significant differences in mean scores on the TFW or TDFW measures; however their scores did decrease slightly on both measures at delayed posttest. Therefore, receiving metalinguistic strategy instruction twice per week (Condition T_1) was sufficient in improving writing skills; however was not robust enough to result in continuous improvement after the conclusion of treatment. It is encouraging,

150

however, that the decreases in scores from posttest to delayed posttest were not significant.

For Condition T_2 the increase from posttest to delayed posttest for TDFW measure was not significant; however there was a significant increase in TFW scores. This finding is important because the total number of details written when describing semantic features not only increased following treatment, but actually continued to improve after treatment ended. This indicates the students retained skills learned during the treatment period as indicated on the delayed posttest. The retention of the EET strategy may have resulted in independent vocabulary learning, which also could have accounted for the increase in scores in the absence of treatment. This finding is consistent with previous research that has suggested direct instruction of metalinguistic strategies can result in independent learning, and subsequent increases in academic skills (Zipke, 2012; Zipoli et al., 2010).

Consequently, while both treatment doses (twice per week and four times per week) were equally effective in improving mean scores on both writing measures students who received treatment four times per week had higher mean scores at posttest. I believe this indicates that the intensity of instruction in Condition T_2 resulted in continued application of metalinguistic strategies beyond the period of direction instruction; while the intensity of instruction in Condition T_1 (twice per week) did not.

Limitations

This study was not without limitations. The first set of limitations relates to the student selection. First, the student sample was drawn from one geographic region. Therefore, it would be difficult to generalize the results of this study to students in other

151

geographic regions. Replication is necessary to determine the impact of metalinguistic strategy instruction on students in other schools and regions. Second, the students in this study who had an identified disability (i.e., autism, specific learning disabilities, speech and language impairments) were included for at least 80% of their school day. It is therefore difficult to determine the impact of the metalinguistic strategy instruction on students with more significant disabilities. Third, the teachers delivered the metalinguistic strategy instruction to students in large groups of at least 16 students. These results cannot be generalized to students receiving similar instruction in smaller groups or individually.

Another limitation was that students in the control condition received the same treatment throughout the course of the study from the pretest through the delayed posttest. Students in Conditions T_1 and T_2 only received their designated treatments from the pretest through the posttest, and then received typical instruction from posttest to delayed posttest. Ceasing treatment provided useful information when comparing students in Conditions T_1 and T_2, because I was able to compare the impact of treatment dose on the students' maintenance of skills once treatment concluded. However, it was difficult to make this same comparison between the control condition and Conditions T_1 and T_2 because the control condition had consistent instruction between the posttest and delayed posttest. This limitation was difficult to avoid, because teachers were ethically obligated to teach students in the control condition. Regardless of this factor, the experimental design of the study allowed me to make comparisons across conditions from pretest to posttest.

152

A final limitation was that the conditions were not equivalent in the writing measures at the beginning of the study. This could have been a potential covariate that could have interacted with the conditions. However, I did not have sufficient numbers to include starting ability level as a covariate in the statistical analyses. In spite of this limitation, including testing time as a factor in the statistical analyses allowed me to determine the interaction of both testing time and condition, which enabled me to determine the interactions of the assigned treatment across the testing times regardless of the students' ability levels at the beginning of the study. Additional analyses also allowed me to determine the changes in mean scores across testing time, which permitted me to determine students' progress over time in addition to comparisons across conditions.

Implications for Practice

An important factor in treatment effectiveness is the attitudes of the professionals implementing them, because this can impact the implementation fidelity and likeliness that other professionals would be willing to incorporate it into practice (Witt & Martens, 1983). Based on the data collected from the social validity measures for the teachers in Conditions T^1 and T^2, I can conclude that both teachers had high levels of acceptability of the EET. This is encouraging, in that the possibility exists that other teachers would feel that this intervention was both feasible and effective. It is necessary for educators to recognize this as an effective strategy so they will be willing to implement it. Feedback from the teachers on the open-ended survey also provided useful information regarding the considerations for implementing metalinguistic strategy instruction such as the EET strategy.

153

Although perceptions were overall positive, one teacher suggested some possible challenges to implementation. She was concerned about the time required to conduct metalinguistic strategy instruction in addition to the time required to teach the school curriculum. This teacher also expressed concern that some teachers may need assistance in choosing vocabulary used in the EET strategy. These concerns lead me to consider whether teachers will understand the importance of integrating metalinguistic strategy instruction rather than seeing it as something "extra".

Additionally, teachers will need training on how to integrate the EET strategy in addition to the guidelines and strategies they are already using; such as writing organizers and vocabulary instructional strategies. Therefore, it is necessary to create protocols for teachers that are both time-efficient and easy to combine with other curricular tools. This was one aspect of the EET that I found lacking. Clearer guidelines are needed for teachers with regards to choosing vocabulary from curricular materials. For this study, I assisted the teachers in choosing vocabulary during our weekly meetings; indicating that other teachers may need similar assistance in implementing the EET instruction should they choose to use it.

More specific guidelines are also needed for integrating EET instruction with graphic organizers aligned with the common core standards. The teachers in this study did not use any other graphic organizers during EET instruction besides the EET materials for writing. The teachers fit this instruction into their schedules because they completed EET instruction during the reading enrichment time, which is a time devoted to extra literacy development in addition to the standard curriculum. If teachers who want to use the EET do not choose to teach it during their allotted reading enrichment

154

time as the teachers did in this study, they will need to know how to integrate it in to other parts of the school day, such as during core reading and writing instruction.

Finally, the teacher in Condition T_1 thought that the intervention might be appropriate as a strategy for students with low to average vocabulary skills but not necessarily for those students with high vocabulary. The teacher in Condition T_2 stated that it was sometimes still difficult for students with lower academic skills to utilize the strategy. These concerns should be thoroughly investigated and this study could be conducted with students with specific disabilities to determine its effectiveness. More clarification is needed to determine how to best utilize this instruction with students of different ability levels.

Student perceptions are also a factor that can impact the effectiveness of treatment (Witt & Martens, 1983). Student ratings of the EET strategy were positive at the conclusion of the study. For Condition T_1, students rated six of ten items as highly acceptable and students in Condition T_2 rated nine out of ten items highly acceptable. These high acceptability ratings from the students were a positive finding, suggesting that other elementary-aged students would have similar ratings.

Teaching strategies must not only be acceptable to teachers and students; but they must also be grounded in research that proves their effectiveness. The findings of this study provided useful information to educators regarding evidence-based practices for students in the area of semantic skills. First, regular curricular activities many teachers utilize, in accordance with the common core standards, are as effective as direct metalinguistic strategy instruction focused on semantic features regarding impact on students' oral describing abilities. This implies that many educators who are

155

implementing instruction aligned with the common core standards are likely providing students with instruction that is effectively improving their oral vocabulary.

With regards to writing, these findings have suggested that teachers should consider implementing direct metalinguistic strategy instruction to improve students' written skills when describing words. Teaching metalinguistic strategies may help students to write more information during writing assignments. It can also help students describe words using more diverse information.

The results of this study also suggest that educators should consider utilizing metalinguistic strategy instruction focused on semantic features for students who are performing below their peers in the area of writing. Teaching them to use a mnemonic device such as the EET to expand and organize their ideas could result in improved describing abilities, as was the case in this study. This could possibly be a way to bridge the gap in performance, as the students in Condition T_2 did when compared to students in the control condition.

Furthermore, because scheduling and availability of time for strategy instruction may be a factor in educational decision making, teachers should be aware that different doses of instruction may yield different results. They should understand that less frequent instructional sessions (e.g., twice per week) may result in similar gains to more frequent instructional sessions (e.g., four times per week); but that greater frequency of instruction may result in better retention and independent application of skills.

