AN INTERACTIVE APPROACH TO VOCABULARY INTERVENTION
FOR SPEECH-LANGUAGE PATHOLOGISTS IN THE CLASSROOM:
MULTIPLE LANGUAGE LEVELS AND MODALITIES

BY

MARY C. KUBALANZA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Speech and Hearing Science
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

       Associate Professor Cynthia J. Johnson, Chair
       Associate Professor Hedda Meadan-Kaplansky
       Associate Professor Laura Segebart DeThorne
       Assistant Professor Raksha Anand Mudar

ABSTRACT

Language and literacy research suggests a strong correlation between reading comprehension and vocabulary knowledge. Intervention research suggests that students who are struggling to read at grade level benefit from additional direct instruction. The present study proposed a vocabulary intervention for struggling readers, led by a speech-language pathologist (SLP) in the classroom setting. Four second and third grade students participated in small-group instruction during the scheduled classroom literacy block.

Two instructional models were compared using a single-case alternating treatments design. During treatment A, small-group instruction integrated interactive, multi-linguistic, and multimodal strategies to support word learning. Research suggests that these instructional techniques may allow students with language and literacy differences greater access to new word knowledge. This treatment is referred to as DRAW, due to one of its multimodal components, in which the students drew pictures. During treatment B, small-group instruction followed a traditional instructional model focused on dictionary definitions and independent composition of sentences using the target words. This treatment is referred to as DICT, due to its use of an online dictionary.

Vocabulary gains were measured by the accuracy of written definitions, word recognition during oral reading, and the use of target words during spoken story retell (i.e., vocabulary retell). A secondary analysis of the number of propositions recalled during the spoken story retells was used to explore the effects of vocabulary intervention on overall reading comprehension (i.e., proposition retell). Following treatment, the social validity of the SLP offering classroom-based vocabulary intervention was determined through a teacher questionnaire, as well as teacher and student interviews.

A treatment effect was demonstrated for the written definition and word recognition tasks, in which all four and three of the students, respectively, acquired a larger number of words following the DRAW treatment. No effect was demonstrated for vocabulary retell or proposition retell (i.e., reading comprehension). Although only a fairly small number of target words were acquired for written definitions following the DICT treatment, all four of the students acquired one to two thirds of the target words on this measure following the DRAW treatment. Similarly, the students gained approximately one third to one half of the target words on the word recognition measure as well. Two of the students showed a bias towards acquiring nouns, as opposed to verbs and adjectives. With regard to social validity, the teacher indicated that the classroom intervention was effective and feasible. The children reported that they enjoyed the treatment, particularly DRAW.

The present study supports the value of SLPs providing intervention or instruction in the classroom (called a "push-in" service delivery model), a practice long advocated by the national professional organization for SLPs, though not yet widely adopted. The study also supports the relatively new concept of in-class, small-group intervention, where students socialize around vocabulary learning tasks, actively focus on multiple linguistic aspects of a word (i.e., how it sounds, how it looks in print, and what it means), and interact with the word through several modalities (listening, speaking, reading, writing, and drawing). With corroboration from the social validity measures, results suggest that an interactive, multi-linguistic, and multimodal intervention is an effective, efficient, feasible, and even enjoyable (for the teacher and students alike) approach to promoting the acquisition of advanced sight vocabulary.

ACKNOWLEDGMENTS

I would first like to thank my brilliant advisor, Dr. Cynthia Johnson, for her guidance and support over the course of my graduate studies. I feel so fortunate to have found a mentor who was truly invested in helping me reach my goals during the doctoral program and find my voice as a researcher, teacher, and clinician. I am also grateful to the members of my committee: Drs. Hedda Meadan-Kaplansky, Laura Segebart DeThorne, and Raksha Anand Mudar, for their time and expertise throughout this project. Thanks to Dr. James Halle for his thoughtful questions and encouraging advice, and thanks to Dr. Holly Shapiro for sharing so many insights from her clinical experience.

This project could not have been completed without the hard work and dedication of many other individuals. The local teachers and students who welcomed me into their classrooms to pilot materials and procedures were invaluable during the preparatory stages of this study. The administrators, teachers, and families at my collaborating school were also a pleasure to work with and could not have been more accommodating. Of course, the four student participants were wonderful, and I will never forget their effort and enthusiasm. I feel lucky to have had the opportunity to learn with them and become a part of their classroom community. Thanks also to the graduate and undergraduate students who assisted with data collection and analysis. I am so appreciative of their careful attention to detail and dependability.