156

Instrumentation and Assessment

There are several factors to consider regarding effective ways to measure student progress and areas of needed investigation. While there are existing standardized measures to assess students' oral language skills, such as the LPT-3 (Richard & Hanner, 2005), or the WORD test-elementary (2nd Ed.; Bowers, Huisingh, LoGuidice, & Orman, 2004), these would be too time-consuming for educators to utilize on a regular basis and were therefore not utilized in this study. They were also not designed to be administered repeatedly over time to measure response to instruction, but rather were designed to identify the presence or absence of language impairments for the purposes of therapeutic decision-making. These standardized measures are therefore not sensitive enough to measure progress over the time period I designated for this study, or that educators commonly use for progress monitoring. Additionally, while some writing CBMs currently exist; such as total words written or correct word sequences (Weissenburger & Espin, 2005), these methods would not measure students' semantic describing abilities in relation to the application of the EET or strategies similar to it.

Semantic skills can have a significant impact on academic performance and it is important for educators to identify students deficient in this area so that students needs can be addressed (Zipke, 2012). Research is needed to determine the most efficient and effective ways to measure students' semantic abilities that are feasible for educators. Without specific protocols for measuring students' semantic describing abilities, it would be difficult to accurately evaluate the effectiveness of interventions targeting these skills. Failing to have specific guidelines for coding of semantic responses could result in

157

differences across raters. While there are specific procedures in published materials for standardized assessments, there is not currently a consistent protocol utilized for measuring students' progress in metalinguistic strategy instruction that targets semantic describing, nor are there commonly used CBM protocols for oral or written semantic describing.

Specifically relating to the EET strategy, the protocol in the manual is not specific enough to ensure that those implementing it will do so consistently across raters. This is partially due to the fact that there are not detailed explanations of acceptable versus unacceptable responses for each semantic feature, and there is also some overlap features which could cause confusion in scoring. For example, there is overlap in the composition, location, and parts features with regard to the way the author has explained them in the manual. One cue for the composition feature students ask themselves "What is it made of?", to remind themselves to describe the composition of an item. However, some of the EET visuals and the EET manual also suggested using the question cue, "Where does it come from?" when students come to this part of the mnemonic device. So, for the word "apple", students could say that it comes from a tree, which could be considered a location, but would also be an appropriate answer to the question asking, "Where does it come from?"

Additionally, the manual states that explaining that a plant comes from a tree could fit in the composition category because it describes the plant's origin, yet it could also be considered a part of the plant. Having an overlap in semantic features is not problematic for students using the EET strategy, because a goal of the EET is to get students to give more information about items. Educators may not be concerned if

158

students provide the same type of information for multiple features of the EET as long as the students are successfully using the strategy to come up with organized and detailed responses. However, failing to distinguish responses between features may be problematic when used to score descriptions of words because it may result in scoring differences across raters.

The protocols I utilized for determining the students' TFO, TDFO, TFW and TDFW revealed acceptable levels of interrater agreement, indicating that the protocols were reliable measures of semantic describing ability across raters. However the process of achieving adequate agreement required considerable training and also necessitated the development of a rubric for scoring student responses. During the training period, differences existed between raters with regards to the extent to which students needed to explain items, as well as the way to categorize responses. The use of language, such as prepositions or verb phrases often changed the way a response could be coded. For example, for the word "apple", a specific verb phrase such as, "grows on a tree" would be coded as a function because the emphasis of that statement is of the apple growing. However, the response, "It's on a tree", would be coded as a location, because that statement does not contain a specific verb, and the emphasis of that response is that the apple can be found in a tree. We only achieved an acceptable level of agreement with the use of a rubric and practice scoring duplicate student responses. Due to these inconsistencies that occurred during interrater training, I conclude that even greater inconsistencies may occur between raters who do not use a rubric and specific scoring guidelines.

159

Although I established acceptable levels of interrater agreement for this study, further investigation is needed to determine the reliability of ratings across a variety of raters beyond those involved in this study. This could help to fine-tune the rubric that I have created, or develop other rubrics that contain acceptable and unacceptable responses, and could aid in clarifying ways to code responses to different semantic features.

Additionally, more research is needed regarding the type of words that educators should use to test semantic describing abilities. Because the words' difficulty level can potentially impact students' performance on semantic describing tasks, it is important to conduct research to specify this information on semantic measurement tools. While I referenced existing protocols to choose words for my data collection protocols, more investigation is needed to determine the best words to use across different ages and ability levels with regards to their appropriateness and potential to accurately measure improvements over time. Therefore, if educators choose to utilize assessments such as those I have utilized in this study, they should take steps to ensure adequate reliability and validity of their measurement tools.

Another factor to consider is the length of time allowed for completing responses during data collection. For the written responses collected to obtain the TFW and TDFW scores, the students had three minutes per word to complete responses. This guideline was based on the time students are typically allowed to write during established writing CBM protocols (Weissenburger & Espin, 2005). However, the possibility exists that more subtle differences in the quality of responses could have been detected had students been allowed a longer response time. For oral responses taken to obtain the TFO and

160

TDFO scores, the students did not have a time limit. Therefore, some students spent a significantly longer time responding than others, resulting in much higher scores.

Just as imposing a time limit may have impacted writing measures; not imposing a time limit for oral measures may have impacted the results. Having a time limit for semantic describing tasks may differentiate the students who have faster word retrieval from those who do not; which is often a skill that can impact academic performance. However, if time limits are too short, even students with higher vocabulary skills may not have enough time to formulate responses, which could result in floor effects. Therefore, future research in assessment of both oral and written describing abilities should investigate whether time limits are appropriate for measures of semantic progress. If time limits are a factor that can help to differentiate between students with low and high semantic abilities, it would be useful for educators to know the appropriate time limits for both oral and written tasks.

It is also important to determine if the measures I utilized in this study are correlated with existing measures of oral and written language. For writing, this may involve determining if semantic describing abilities are correlated with writing CBMs such as correct word sequences or total words written. For oral language, it would be useful to know if these protocols are correlated with standardized language assessments. Determining correlations with existing measures could help to determine concurrent validity, and thus establish effective protocols for measuring oral and written language skills. Establishing validity of additional measures of semantic describing could potentially provide educators with a wider range of time-efficient options for identifying students in need of further language interventions or more in-depth evaluations.

161

Frequency and Duration of Metalinguistic Strategy Instruction

While the treatment period in this study lasted nine weeks, it would be important to determine the impact of metalinguistic strategy instruction over a longer period of time. Both of the conditions that received treatment (T_1 and T_2) improved more on both the oral and written measures when compared to the control condition, yet these differences were not statistically significant. Future research is needed to determine if this pattern of improvement would persist over time, eventually resulting in significant differences for students who have received metalinguistic strategy instruction.

Additionally, future research could investigate the impact of the frequency of instruction over a longer time period. On the writing measures, students in Condition T_2 improved more than students in Condition T_1, yet differences across conditions were not significant. More research is needed to determine if significant differences in writing performance may occur in students who receive different frequencies of instruction. Finally, students in Condition T_2 had better retention and improvement of skills following a cessation of treatment as compared to students in Condition T_1. Further investigation is needed to determine if less frequent instructional sessions, such as those delivered in Condition T_1, will eventually lead to gains comparable to those seen in students who received more frequent sessions. Answering this question would be important to educators, because time constraints are a common issue in educational settings, and it would be worthwhile to determine if less frequent strategy instruction is as effective as more frequent strategy instruction.

162

Instructional Practices and Metalinguistic Strategy Instruction

I developed protocol for this study based on my interpretation of the procedures outlined in the EET manual (Smith, 2010), telephone conversations with the author, and two of the author's live lecture presentations (Smith, 2011). The EET manual provides general descriptions of how to use the strategy, in addition to a number of worksheets and student samples. However, it does not provide a step-by-step protocol for implementing strategy instruction, such as recommended components of lessons, specific methods of explaining strategy use and modeling, how to choose vocabulary words to target, and how to transition from oral to written language. Therefore, others may choose to implement EET strategy instruction using other materials, procedures, or vocabulary which may result in different outcomes than those in this study. Research on a number of factors is necessary to determine effective ways to teach students to use the EET.

One factor to consider is the emphasis and time spent on oral versus written language skills during metalinguistic strategy instruction. During the weekly meetings with the teachers assigned to Conditions T_1 and T_2, both teachers reported that some students were not correctly utilizing the EET strategy to describe words orally after several weeks of metalinguistic strategy instruction. Because of this, the teachers spent several weeks of intervention teaching students to orally describe words using the strategy before showing them how to use it during writing. The teachers reported they felt some students still did not have a full understanding of how to use the EET strategy to describe words orally before they began practicing it in writing; however the teachers felt it was appropriate to move on to writing given the necessary scaffolding because writing was also one of the treatment goals. The possibility exists that a greater emphasis

163

on oral describing skills could have resulted in greater gains by students in Conditions T_1 and T_2.

Although the teachers thought that the students could have used more practice using the EET in oral language than was allotted, the teachers still ended up spending more time on oral language that we had originally projected. Because we extended the time period for oral language instruction, we did not have as much time to work on using the EET in writing. Students did not have the opportunity to use the EET independently in paragraph writing until the last several weeks of the intervention. Just as with the oral language, teachers felt that students needed more time to practice using the EET in writing than they had during the nine weeks of intervention. It is therefore necessary to determine the appropriate length of time to spend on oral and written language. A greater emphasis on writing, and providing the students more time to use the EET strategy independently could have resulted in different outcomes, and is an important consideration for future studies.

Another factor that may have influenced the results is the choice of words used when teaching students to use the EET. When teaching students to use the EET in both oral and written language, teachers had students practice the strategy using common nouns that most elementary-aged students know. Because the focus was strategy instruction rather than content instruction, the teachers focused on words that would be easy to describe using the EET in order to facilitate independence with the strategy. The goal was for students to understand how to use the EET, and then allow them to practice the strategy with more difficult words in order to improve generalization of strategy use. However, due to time constraints of this study, students did not get the opportunity to use

164

the strategy with grade level content vocabulary. It is therefore possible that treatment would have had a greater impact on vocabulary development had students had the chance to apply it to more difficult words.

It is also necessary to investigate the impact of direct metalinguistic strategy instruction on small versus large groups. Research had indicated that intervention conducted in small groups can have different effects than large group instruction (Ehri, Nunes, Willows, Schuster, Yaghoub-Zadeh, & Shanahan, 2001). While this study focused on the impact of instruction delivered in a large group, there may have been a different impact had the groups been smaller.

Learner Characteristics and Response to Metalinguistic Strategy Instruction

In this study, the writing skills of the students in Condition T_2 began with significantly lower written semantic describing abilities compared to students in the control condition. The students in Condition T_2 made gains significant enough to catch up to the students in the control condition following treatment. This indicated the lower performing students were able to successfully improve their writing skills given metalinguistic strategy instruction. Metalinguistic strategy instruction has potential to bridge gaps in writing performance.

However, more information is needed regarding the impact of metalinguistic strategy instruction on students of varying ability levels. Because a Matthew Effect is often present in interventions that focus on semantic abilities (Christ & Wang, 2011; Marulis & Neuman, 2010), it is necessary to determine if students with lower oral and written describing abilities would achieve similar gains to those with higher abilities. It would be interesting to see if students with lower abilities could catch up to students with

165

higher abilities if both received metalinguistic strategy instruction. There is a possibility that students with higher abilities may not achieve gains as great as students with lower abilities because they may already possess these skills prior to intervention. Other literature has suggested students with higher abilities may improve at a greater rate than students with lower abilities because they have a more solid foundation of prior knowledge (Ewers & Brownson, 1999; Zipke, 2012).

More research is needed to determine the impact of metalinguistic strategy instruction on students with disabilities and socioeconomic levels. It is important to note that students who come from low socioeconomic backgrounds often come to school with less vocabulary knowledge that their same-aged peers (Hart & Risley, 1995). Socioeconomic status (SES) was not included in this data analysis. However, SES would be an important factor for future investigation.

Additionally, the presence of a disability can also impact a student's ability to learn new words (Marulis & Neuman, 2010). Most students in this study did not have an identified disability and those students who did have an identified disability and an IEP were in the general education setting for at least 80% of their school day. While the typical instructional activities conducted in the control condition for this study were effective in improving both oral language measures and one of the written language measures, I cannot generalize this finding to students with more severe disabilities. Oftentimes, general education instruction does not meet the specific needs of students with disabilities, and they may receive supplemental or direct instruction in metacognitive strategies to help them organize their oral and written language outputs (Cain et al., 2004). The presence or absence of a disability or low SES may often interact

166

with the impact of other factors; such as frequency or intensity of instruction, as students from these populations may respond differently to large group versus small group instruction (Ehri et al., 2001). Therefore, future research should focus on determining the appropriate frequency, intensity, and group size for students of different disability populations, abilities, and socioeconomic levels.

Summary

In summary, the metalinguistic strategy instruction focused on teaching students to use the EET was not significantly more effective than the regular curricular activities aligned with common core standards with regards to their impact on oral semantic describing abilities; regardless of the frequency of metalinguistic strategy instruction. Students who received metalinguistic strategy instruction four times per week improved written semantic describing abilities at a greater rate than students who did not receive metalinguistic strategy instruction. Receiving strategy instruction twice per week was just as effective as instruction delivered four times per week for both oral and written language; however students who received metalinguistic strategy instruction four times per week continued to improve the number of semantic features used in writing after treatment concluded compared to those who only received metalinguistic strategy instruction twice per week. More research is needed to determine the most efficient ways to deliver metalinguistic strategy instruction for both oral and written language, as well as ways to effectively assess these skills. Future investigation regarding factors such as treatment length, frequency, and student characteristics is necessary to inform effective educational practices.

REFERENCES

Alt, M., Plante, E., & Creusere, M. (2004). Semantic features in fast-mapping: Performance of preschoolers with specific language impairment versus preschoolers with normal language. *Journal of Speech, Language, and Hearing Research, 47,* 407-420. doi: 10.1044/1092-4388(2004/033)

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

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36,* 189-208. doi:10.1016/S0021-9924(03)00019-4

Beck, I.L., & McKeown, M.G. (2007). Increasing young low-income children's oral vocabulary repertoires through rich and focused instruction. *The Elementary School Journal, 107,* 251-271. doi:10.1086/511706

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

Biemiller , A. (2001). Teaching vocabulary: early, direct, and sequential. *American Educator, 25 (1).* Retrieved from: http://archive.aft.org/pubsreports/american_educator/spring2001/vocab.html

Biemiller, A., & Boote, C. (2006). An effective method for building meaning vocabulary in primary grades. *Journal of Educational Psychology, 98,* 44-62. doi:10.1037/0022-0663.98.1.44

Bordens, K. S., & Abbott, B. B. (2005). *Research design and methods: A process approach. (6th Ed.).* New York, NY: McGraw Hill.

Bowers, L., Huisingh, R., LoGuidice, C., & Orman, J. (2004). *The WORD test-elementary (2nd Ed).* East Moline, IL: Linguisystems, Inc.