Finally, I would like to thank my family and friends for their love and support. A special thanks to my mom, whose confidence in me helped me have confidence in myself. Thank you.

iv

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

The National Reading Panel identified vocabulary as one of five key areas in reading development, and stated that vocabulary was important both in learning to read and in comprehending text (NICHD, 2000). The *simple view of reading* describes reading comprehension as the result of decoding and spoken language comprehension, which is highly dependent on vocabulary knowledge (Gough & Turner, 1986). During the early stages of literacy development, reading comprehension and language comprehension have a large amount of overlap, but this overlap diminishes as children progress in school. As text complexity increases, students encounter a higher proportion of low-frequency words that would not typically be used in spoken language (Vadasy & Nelson, 2012). Consideration of this shift within the framework of the simple view of reading would suggest that as children begin to master decoding skills and are introduced to more complex texts, language comprehension becomes the primary factor in reading comprehension (Hoover & Gough, 1990).

The correlation between reading comprehension and vocabulary knowledge has been reported in reading research for many years (Davis, 1944). Throughout the lifespan, individuals who demonstrate higher vocabulary skills also tend to demonstrate higher reading comprehension skills (Nation, 2009), and this correlation becomes stronger as individuals get older (Torgesen, Wagner, Rashotte, Burgess, & Hecht, 1997). Scarborough (2001) showed that performance on vocabulary assessments in kindergarten and first grade are strongly predictive of reading comprehension in third and fourth grade. Additionally, Cunningham and Stanovich (1997) suggested that results of early vocabulary assessments can be predictive of reading performance as late as eleventh grade. Conversely, research also shows that reduced vocabulary

knowledge may contribute to an individual's failure to comprehend a text (Cromley & Azevedo, 2007). While researchers have suggested a variety of other factors that may be related to reading comprehension, such as working memory (Cain, Oakhill, & Bryant, 2004), the correlation and predictive relation between vocabulary knowledge and reading comprehension has persisted (Wagner, Muse, & Tannenbaum, 2007).

Vadasy and Nelson (2012) stated that vocabulary knowledge is the strongest predictor of reading comprehension. During early childhood, most vocabulary is introduced through oral communication, and spoken language forms the foundation for emergent literacy (Nagy & Anderson, 1984). Consequently, reading comprehension and spoken language comprehension are highly correlated in young readers (Sticht, Beck, Hauke, Kleiman, & James, 1974). Many researchers have suggested that it is not until third or fourth grade that vocabulary knowledge begins to influence reading comprehension, as the texts begin to incorporate more words that are uncommon in spoken language (Chall, Jacobs, & Baldwin, 1990; Scarborough, 2001; Storch & Whitehurst, 2002). However, more recent research has highlighted the impact of vocabulary knowledge on reading comprehension at even younger ages.

Lindsey, Manis, and Bailey (2003) examined a sample of 249 kindergartners who were native Spanish-speaking, English language learners. This study followed the students from the beginning of kindergarten through the end of first grade. Lindsey and colleagues found that expressive vocabulary measured at the beginning of kindergarten was a good predictor of word identification skills measured at the end of the first grade. It is important to note that this pattern of results was bidirectional; the results were consistent both when predicting from English to Spanish and from Spanish to English. This supports the theory of cross-linguistic transfer of

literacy skills and suggests that the importance of vocabulary knowledge is common across the two languages.

Extending the research of Lindsey et al. (2003), Chiappe, Chiappe, and Gottardo (2004) examined both the expressive and receptive vocabulary skills of good readers and poor readers in first to third grade. Poor readers ($n = 13$) were defined as students who achieved word identification scores below the 26th percentile, and good readers ($n = 49$) were defined as students who achieved word identification scores above the 29th percentile. Results suggested that expressive vocabulary skills were more highly correlated with measures of phoneme blending ($r = .35$), phoneme deletion ($r = .48$) and both word ($r = .44$) and nonword identification ($r = .41$) than receptive vocabulary ($r = .22$, $r = .11$, $r = .19$, and $r = .22$, respectively), when controlling for age. It was also found that the poor readers performed significantly lower on measures of expressive vocabulary than the good readers, although the two groups did not differ significantly on measures of receptive vocabulary. These findings may suggest that expressive vocabulary skills are more influential to literacy development.

Graves and Silverman (2011) identified four different types of vocabulary knowledge that are characterized by two distinguishing features: knowledge that is expressive or receptive, and knowledge that is spoken or written. Therefore, each individual has vocabulary knowledge of words produced in spoken language, words heard in spoken language, words written in printed text, and words read in printed text. The authors explained that the distribution of words across these categories changes for individuals over time, but both children and adults consistently maintain larger receptive than expressive vocabularies. Although receptive vocabulary skills may encompass a greater number of total words, Chiappe et al. (2004) hypothesized that expressive, spoken vocabulary tasks may correlate more highly with literacy because oral

production of a word demands a more complete phonological representation than both written production tasks and receptive tasks that rely on recognition.

The influence of vocabulary knowledge on reading comprehension has also been documented beyond the controlled environments of research. The Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good & Kaminski, 2002) is one of the most common tools used to assess students' early progress in learning to read in elementary schools across the United States. Riedel (2007) found that approximately 15% of students, in his sample of 1,518 first graders from a large urban school district, that obtained satisfactory Oral Reading Fluency (ORF) DIBELS scores at the end of first grade did not achieve satisfactory comprehension scores. The distinguishing feature among students with satisfactory ORF but poor comprehension was poor vocabulary. These findings suggest that vocabulary plays an essential role in reading comprehension, and demonstrate the need for additional, enhanced vocabulary intervention within the literacy curriculum.

CHAPTER 2

LITERATURE REVIEW

Theoretical Framework

Although the correlation between reading comprehension and vocabulary knowledge is clear, there is some debate over how to interpret the relation. Some research suggests that an individual's reading ability is a determining factor of his or her vocabulary development, such that weaker reading abilities may lead to less print exposure and, over time, reduced vocabulary development (Cunningham & Stanovich, 1997; Nagy & Scott, 2000). Other research suggests that an individual's vocabulary knowledge is a determining factor of his or her reading comprehension, such that reduced vocabulary knowledge may result in difficulty accessing word meanings while reading, and consequently limit overall reading comprehension (Beck, Perfetti, & McKeown, 1982).

The present study is aligned with a third interpretation of this correlation; vocabulary knowledge and reading comprehension are correlated because they share underlying language processes. McGregor (2004) explained that vocabulary acquisition is dependent on access to cognitive mapping that links phonology and semantics. Similarly, Perfetti (2007) suggested that reading comprehension, and reading ability more broadly, depends on access to cognitive mapping that links phonology and semantics to an orthographic code. This theory supports an integrated vocabulary and reading intervention.

Multi-linguistic approach. Children's proficiency in multiple aspects of linguistic awareness contributes to their literacy development (Apel & Masterson, 2001). Researchers have reported that phonological awareness, morphological awareness, and vocabulary knowledge influence children's ability to read and spell (Castles & Coltheart, 2004; Wolter,

Wood, & D'zatko, 2009). However, much of the current research regarding the influence of linguistic awareness skills on the development of literacy abilities has explored only one of these linguistic skills, in an isolated context (Kim, Apel, & Al Otaiba, 2013). While these studies have contributed valuable information regarding the individual roles that component skills of linguistic awareness play in successful literacy development, additional research is needed to explore how these skills interact and influence literacy development concurrently as an individual learns to read.

Consistent with Perfetti (2007), the connectionist model of reading outlines three linguistic systems that are essential for successful reading: phonology, orthography, and semantics (Plaut, McClelland, Seidenberg, & Patterson, 1996). The role of phonology in reading is prevalent in the literature and widely accepted (NICHD, 2000) and vocabulary knowledge (semantics) is often hypothesized to interact with phonology to contribute to students' reading performance (Nation & Snowling, 2004; Oullette, 2006). The connectionist model also highlights the need for orthographic awareness during literacy development, because both reading and spelling involve encoding and decoding letter patterns. However, due to inconsistent grapheme-phoneme correspondence in English (Kim et al., 2013), the present study proposed that morphological awareness might have the potential to contribute more meaningful information than orthographic awareness when learning new words.

Berninger and colleagues (2003) compared a phonological awareness intervention with an integrated phonological and morphological awareness intervention. The phonology treatment emphasized word building through listening for and blending sounds. The integrated treatment also highlighted sounds, but added a second level of word analysis by highlighting grammatical morphemes. Twenty students in fourth to sixth grade who read at least one standard deviation

below their verbal IQ were randomly assigned to either the phonological or integrated treatment group. Results suggested that an integrated intervention including phonological and morphological awareness improved word recognition abilities of students with persistent reading difficulties.