Brown, D. (2006). What aspects of vocabulary knowledge do textbooks give attention to? *Language Teaching Research, 15,* 83-97. doi:10.1177/1362168810383345

Cain, K. (2007). Deriving word meanings from context: Does explanation facilitate contextual analysis? *Journal of Research in Reading, 30,* 347-359. doi:10.1111/j.1467-9817.2007.00336.x

168

Cain, K., Lemmon, K., & Oakhill, J. (2004). Individual differences in the inference of word meanings from contexts: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96,* 671-681. doi:10.1037/0022-0663.96.4.671

Chalk, J.C., Hagan-Burke, S., & Burke, M.D. (2005). The effects of self-regulated strategy development on the writing process of high school students with learning disabilities. *Learning Disability Quarterly, 28,* 75-87.

Christ, T., & Wang, C. (2011). Closing the vocabulary gap?: A review of research on early childhood vocabulary practices. *Reading Psychology, 32,* 426–458. doi:10.1080/02702711.2010.495638

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Coyne, M.D., McCoach, D.B., Kapp, S. (2007). Vocabulary intervention for Kindergarten students: Comparing extended instruction to embedded instruction and incidental exposure. *Learning Disabilities Quarterly, 30,* 74-88.

Cunningham, A.E., & Stanovich, K.E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*, 934-945. doi: 10.1037//0012-1649.33.6.934

Delogu, F., Raffone, A., & Belardinelli, M. O. (2009). Semantic coding in working memory: Is there a (multi)modality effect? *Memory, 17,* 655-663. doi: 10.1080/09658210902998054

Dreher, M.J., & Zenge, S.D. (1990). Using metalinguistic awareness in first grade to predict reading achievement in third and fifth grades. *Journal of Educational Research, 84,* 13-21.

Duff, F. J., Fieldsend, C.B., & Hulme, C. (2008). Reading with vocabulary intervention: Evaluation of an instruction for children with poor response to reading intervention. *Journal of Research in Reading, 31,* 319-336. doi:10.1111/j.1467-9817.2008.00376.x

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th Ed.). Minneapolis, MN: NCS Pearson.

Dynavox Mayer-Johnson. (1981-2011). *The picture communication symbols.* Pittsburgh, PA: Dynavox Mayer-Johnson, LLC.

169

Ehri, L.C., Nunes, S.R., Willows, D.M., Schuster, B.V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly, 36,* 250-287.

Elliot, J. L., & Gentile, J. R. (1986). The efficacy of a mnemonic technique for learning disabled and nondisabled adolescents. *Journal of Learning Disabilities, 19,* 237-241.

Ewers, C. A., & Brownson, S. M. (1999). Kindergarteners' vocabulary acquisition as a function of active vs. passive storybook reading, prior vocabulary, and working memory. *Journal of Reading Psychology, 20,* 11-20.

Frishkoff, G. A., Perfetti, C. A., & Collins-Thompson, K. (2010). Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology, 35,* 376-403. doi:10.1080/875656412010480915

Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research, 71*, 279-320. doi: 10.3102/00346543071002279

Graham, S. & Perin, D. (2007). *A meta-analysis of writing instruction for adolescent students. Journal of Educational Psychology, 99,* 445-476. doi: 10.1037/0022-0663.99.3.445

Graves, M.F. (2006). *The vocabulary book: Learning and instruction*. New York, NY: Teachers College Press.

Green, S. B., & Salkind, N. J. (2005). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (4th Ed.). Upper Saddle River, New Jersey: Pearson.

Gresham, F.M. (2004). Current status and future directions of school-based behavior interventions. *School Psychology Review, 33,* 326-343.

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

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences (5th Ed.).* Boston, MA: Houghton Mifflin Co.

IDEA. (2004). *Individuals with Disabilities Education Act Amendments of 2004*. Washington, DC.

170

Justice, L. M., Meier, J., & Walpole S. (2005). Learning new words from Storybooks: An efficacy study with at-risk Kindergartners. *Language, Speech, and Hearing Services in Schools, 36,* 17-32. doi:10.1044/0161-1461

Kibby, M. Y., Marks, W., Morgan, S., & Long, C. J. (2004). Specific impairment in developmental reading disabilities: A working memory approach. *Journal of Learning Disabilities, 37,* 349-363.

Li, Y., Wang, G., , Long, J., Yu, Z. Huang, B., Li, X., Yu, T., Sun, P. (2011). Reproducibility and Discriminability of Brain Patterns of Semantic Categories Enhanced by Congruent Audiovisual Stimuli. *PLosOne, 6,* 1-14 doi:10.1371/journal.pone.0020801

Little, M.A., Lane, K. L., Harris, K. R., Graham, S., Story, M., & Sandmel, K. (2010). Self-regulated strategies development for persuasive writing in tandem with schoolwide positive behavioral support: Effects for second-grade students with behavioral and writing difficulties. *Behavioral Disorders, 35,* 157-179.

Marulis, L. & Neuman, S. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of educational research, 80,* 300-335. doi:10.3102/0034654310377087

McGregor, K.K., Newman, R.M., Reilly, R.M., & Capone, N.C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research, 45,* 998-1014. doi:1092-4388/02/4505-0998

Munro, N., Lee, K., & Baker, E. (2008). Building vocabulary knowledge and phonological awareness skills in children with specific language impairment through hybrid language intervention: A feasibility study. *International Journal of Language, Communication Disorders, 43,* 662-682. doi: 10.1080/13682820701806308

Nagy,W.E., & Scott, J. (2000). Vocabulary processes. In M. Kamil, P. Mosenthal, P.D. Pearson, & R. Barr (Eds.)*, Handbook of reading research Volume III* (pp. 269-284) Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Newman, S.B., & Dwyer, J. (2009). Missing in action: Vocabulary instruction in Pre-K. *The Reading Teacher, 62,* 384-392. doi: 10.1598/RT.62.5.2

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11,* 375-383. doi:10.1080/10888430701530730

171

Richard, G. J., & Hanner, M. A. (2005). *Language processing test: Elementary (3rd Ed.).* East Moline, IL: Linguisystems, Inc.

Roehr, K. (2008). Linguistic and metalinguistic categories in second language learning. *Cognitive Linguistics, 19,* 67-106. doi: 10.1515/COG.2008.005

Scarborough, H.S. (2001). Connecting early language and literacy to letter reading (dis)abilities: Evidence, theory, and practice. In S. B. Neuman & D.K. Dickinson (Eds.)*, Handbook of early literacy research* (pp.97-110). New York, NY: Guilford Press.

Smith, S. L. (February, 2010). Expanding expression: A multisensory approach for improved oral expression and writing. *Illinois Speech-Language and Hearing*

Smith, S. L. (2011). *Expanding expression: A multisensory approach for improved oral and written language* (2nd Ed.). Bay City, MI: Expanding Expression.

Sousa, D. (2011). *How the brain learns (4th Ed.).* Thousand Oaks, CA: A SAGE company.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21,* 360-406.

Sugihara, T., Diltz, M.D., Averbeck, B. B., & Romanski, L. M. (2006). Integration of auditory and visual communication information in the primate ventrolateral prefrontal cortex. *The Journal of Neuroscience, 26* (43), 11138-11147. doi:10.1523/JNEUROSCI.3550-06.2006

Taylor, D.B., Mraz, M., Nichols, W.D., Rickelman, R.J., & Wood, K.D. (2007). Using explicit instruction to promote vocabulary learning for struggling readers. *Reading and Writing Quarterly, 25,* 205-220. doi: 10.1080/10573560802683663

Therrien, W. J., Hughes, C., Kapelski, C., & Mokhtari, K. (2009). Effectiveness of a test-taking strategy on achievement in essay tests for students with learning disabilities. *Journal of Learning Disabilities, 42,* 14-23. doi: 10.1177/0022219408326218

Vogt, W. P. (2007). *Quantitative research methods for professionals.* Boston, MA: Pearson Education, Inc.

Walsh, B.A., & Blewitt, P. (2006). The effect of questioning style during storybook reading on novel vocabulary acquisition of preschoolers. *Early Childhood Education Journal, 33,* 273-278. doi:10.1007/s10643-005-0052-0

172

Weissenburger, J. W., & Espin, C.A. (2005). Curriculum based measures of writing across grade levels. *Journal of School Psychology, 43,* 153-169. doi:10.1016/j.jsp.2005.03.002

Wiersma, W., & Jurs, S. G. (2005). *Research methods in education: An introduction.* (8th Ed.). Boston, MA: Allyn & Bacon.

Williams, K. T. (2007). *Expressive vocabulary test* (2nd Ed). Minneapolis, MN: NCS Pearson.

Witt, J. C., & Martens, B.K. (1983). Assessing the acceptability of behavioral interventions in used in classrooms. *Psychology in the Schools, 20,* 510-517. Vogt, P. (2007). *Quantitative research methods.* Boston, MA: Allyn & Bacon.

Zipke, M. (2012). First graders receive instruction in homonym detection and meaning articulation: The effect of explicit metalinguistic awareness practice on beginning readers. *Reading Psychology, 32,* 349-371 doi: 10.1080/02702711.2010.495312

Zipke, M. (2007). The role of metalinguistic awareness in the reading comprehension of sixth and seventh graders. *Reading Psychology, 28,* 375-396. doi: 10.1080/02702710701260615

Zipke, M., Ehri, L.C., & Cairns, H.S. (2009). Using Semantic Ambiguity Instruction to Improve Third Graders' Metalinguistic Awareness and Reading Comprehension: An Experimental Study. *Reading Research Quarterly, 44,* 300-321 doi: 10.1598/RRQ.44.3.4

Zipoli, R. P., Coyne, M. D., & McCoach, D. B. (2011). Enhancing vocabulary intervention for Kindergarten students: Strategic integration of semantically related and embedded word review. *Remedial and Special Education, 32,* 131-143. doi: 10.1177/0741932510361262

APPENDIX A

PRINCIPAL APPROVAL LETTER

Dear Kelly Throneburg,

My name is Karen Dudek and I am a doctoral candidate in the Department of Special Education at Illinois State University. I am completing research under the advisement of Dr. Julia Stoner and am requesting your assistance. My research will investigate the impact of an instructional method focused on teaching a metalinguistic strategy to children to improve their expressive oral and written language. This study will require the third grade teachers to conduct 20 minute group lessons in their classrooms with their students either twice or four times per week targeting metalinguistic strategy use for a 9 week period. It will also require teachers to attend several initial training sessions one hour in length to learn to conduct the instructional technique, as well as short weekly meetings lasting approximately 15 to 20 minutes in length. I will complete assessment procedures with students administered through both individual and group administration for purposes of measuring student progress, which will require students to miss some class time. Students will miss 30 minutes for procedures to establish group equivalency prior to study initiation. They will also miss 20-30 minutes during three different testing phases to participate in assessments. I will consult with teachers to determine the most appropriate times to complete assessments. Both teachers and students will be asked to take a short survey about their feelings about the intervention at two different times during the study, which will take approximately 10-15 minutes each. Audio recording of intervention and data collection sessions will be done and used for the purpose of ensuring implementation fidelity and will not be viewed by anyone not involved in this research. All audio recordings taken will be used solely for the purpose of the study and will not be viewed by anyone not involved in this research.

Students and teachers may incur some benefits to participating in this study due to the fact that they will learn a strategy that can improve student's language use in academic settings. Their participation in this study may also benefit other educators or children in that they will be assisting in research aimed at designing more effective teaching methods. There are minimal risks associated with participation in this study. To minimize loss of confidentiality, no other persons other than those associated with this study will listen to audiorecorded sessions. I will store assessment protocols, audio files, and other electronic data associated with this study in a locked office or password protected computer. Data will be purged five years following the study after results are shared.

I will answer any questions about this research now or at a later time during the course of this study and can provide a copy of its approval from the Institutional Review Board at Illinois State University upon request. I can be contacted at 309-527-4405 ext. 6772 or dudekk@unit11.org. You can also contact Dr. Julia Stoner at 309-438-5993 or jbstone@ilstu.edu with further questions.

Please sign below if you agree to allow this research at Jefferson Park School.

I, as the principal at Jefferson Park School, agree to allow Karen Dudek to conduct the research described above under the supervision of Dr. Julia Stoner at Jefferson Park School in El Paso, Illinois.

_____   4/18/13
Signature   Date

Sincerely,

Karen Dudek

Research at Illinois State University that involves human participants is overseen by the Institutional Review Board. Any questions about this research can be addressed by the Chairperson of Institutional Review Board, Research Ethics, and Compliance Office at Illinois State University, Illinois State University, Campus Box 3330, Normal, IL 61790-3330 or 309-438-2529.

174

APPENDIX B

IRB APPROVAL

May 15, 2013

Julia Stoner
5910 Special Education

Thank you for submitting the IRB protocol titled **"The Effect of Metalinguistic Strategy Instruction on the Oral and Written Expression of School-Aged Children"** for review by the Illinois State University Institutional Review Board (IRB). The IRB has **Approved** this research protocol following an **Expedited Review** procedure. You may begin this research.

This protocol has been given the IRB number **2013-0164**. This number should be used in all correspondence with the IRB. You may proceed with this study from **5/13/2013** to **5/7/2014**. You must submit a continuation request and receive approval prior to continuing your research beyond this expiration date.

Please also note that research protocols may be approved for continuation for a maximum of three years from the original date of approval in periods not to exceed one year. Research protocols having had three years of approval must be resubmitted and reviewed as new proposals.

This approval is valid only for the research activities, timeline, and subjects described in the above named protocol. IRB policy requires that any changes to this protocol be reported to, and approved by, the IRB before being implemented. You are also required to inform the IRB immediately of any problems encountered that could adversely affect the health or welfare of the subjects in this study. Please contact Kathy Spence, Assistant Director of Research at 438-2520 or myself in the event of an emergency. All other correspondence and questions should be addressed to:

Institutional Review Board
Campus Box 3330
Hovey Hall, Room 307
Telephone: 438-2529
E-mail: rec@IllinoisState.edu

175

It is your responsibility to notify all co-investigators **(Karen Dudek)**, including students, of the approval of this protocol as soon as possible.
Thank you for your assistance, and the best of success with your research.

Gary Creasey, Chairperson
Institutional Review Board
Telephone: 438-8139

APPENDIX C

TEACHER INFORMED CONSENT

Teacher Informed Consent

I _____, agree to participate in the study conducted by Karen Dudek, a doctoral candidate in the Department of Special Education at Illinois State University, under the advisement of Dr. Julia Stoner. I understand that participation is voluntary and that I can withdraw my permission at any time without penalty.

I understand that the purpose of this research is to investigate the impact of an instructional method focused on teaching a metalinguistic strategy to children to improve their expressive oral and written language. I understand that participation in this study may require me to conduct 20 minute group lessons in my classroom either conducted twice or four times per week targeting metalinguistic strategy use. It will also require me to attend an initial training session to learn to conduct the instructional technique, as well as short weekly meetings lasting approximately 15 to 20 minutes in length.

I understand that these instructional sessions may be observed and that audio recording will be taken of instructional sessions for the purpose of ensuring that the method is delivered correctly. I understand that any audio recordings taken will be used solely for the purpose of the study and will not be viewed by anyone not involved in this research.

I understand that I may incur some benefits to participating in this study due to the fact that I will learn a strategy that can improve my student's language use in academic settings. My participation in this study may also benefit other educators or children in that I will be assisting in research aimed at designing more effective teaching methods.