Kirk and Gillon (2009) evaluated the effects of a similar literacy intervention program that highlighted morphological awareness using a multiple linguistic approach with younger children. Sixteen children between the ages of 8 and 11 years old who demonstrated spelling difficulties were randomly assigned to either an experimental or control group. The experimental group received an intervention focused on morphological awareness including orthographic patterns that emerge when adding suffixes to base words. The control group did not receive intervention until after the experimental group had completed the program. Results showed that the experimental group made significantly greater gains in reading and spelling accuracy than the control group on both experimental and standardized measures. Participants in the experimental group also demonstrated generalization of morphological knowledge to new words. Kirk and Gillon concluded that developing morphological awareness within a multiple linguistic approach can improve literacy intervention. Based on this work, the present study adopted a modified connectionist framework to incorporate a multi-linguistic approach to vocabulary intervention that emphasized phonological, morphological, and semantic knowledge to support literacy development.

Multimodal approach. It is well documented that spoken language and literacy skills demonstrate reciprocal benefits such that improvement in one, positively influences the other (Sénéchal, LeFevre, Smith-Chant, & Colton, 2001). Research concerning students with autism (Lanter & Watson, 2008) and English language learners (Mathes, Pollard-Durodola, Cárdenas-

7

Hagan, Linan-Thompson, & Vaughn, 2007) has incorporated a multimodal approach to literacy instruction. The present study suggested extending this application to all students struggling with reading in the general education classroom.

Lanter and Watson (2008) composed a tutorial for speech-language pathologists (SLPs) working on language and literacy goals with school-aged students and adults with autism, ranging in reading ability from the emergent reading stage to a skilled reading stage consistent with typically developing fourth graders. The strategies that the authors suggested were intended to support the development of both reading and oral language skills, as these are commonly designated areas of need among many students with autism. Story retelling was recommended as a strategy to support both independent reading and listening comprehension skills, as well as oral storytelling of past events because it provides students an opportunity to organize information in an expressive modality while reinforcing a narrative story schema (Roth & Baden, 2001). In addition to practice utilizing the content information using spoken language, the retell task improves schema recognition, which has been shown to improve narrative comprehension in typically developing students (Mandler & Johnson, 1977). Lanter and Watson also suggested utilizing discourse to support oral language skills during literacy intervention and encouraged teachers and SLPs to integrate visual aids to help students access information through multiple modalities.

Mathes et al. (2007) synthesized the results from four Tier 2 vocabulary intervention studies in order to provide a resource for SLPs working with native Spanish-speaking first graders who are struggling to develop early literacy skills. Tier 2 words are academic vocabulary that students might encounter in multiple subject areas (Beck, McKeown, & Kucan, 2002). The instructional components were organized into *strands* of reading-related abilities,

8

according to design principles including: phonemic awareness, orthophonemic knowledge (letter-sound correspondence), word recognition, fluency, and comprehension. Instruction in each strand incorporated visual aids, such as images that depicted target words, and consistent spoken language modeling. In addition to activities that addressed each strand, the final component of the intervention was a 10-minute storybook routine intended to provide students an opportunity to listen for new words, orally retell the story, and engage in dialogue with the teacher about the story using the new words. Mathes and colleagues concluded that English language learners benefited from an integrated instructional model that targeted multimodal interaction with language through visual aids, consistent modeling, structured literacy routines, and informal conversation.

Bereiter and Scardamalia (1982) suggested that the presence of words and sentences on the page may, at times, hinder children's ability to creatively construct ideas, and may even prohibit alternative ways of expressing those ideas beyond linguistic modalities, which is particularly important when considering students with different language-learning needs. Drawing is a nonlinguistic alternative to writing that prioritizes concept formation over linguistic demands (Ukrainetz, 1998). In classrooms, drawing is often recommended as a pre-writing strategy for kindergartners and first graders in order to support brainstorming and organizing ideas, and it has also been observed that students refer back to their drawings for visual reminders of content as they compose written texts (Myers, 1983). In this way, incorporating drawing into a multimodal approach to vocabulary intervention can allow a nonlinguistic avenue to both synthesize and express knowledge related to new words and act as a visual reminder when students try to access that knowledge at a later time.

Consistent with Beck, McKeown, and Kucan's (2002) discussion of the levels of semantic knowledge expressed within a linguistic context, McGregor, Friedman, Reilly, & Newman (2002) explained that children's drawings demonstrated a similar continuum of semantic knowledge. The authors conducted two studies of typically developing 5 year olds ($n = 25$ and 16) to compare picture naming and picture drawing performance with relation to semantic knowledge. They asked participants to name (orally label), draw, and define age-appropriate objects. Results suggested that drawings might be a valuable source of information when evaluating children's semantic knowledge. Given the Theoretical Model of Lexical Storage and Processing described by the authors, in which drawings utilize semantic representations via the visual modality, this type of nonlinguistic processing may be especially important when considering intervention with students who have a variety of language and learning differences.

Interactive approach. Vygotsky's (1978) social interaction theory provides a framework for understanding how interactive activities can create meaningful opportunities for students to learn new words. For many years, developmental research focused on decontextualized skills of children, with the skills and child isolated from natural interactional contexts (Rogoff, 1990). Rogoff argued that children's development is greatly influenced by collaboration with peers and more-skilled partners within a sociocultural context. Her concept of *guided participation* suggested that elements of both leadership from a partner and active engagement from the child are essential to a collaborative learning model. Rogoff continued to explain that this collaborative process encourages children to build upon their current knowledge and abilities in order to reach new levels of understanding and skill.

Fleming and Alexander (2001) investigated the benefits of peer collaboration among 31 fourth-grade students in two public schools in Sydney, Australia. Fifteen of the subjects experienced the peer-collaborative intervention while the remaining sixteen subjects served as a control group. Participants were matched and randomly assigned to the treatment or control group. Then, small groups of three or four students each were formed to participate in collaborative teams. The primary dependent variable was a word recall task in which students dictated words from memory to the researcher. Results replicated similar findings from a U.S. sample (Manion & Alexander, 1997; as cited in Fleming and Alexander, 2001). The authors concluded that talking about academic content in collaborative groups resulted in children gaining a better understanding of the information, and improved recall when compared to children working alone.

Applying this concept of collaboration to children with language impairments, Brinton and Fujiki (2006) proposed a collaborative learning model within the classroom as a means through which to overcome language deficits. The authors explained that students with language impairments often struggle with classroom activities and could benefit from working closely with other students. Allowing students multiple opportunities to learn with and from peers, could support progress towards both academic and social goals.

Although the potential benefits of collaboration within the classroom have been discussed in research for many years, much of the classroom curriculum continues to rely on independent work, with peer collaboration even viewed as cheating in some classrooms (Dyson, 2013). The present study suggested incorporating an interactive instructional activity within small-group vocabulary intervention to support robust semantic mapping through meaningful social interactions.

Effective Vocabulary Intervention

Nagy and Anderson (1984) calculated that printed school English contains approximately 88,500 word families. When proper nouns, words with multiple meanings, and idioms were included in the calculation, that number increased to an estimated 180,000 words in printed school English (Anderson & Nagy, 1992). Considering the immense number of words and the variety of experiences students have with print, Beck et al. (2002) listed five levels of word knowledge:

1. no knowledge;

2. a general sense of meaning or connotation;

3. a narrow, context-bound knowledge;

4. a basic knowledge of and ability to use in meaningful and appropriate ways;

5. a rich, decontextualized knowledge and extension to related ideas (p.10).

The extant literature includes words that reach or exceed basic knowledge (Level 4) to be indicative of the student's vocabulary (Graves & Silverman, 2011).

Nagy and Herman (1987) estimated that third graders typically have a reading vocabulary of about 10,000 words and learn 3,000 words each year. White, Graves, and Slater (1990) extended this research to investigate spoken vocabulary skills of first to fourth grade students. Students who attended a low-SES school had vocabularies of approximately 13,000 words, while students who attended a higher-SES school had vocabularies of approximately 19,000 words. While these types of large vocabulary measures certainly have limitations, they tend to be biased towards vocabulary that is necessary for success within the current educational system. For example, Chall et al. (1990) observed that students attending low-SES schools began to significantly fall behind in reading performance compared to higher-SES peers in fourth grade.

This gap continued to widen as students got older, and the component of reading development that showed the greatest decline was semantic knowledge.