I understand that there are minimal risks associated with participation in this study. To minimize loss of confidentiality, no other persons other than those associated with this study will listen to audiorecorded sessions. I understand that Karen Dudek will store assessment protocols, audio files, and other electronic data associated with this study in her locked office or password protected computer. Data will be purged five years following the study after results are shared.

Karen Dudek will answer any questions about this research now or at a later time during the course of this study. She can be contacted at 309-527-4405 ext. 6772 or dudekk@unit11.org. I can also contact Dr. Julia Stoner at 309-438-5993 or jbstone@ilstu.edu.

I have been informed that if I have any questions about this research or my rights as a participant, I can contact the Chairperson of Institutional Review Board, Research Ethics, and Compliance Office at Illinois State University, Illinois State University, Campus Box 3330, Normal, IL 61790-3330 or 309-438-2529.
My signature below indicates my consent to participate in this study.

My signature below indicates my informed consent to participate in this study:

_____
Signature   Date

APPENDIX D

TEACHER DEMOGRAPHIC INFORMATION FORM

Teacher Demographic Information Form

1.   Age:_____

2.   Highest degree attained (circle one): Bachelor's   Master's   Ed.D

Ph.D.

3.   Number of years teaching:_____Number of years teaching at current

district_____

4.   What certificates do you currently hold?

_____

5.   Additional

Information:_____

_____

APPENDIX E

PARENTAL PERMISSION FORM

Parental Permission Form

I agree to allow my child, _____, to participate in the study conducted by
Karen Dudek, a doctoral candidate in the Department of Special Education at Illinois
State University, under the advisement of Dr. Julia Stoner. I understand that participation
is voluntary and that I can withdraw my permission at any time without penalty.

I understand that the purpose of this research is to investigate the impact of an
instructional method focused on teaching a metalinguistic strategy to children to improve
their expressive oral and written language. I understand that participation in this study
may involve my child participating in 20 minute group lessons either conducted twice or
four times per week targeting metalinguistic strategy use. It will also involve
participation in assessment procedures administered through both individual and group
administration for purposes of measuring student progress. Testing will include a 30
minute assessment at the beginning of the study, and 45 minute assessments conducted at
3 different times during the study. I understand that my child will miss some class time to
participate in these assessments.

I understand that my child may be observed and that audio recordings will be taken of
instructional sessions for the purpose of ensuring that the method is delivered correctly. I
understand that any audio recording taken will be used solely for the purpose of the study
and will not be viewed by anyone not involved in this research.

My child may incur some benefits to participating in this study due to the fact
metalinguistic strategy instruction can improve children's language use in academic
settings. My child's participation in this study may also benefit other children in that
he/she will be assisting in research aimed at designing more effective teaching methods.

I understand that there are minimal risks associated with participation in this study. To
minimize loss of confidentiality, no other persons other than those associated with this

study will listen to audio recorded sessions. Additionally, any assessment protocols will be assigned a code to protect my child's identity. No other individuals other than those involved in the research will have access to these codes and protocols. I understand that Karen Dudek will store assessment protocols, audio files, and other electronic data associated with this study in her locked office or password protected computer. Data will be purged five years following the study after results are shared.

Karen Dudek will answer any questions about this research now or at a later time during the course of this study. She can be contacted at 309-527-4405 ext. 6772 or dudekk@unit11.org. I can also contact Dr. Julia Stoner at 309-438-5993 or jbstone@ilstu.edu.

I have been informed that if I have any questions about this research or my child's rights as a participant, I can contact the Chairperson of Institutional Review Board, Research Ethics, and Compliance Office at Illinois State University, Illinois State University, Campus Box 3330, Normal, IL 61790-3330 or 309-438-2529.

My signature below indicates my permission to allow my child to participate in this study.

_____

Signature   Date

APPENDIX F

STUDENT ASSENT SCRIPT

Hi Class,

My name is Ms. Dudek and I know some of you have seen me working your school. I am going to school, just like you are, at Illinois State University. I am doing research and would like you to be a part of it. I am asking every third grade student at Jefferson Park to be in my study.

I am asking you to look at some pictures with graduate student friends from ISU. Then I will look at that information and choose two classes to learn a way to remember vocabulary words. I or my graduate student friends from ISU will ask you some more questions 3 times during the next couple of months. The way to learn the vocabulary is called EET!

If you are in one of those classes your teacher will show you the new way to learn vocabulary.
If you don't want to be in my study that is fine but you will still learn the EET if you are in one of the chosen classrooms. If you don't want to go with one of my graduate student friends and look at pictures that is fine too-you don't have to.

Do you have any questions for me about my study?

Would you like to be in my study? (The classroom teacher will verify assent or non assent from the students).

Thank you so much.

APPENDIX G

IMPLEMENTATION FIDELITY CHECKLIST

Implementation Fidelity Checklist

Teacher:_____

Date:_____

For the following components please indicate if the components was: present=1, or not present=0

1. Stated name of strategy	
2. Explain the rationale	
3. Exposed to semantic features	
4. Provided models of strategy use	
5. Provided opportunities for strategy use	
6. Provided access to mnemonic device cues	

APPENDIX H

ORAL DESCRIPTION PROMPT

ORAL DESCRIPTION PROMPT

Administration Script: For each word say, "I want you to pretend like I don't know anything about these words. I want you to tell me everything you know about them." For each individual item say, "Tell me everything you know about a _____." (Allowable prompt: "Can you tell me more?").

1. Apple	
2. Television	
3. School	
4. Bus	
5. Bed	
6. Bumblebee	

184

APPENDIX I

SEMANTIC FEATURES RUBRIC

Semantic Features Rubric
Oral Prompt Words

	Acceptable	Unacceptable
Apple		
Category	Fruit, food, produce, specific brand or kind of apple, snack	Vegetable, any other incorrect food category
Function	Eat it, bake it, make apple pie, make apple sauce/apple juice, peel it, pick it, slice it, grow it/it grows on a tree, squeeze/juice it, get bruised, ripen	Any general verb that is not specifically associated with the intended purpose of apples (e.g., throw it, play with it)
Physical Attributes	Round, edible, specific colors of an apple (i.e., red, yellow, green),(turn) brown, hard, sticky, juicy, crunchy, sweet, sour, healthy, specific appropriate size comparison to another object (e.g., bigger than a ___, smaller than a___)	Mention of color without listing specific color name (e.g., any color, many colors), general statement of size without specific dimension or comparison (e.g., big, small)
Composition	Juice, vitamins (or specific types), nutrients (or specific types), sugar, carbohydrates, any other appropriate components	
Parts	Seed, stem, skin, leaf, core	
Associations	The following items listed without a specific verb or prepositional phrase: apple pie, apple juice/sauce, worm, types of utensils, serving dishes/containers	

185

Location	(Find/get/come) from or on a tree, in a (specific type of serving or storage container), (buy/get)at the store, specific type or department of a store, in the kitchen, at an orchard *Statements such as the following stated as prepositional phrases should count as locations rather than associations: "in applesauce", "in juice", "in an apple pie." *The statement "Grows on a tree/from a seed" should be given one point for a function.	
School		
Category	Building, facility, specific name or type of school	
Function	Any of the following listed in a verb phrase" Kids learn, teachers teach, read (books), learn/do homework/work/study; specific subjects or activities listed as verb phrases (e.g., do math, learn to read, do crafts, eat lunch, play at recess)	
Physical Attributes	Specific size dimensions; specific appropriate size comparison to another object, specific length of time (e.g., 6-8 hours, stay until 3:30).	Mention of color without listing specific color name (e.g., any color, many colors), general statement of size without specific dimension or comparison (e.g., big, small)
Composition	Wood, bricks, any other appropriate material used for construction of schools	
Parts	Classrooms, specific classrooms or places, (e.g, library, music room, hallway, lockers, bathrooms, cubbies, offices, gym, playground)	