Due to the overall size of the lexicon, direct vocabulary instruction cannot function as the only source of new word learning, but may have the potential to support student performance within an academic context, and bolster skill development for students with language and literacy differences. Nagy and Herman (1987) stated that many studies of vocabulary instruction have resulted in reliable gains in reading comprehension. Based on this research, the authors suggested five principles for direct vocabulary instruction: multiple exposures to target words, experience with words in meaningful contexts, rich information about each word, connections to students' background knowledge or personal experiences, and active participation of students.

Building upon these foundational principles, Graves (2006) developed an instructional model in which four elements of effective vocabulary intervention were highlighted: providing rich and varied language experiences, teaching individual words, teaching word learning strategies, and fostering word consciousness. The first component, providing rich and varied language experiences, requires teachers to create a classroom literacy environment in which students are exposed to words in multiple contexts through active engagement with written and spoken language, and are given opportunities to co-construct vocabulary knowledge through social interactions (Stahl & Nagy, 2006).

The second component of Graves' (2006) instructional model, teaching individual words, requires teachers to provide instruction on specific high-utility academic words for mature language users (Tier 2) and more domain-specific words that are central to understanding a particular text or topic area (Tier 3; Beck et al., 2002). Nagy and Townsend (2012) suggested that academic vocabulary can be used as a tool for communicating and thinking about curricular

13

content, and interventions that effectively teach these words have shown success in helping students learn more broadly.

The third component of Graves' (2006) instructional model, teaching word learning strategies, requires teachers to encourage students to analyze linguistic and contextual clues to infer word meanings. Beck and McKeown (2001) suggested using discourse to highlight word features and help students meaningfully incorporate their background knowledge to connect ideas, and deepen their understanding of new words. The final component of Graves' (2006) framework is fostering word consciousness. Graves and Watts-Taffe (2002) defined word consciousness as "an interest in and awareness of words and their meanings" (p. 144). This component requires teachers to integrate engaging language activities that motivate students to become fascinated with new words (Cunningham, 2014).

Multimodal instruction is important to consider when fostering word consciousness. In order for students to successfully engage with vocabulary, they need to recognize and interact with language across meaningful reading, writing, speaking, and listening opportunities. An in-depth understanding of the bridge between receptive and expressive language abilities and literacy development positions SLPs as knowledgeable, skilled resources capable of supporting students with vocabulary difficulties within the literacy curriculum (Lanter & Watson, 2008).

McKeown, Beck, Omanson, and Pople (1985) worked with fourth grade students to compare three instructional conditions: (a) traditional vocabulary instruction with an emphasis on learning dictionary definitions, (b) rich vocabulary instruction focused on exploring multiple aspects of word meanings, and (c) extended instruction which combined rich instruction and word consciousness. In this study, word consciousness was encouraged through discussion of students' knowledge of and experiences with target words, including examples of encounters

with target words in contexts outside of the classroom. Participants were students in four separate classrooms within the same small urban school district. Each classroom was randomly assigned to a condition: one classroom was designated for each of the instructional conditions and the fourth classroom was designated as the control group. Effects of instructional conditions were evaluated through a multiple choice definition test and a story retell. Results indicated that students receiving the extended instruction performed highest on measures of definitional knowledge and story comprehension.

Blachowicz and Obrochta (2005) incorporated word consciousness into vocabulary instruction with first graders. Their intervention utilized a topic-based approach and emphasized engaging students in discussions about the topic, generating related ideas and words, and using target words in writing activities. Teachers in this study often used photographs and illustrations to stimulate discussion about students' background knowledge related to the topic and to address any questions they had about specific words. The effect of the intervention was measured by comparing performance on a pre- and post-intervention task requiring students to list, in writing, all of the words that they knew, related to each topic. Students generated significantly more words during the post-intervention assessment. A significant limitation of this study is the absence of a control group. It is unclear what the student performance would have looked like if provided instruction that did not emphasize word consciousness.