186

Associations	Any of the following listed in absence of a verb phrase: Specific school staff (e.g., teachers, principals), students, specific school materials/equipment (e.g, desks, pencils, flags, white boards, computers), subjects (or listed specific subjects). *if student says, you can "See/find (associated item)" give point for associations. *If student lists a general category (e.g., supplies, subjects), and then proceeds to list specific items in this category, award a point for the initial mention of the category and a point each for every item mentioned within that category.	*If student lists the word subjects and then lists examples, do not give a point for the initial statement of "subjects"
Location	Specific place of a school (e.g., street name, town), in the country, in the city, in towns	
Television		
Category	Machine, appliance, electronic device/equipment, specific brand or kind of television (e.g., plasma,	Toy, general nouns (e.g., thing), TV
Function	Watch shows/TV/movies, play (video) games, learn information, entertain, do exercise videos/programs, plug it in, change channels/control (with remote), record show/movies/programs, repair it	Any vague explanation such as "Shows you stuff"

187

Physical Attributes	Black or other appropriate color, square, rectangular, flat (screen), specific size dimensions or comparisons, electronic, colored or black and white (in reference to type of screen display)	General statements of size (e.g., big, small, large, tall), the phrase "It can be many colors" stated not referencing specific type of screen display (e.g., color screen versus black and white screen)
Composition	Picture, sound, lectricity, glass, plastic, metal, plasma, any other appropriate material used for constructing a television	
Parts	Antenna, screen, cords, wires, buttons, jacks, tuner, display, any other appropriate television components	Any misnamed or nonspecific television component
Associations	Any associated item listed without a verb phrase: DVDs, DVD player, channels, videos, shows, movies, commercials, game system, video games, remote, cable, dish, piece of furniture any other appropriate associated item *The statement "Change channel/control with remote" should get a point for function (change channel) and a point for association (remote).	Any misnamed or nonspecific television component
Location	(Get/buy) from the store (or specific type or department of a store, any appropriate room for a television (e.g., living room, bedroom), on a TV stand, in the entertainment center, any other appropriate place of purchase	

188

Bus		
Category	Transportation, vehicle, specific type or brand name of bus	Car, truck, other types of vehicles
Function	Drive it, ride in it, takes kids to school/home, takes people places, carries people	Other actions not specific to riding a bus (e.g., sitting, read a book, put stuff under the seat)
Physical Attributes	Yellow or other appropriate color, specific size dimensions or comparisons (e.g., bigger than a car), specific common colors of buses not used for school (e.g., "Some buses can be white").	General statements of size (e.g., long, big, small, short), it can be any color
Composition	Metal, rubber, plastic, glass, any other appropriate material	
Parts	Door, seats, steering wheel, windows, rearview mirrors, emergency exits, engine, any other appropriate parts	
Associations	Any associated parts not listed as part of a verb or prepositional phrase (e.g., kids, students, passengers, bus driver), school supplies, listing specific supplies, book bag *The statement "Kids ride to school" should get a point for associations (kids) and function (ride to school).	Nonspecific items not specifically associated with a bus (e.g., people, stuff)
Location	Bus barn, bus stop, schools, on the street	
Bed		
Category	Furniture, specific type of bed (e.g, bunk bend, water bed, Queen, futon)	

189

Function	Sleep, nap, rest, lay in it, watch television, read a book, make it, wake up in it	Any other verb not specific to the function of a bed (e.g., play on, jump on, do homework, sit on)
Physical Attributes	Soft, hard, comfortable, specific size dimension or comparison, square, rectangular, appropriate colors listed in association with specific parts of beds (e.g., bed frame can be brown/black/white, mattress is white, comforters can have many different colors/prints)	General statement of size without specific dimension or comparison, the phrase "It can be any color" without reference to a specific part of a bed
Composition	Wood, metal, fabric, cotton, water, any other specific material used to make a part of a bed	
Parts	Mattress, bed frame, blankets, comforter, pillows, pillow cases, wheels, springs, head board, sheets, any other appropriate part of a bed	
Associations	Additional items often placed on or near bed (e.g. stuffed animals, dressers, nightstands, alarm clocks, night lights)	
Location	Bedroom, stores or appropriate department of stores, in your house, in a hospital, any other common location for a bed	
Bumblebee		
Category	Bugs, insects, specific kinds of bees (e.g., larva, worker bees, drones, Queen bees, honey bees, sweat bees)	Any misnamed kind of bees or different kind of insect (e.g, wasps, hornets, yellow jackets, nonspecific type of bee such as "baby" or "little" bees)

Function	Sting you, die after they sting you, collect/drink/eat pollen from flowers, you swat them, have allergic reaction to sting, collect/eat/drink/suck nectar, pollinate, make/build hives, (worker bees) make honey, queen lays eggs (count as one function if student mentions specific type of bee and a specific role in one verb phrase), buzz, fly	Nonspecific explanations such as moves, makes noise, bothers you, scares you, have babies
Physical Attributes	Yellow, black, fuzzy, specific size dimension or comparison, have stripes	General statement of color/shape/size without specific dimension or comparison
Composition	Cells, blood, fur, skin, bristles, venom	
Parts	Antennae, wings, stinger, thorax, stomach/abdomen, head, eyes, tongue/proboscis, legs	Any other incorrect body parts
Associations	Any associated items not listed in a specific verb or prepositional phrase (e.g., flowers, pollen, nectar)	
Location	On flowers (when stated in a prepositional phrase), in a hive, outside, in nature, on a bee farm, *the statement "Live in a hive" should receive one point for location only	

191

Written Prompt Words

	Acceptable	Unacceptable
Banana		
Category	Fruit, produce, food, specific brands or types of bananas, snack	Vegetable or any other incorrect food category
Function	Eat, peel, slice, edible, the following types of responses when listed as verb phrases: make banana bread , grow on a tree, blend in smoothies, get bruised, ripen	
Physical Attributes	Yellow, round, brown/black, green, long, skinny, soft, mushy, healthy, any specific size dimension or comparison, bruised, nutritious, squishy, edible	Vague statement of size such as big or small
Composition	Vitamins (or specific types), nutrients (or specific types), sugar, carbohydrates, fiber, any other appropriate components	
Parts	Skin, stem, seeds	
Associations	The following types of responses when listed without specific verb or prepositional phrases: banana bread, smoothies, blenders, types of utensils, serving dishes	
Location	On a tree, in a store, in a specific type of department of store, on/in a (specific serving or cooking container), in a field, on a farm	
Flower		
Category	Plant, decoration, specific type of flower	

192

Function	Grows, make bouquet/corsage/boutonniere (when listed as a verb phrase), plant, smell it, garnish with it , give it to someone for (specific purpose or event such as wedding, birthday, anniversary), gives oxygen/air, blooms, animals eat	
Physical Attributes	Specific dimensions or size comparisons, can be many colors, specific appropriate colors, fragrant/smells good/ beautiful/pretty (count as one point if list both words)	General statements of size without specific dimensions (e.g., big, small, short, tall)
Composition	Chlorophyll, cells, pigment, fragrance, nectar, pollen, any other appropriate components	
Parts	Stem, leaf, root, petals, pollen, any other appropriate parts, scent	
Associations	Any appropriate associated object listed without a verb or prepositional phrase: bees, seeds, bouquet/corsage/boutonniere, weeds , sun, water, soil	
Location	In a flower pot, in the dirt/soil/ground, at a store/nursery, appropriate department of a store, in a (bouquet, corsage/boutonniere/vase), in a garden/yard, at a (any appropriate location for an event using flowers as a decoration such as a church or banquet hall)	

Train		
Category	Transportation, vehicle, machine, specific type of train (e.g., commuter, cargo)	
Function	Transports people, commute, travel in it, drive it, ride in it, takes you places (or listing of specific places, such as vacation, work), goes fast, blows horn/makes loud noise	Any other activity not specifically associated with the purpose of a train (e.g., sit in it, take a nap, read a book)
Physical Attributes	Specific appropriate colors or dimensions, appropriate size comparisons (e.g., longer/bigger than a car/bus/van), loud, powerful	General statements of size (e.g., long, big)
Composition	Metal, plastic, glass, any other appropriate component	
Parts	Engine, wheels, motor, seats, doors, compartments, horn, windows, any other appropriate parts of a train	
Associations	Appropriate items associated with trains not listed in a prepositional or verb phrase (e.g., passengers, tracks, coal)	
Location	Specific prepositional phrases such as: at the train station, on the track, specific location of a train station or destination, railroad,	

APPENDIX J

WRITTEN PROMPT SCRIPT

Written Prompt Script

Administration Script: For each word say, "I want you to pretend like I don't know anything about these words. I want you to write everything you know about them using complete sentences. You will have three minutes to write about each word." For each individual item say, "Tell me everything you know about a _____. You have three minutes. You may begin."