Bowers and Kirby (2010) examined the effects of a morphological vocabulary intervention targeting rich and varied exposure to target words for 81 fourth and fifth grade students, in 4 classrooms, who had been randomly assigned by classroom to either treatment or control conditions. The control group continued with typical instruction. The treatment group received instruction from the first author that encouraged students to discover spelling-meaning

15

connections between words using a structured inquiry approach that allowed students to study words in engaging and interactive activities. During instruction, students utilized both spoken and written language, and graphic representations of morphological units were used as visual cues to reduce the demand on working memory. Vocabulary acquisition was measured using two tasks. During the first task, students were asked to circle the main part of a word in order to assess their ability to identify the base word. During the second task, students were asked to provide a spoken definition of the target word to assess their semantic knowledge. Data analyses that controlled for vocabulary knowledge prior to instruction, showed significantly higher performance for students in the treatment condition. Results support instructional techniques that incorporate direct instruction on morphology and provide deep, rich instruction about specific words.

One important factor to consider when teaching specific vocabulary is the structure and content of the word meanings being taught. Graves and Silverman (2011) observed that dictionary definitions often define simple words with more complex words, and are not designed to teach the word to a student who is not already familiar with the meaning. McKeown (1993) investigated the word learning of fifth grade students to compare the effectiveness of traditional dictionary definitions to revised definitions intended to be more student-friendly (Beck et al., 2002). Student-friendly definitions sought to define words using more common terms and simpler syntactic structures. McKeown asked students to perform two tasks after being presented with traditional dictionary definitions or student-friendly definitions. The first task was to write a sentence using the word, and the second task was to respond, in writing, to two questions requiring students to generate associated words or concepts. Results suggested that the revised definitions were significantly more effective at teaching students new words.

Graves (2006), suggested the following guidelines for constructing student-friendly definitions:

1. Use simple language; aim to use words and syntax that are familiar to students.

2. Do not include all possible definitions of a word; focus on what will be useful and meaningful for the students.

3. Use complete sentences.

4. Include an example of the word used in a sentence.

Written definitions. Marinellie and Johnson (2003) investigated the development of both the form and content of students' definitions of adjectives with respect to age and word frequency. Participants were recruited in three age groups: 6th graders, 10th graders and college students. There were 50 individuals in each group, for a total of 150 participants. Target words were determined to be high-frequency if they occurred between 100 and 400 times in the corpus developed by Kučera and Francis (1967). Target words were determined to be low-frequency if they occurred less than 45 times in the same corpus. Participants were asked to compose written definitions for 12 high-frequency adjectives and 12 low-frequency adjectives. Results suggested that all participant age groups were competent in expressing content in written definitions of adjectives. Additionally, participants were more likely to produce a basic definition, such as a synonym, for low-frequency words, but compose more elaborated definitions for high-frequency words. Use of a synonym was most common for sixth graders, and became less common as students got older. Older participants were more likely to produce elaborated definitions that extended beyond a single adjectival synonym.

Marinellie and Chan (2006) extended this research to explore the effect of word frequency on the written definitions of nouns and verbs for 4th, 7th, and 10th grade students.

17

There were 30 individuals in each group. Similar to Marinellie and Johnson (2003), target nouns and verbs were determined to be high- or low-frequency based on the Kučera and Francis (1967) norms. The mean frequency for high-frequency nouns was 334 per million, and for low-frequency nouns was 8.83 per million. The mean frequency for high-frequency verbs was 181.33, and for low-frequency verbs was 2.8. Participants were asked to compose written definitions for six high- and low-frequency nouns and six high- and low-frequency verbs for a total of 24 words. Results supported the conclusions of Marinellie and Johnson (2003) with regard to the influence of word frequency on form and content and included fourth graders as competent writers in definitional tasks. Results also suggested that as students got older, definitions became more socially shared, diverging from the highly concrete and context-dependent definitions of younger students.

The present study extended this research to explore the accuracy of written definitions for nouns, verbs, and adjectives composed by second and third grade students, with and without an interactive vocabulary intervention intended to support socially constructed meanings for new words.

Trends in Reading Comprehension Intervention

Duke, Pressley, and Hilden (2004) provided a comprehensive review of reading comprehension research, revealing four factors that influence comprehension abilities: decoding skills, spoken language proficiency, utilization of comprehension strategies, and motivation. Recent intervention designs tend to focus on at least one of these factors. Almasi, Palmer, Madden, and Hart (2011) compiled a review of reading comprehension intervention research and concluded that interventions could be grouped in three primary categories: directly targeting

comprehension, indirectly targeting comprehension through oral reading fluency, and targeting

comprehension as one component of reading within a larger literacy program.