APPENDIX K

ADAPTED TEACHER IRP FOR METALINGUISTIC AWARENESS INSTRUCTION

Teacher IRP (adapted) for Metalinguistic Awareness Instruction (Witt & Martens, 1983).

		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	This would be an acceptable type of instruction for improving expressive language skills.					
2	Most teachers would find this instructional method appropriate for addressing potential expressive language delays.					
3	This method should prove effective for increasing expressive language skills.					
4	I would suggest the use of this method to other teachers.					
5	I would be willing to use this method in the classroom setting.					

196

6	This method would not result in negative effects for students.					
7	This method would be appropriate for a variety of students.					
8	This method is consistent with those I have used in the classroom setting.					
9	This instructional method is a reasonable way to improve students' expressive language skills.					
10	I like the procedures in this instructional method.					
11	This method is a good way to address potential delays in expressive language skills.					
12	Overall, this method would be beneficial for students.					

APPENDIX L

ADAPTED STUDENT IRP FOR METALINGUISTIC AWARENESS INSTRUCTION

Student IRP (adapted) for Metalinguistic Awareness Instruction (Witt & Martens, 1983).
The Picture Communication Symbols (c)1981-2013 by DynaVox Mayer-Johnson LLC.
All Rights Reserved Worldwide. Used with permission.

		Disagree	Not sure	Agree
1	Using the EET is a good way for me to learn and explain new words.			
2	Most kids would like to use the EET to help them learn and explain new words.			
3	Using the EET can help me to write and speak more easily.			
4	I would tell my friends to use the EET to help them learn new words.			

5	I would tell me friends to use the EET to help them think of things to write.			
6	I would like to use the EET to learn new words.			
7	I would like to use the EET to explain what words mean.			
8	I would like to use the EET to think of things to write.			
9	Using the EET would be a good way to help me learn and explain difficult words.			
10	Using the EET will help me perform better in school.			

APPENDIX M

IRP COPYRIGHT PERMISSION

This is a License Agreement between Karen L Dudek ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3272600446284
License date	Nov 19, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Psychology in the Schools
Licensed content title	Assessing the acceptability of behavioral interventions used in classrooms
Licensed copyright line	Copyright © 1983 Wiley Periodicals, Inc., A Wiley Company
Licensed content author	Joseph C. Witt,B. K. Martens
Licensed content date	Feb 14, 2006
Start page	510
End page	517
Type of use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Table 2
Will you be translating?	No
Total	0.00 USD
Terms and Conditions	

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this license must be completed within two years of the date of the grant of this license (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

201

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

Wiley publishes Open Access articles in both its Wiley Open Access Journals program [http://www.wileyopenaccess.com/view/index.html] and as Online Open articles in its subscription journals. The majority of Wiley Open Access Journals have adopted the Creative Commons Attribution License (CC BY) which permits the unrestricted use, distribution, reproduction, adaptation and commercial exploitation of the article in any medium. No permission is required to use the article in this way provided that the article is properly cited and other license terms are observed. A small number of Wiley Open Access journals have retained the Creative Commons Attribution Non Commercial License (CC BY-NC), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Online Open articles - Authors selecting Online Open are, unless particular exceptions apply, offered a choice of Creative Commons licenses. They may therefore select from the CC BY, the CC BY-NC and the Attribution-NoDerivatives (CC BY-NC-ND). The CC BY-NC-ND is more restrictive than the CC BY-NC as it does not permit adaptations or modifications without rights holder consent.

Use by commercial "for-profit" organizations

Use of non-commercial Wiley Open Access articles for commercial, promotional, or marketing purposes requires further explicit permission from Wiley and will be subject to a fee. Commercial purposes include:

- Copying or downloading of articles, or linking to such articles for further redistribution, sale or licensing;

- Copying, downloading or posting by a site or service that incorporates advertising with such content;

- The inclusion or incorporation of article content in other works or services (other than normal quotations with an appropriate citation) that is then available for sale or licensing, for a fee (for example, a compilation produced for marketing purposes, inclusion in a sales pack)

- Use of article content (other than normal quotations with appropriate citation) by for-profit organizations for promotional purposes

- Linking to article content in e-mails redistributed for promotional, marketing or educational purposes;

- Use for the purposes of monetary reward by means of sale, resale, license, loan, transfer or other form of commercial exploitation such as marketing products

- Print reprints of Wiley Open Access articles can be purchased from: corporatesales@wiley.com

The modification or adaptation for any purpose of an article referencing the CC BY-NC-ND License requires consent which can be requested from RightsLink@wiley.com .

Other Terms and Conditions:

BY CLICKING ON THE "I AGREE..." BOX, YOU ACKNOWLEDGE THAT YOU HAVE READ AND FULLY UNDERSTAND EACH OF THE SECTIONS OF AND PROVISIONS SET FORTH IN THIS AGREEMENT AND THAT YOU ARE IN AGREEMENT WITH AND ARE WILLING TO ACCEPT ALL OF YOUR OBLIGATIONS AS SET FORTH IN THIS AGREEMENT.

v1.8

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK501163006.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

APPENDIX N

PERMISSION TO REPRINT BOARDMAKER® SYMBOLS

Dear Karen,

Thank you for your email. Please use up to 50 symbols with our permission.

The Picture Communication Symbols (c)1981-2013 by DynaVox Mayer-Johnson LLC.
All Rights Reserved Worldwide. Used with permission.

Please let me know if you need anything else.

Thank you,
Alicia Trax
Reimbursement Manager
_____

From: kldudek@ilstu.edu [kldudek@ilstu.edu]
Sent: Wednesday, October 16, 2013 4:50 PM
To: Mjq
Subject: Permission to print symbols

Hello,

 My name is Karen Dudek, and I am a doctoral candidate at Illinois State University. I am
contacting you to request permission to print a survey that I have created for my
dissertation study that includes 3 Boardmaker symbols. The symbols I have used include
a happy face, a confused face, and a sad face as part of a Likert scale that I will have used
with elementary aged children in order to measure the social validity of a language
intervention. If you could indicate what I will need to do in order to get permission to use
these symbols, that would be great. I am requesting to reprint and use these symbols as
part of the survey that I have used for my dissertation research only.

Thank you,

Karen Dudek

207

APPENDIX O

TEACHER PERCEPTIONS OF INTERVENTION QUESTIONNAIRE

Teachers Perceptions of Intervention Questionnaire

1. How do you feel this intervention will address students' oral and written language expression skills?

2. What feedback do you have regarding the appropriateness of this intervention in addressing language expression in the classroom?

3. What questions/concerns do you have regarding the implementation of the intervention?

4. Do you think that the goals of this intervention will be met?

5. Do you think the students will benefit from this intervention?

6. How easy were the procedures for implementation?