Throughout the 1980s, top-down frameworks that emphasized the importance of higher-

level cognitive processes in overall reading comprehension became more prevalent (Stanovich,

1980). Researchers began to consider instructional techniques that targeted individual

comprehension strategies that could be directly taught and then self-regulated by students during

the reading process (Almasi et al., 2011). For example, several studies during this time utilized

single-case designs to explore the effectiveness of story grammar instruction on reading

comprehension.

Newby, Caldwell, and Recht (1989) measured the number of story elements children with

dyslexia included in their narrative retells prior to and following direct instruction focused on

story grammar. Similarly, Idol and Croll (1987) investigated the effects of story grammar

instruction for struggling readers, by comparing pre- and post-intervention results of reading

comprehension, listening comprehension, and narrative retell tasks. Both studies had

inconclusive results. Almasi et al. (2011) reviewed intervention studies that examined teaching

struggling readers a specific text-level strategy, such as recognizing story grammar, summarizing

the story, or visualizing events in the story. These authors concluded that while such

interventions may have resulted in short-term improvement in reading comprehension, they were

not successful for long-term maintenance.

In the late 1980s, the United States Department of Education's Office of Educational

Research and Improvement began to focus on phonics instruction, and by 2000 both the National

Reading Council and the National Reading Panel had published reports that discussed the

relationship between fluency and comprehension (Almasi et al., 2011). These events contributed

to a greater number of researchers examining the effect of oral reading fluency on comprehension. Allinder, Dunse, Brunken, and Obermiller-Krolikowski (2001) gathered reading comprehension data on 50 seventh grade students, including 14 students with learning disabilities, in two groups. One group received instruction related to reading fluency (i.e., reading at an appropriate rate) while the second group did not receive any fluency related instruction. There was no significant difference in reading comprehension between groups following instruction. Based on previous research indicating a high correlation between oral reading fluency and comprehension, the authors concluded that fluency instruction may benefit students' reading comprehension, but perhaps only if administered within a balanced literacy curriculum.

Recent research has addressed more comprehensive literacy programs that target multiple aspects of literacy development. Moats (2004) evaluated the effects of an instructional program that integrated phonology, semantics, syntax, pragmatics, and discourse processing skills on reading comprehension of 6th to 10th grade students characterized as struggling readers. Moats compared pre- and post-intervention comprehension scores on standardized assessments and found significant gains at each grade level. This study did not include a control group, so it is uncertain what proportion of gain was a direct result of the instructional program.

Almasi et al. (2011) stated that many studies of comprehensive literacy programs incorporate direct instruction from an additional professional or tutor within the academic setting. The authors explained that research suggests interventions that utilize highly trained teachers, tutors, or other professionals have consistently found statistically significant differences for low achieving students on standardized measures of comprehension. Tutoring programs that used untrained volunteers or tutors were less successful.

Spoken story retells. Oral story retelling is a common measure of reading comprehension in recently published, informal reading inventories (Leslie & Caldwell, 2006; Stieglitz, 2002). However, there is concern that retellings may underestimate the comprehension of some children because of linguistic production demands required by the retelling process (Francis, Fletcher, Catts, & Tomblin, 2005). Nevertheless, assessments are increasingly including retelling tasks because of criticism that comprehension questions are often dependent on additional knowledge beyond what is presented in the text, or do not match specific levels of comprehension appropriate for students of a given age and reading ability (Fuchs, 1992). Although story retelling is a more open-ended and less formalized task, Fuchs, Fuchs, and Maxwell (1985) found correlations between 0.76 and 0.82 for retelling procedures scored based on total number of words, percent of content words, or percent of idea units, and a standardized reading comprehension test. Leslie and Caldwell (2006) also found significant correlations between spoken retelling and reading comprehension as measured by direct questions, for middle school and high school texts.

Utilizing oral retelling tasks to assess reading comprehension is also common in the existing reading literature. Kintsch and Van Dijk's (1978) model of comprehension evaluated performance on the story retell task based on the number of propositions (i.e., a predicate and one or more arguments such as agent, object, or goal) retold by the reader. Stein and Glenn, (1979) discussed measuring comprehension through identification of the elements of story grammar (i.e., character, setting, goal, problem, events, and resolution) that were included in the retell. Other researchers have examined recall based on the total number of words retold (Fuchs, Fuchs, & Hamlett, 1989), and the percentage of content words (exact matches or synonyms of words in the text) retold (Fuchs et al., 1985).

21

Fuchs et al. (1989) monitored reading growth by measuring student performance on a story retell task after reading 400-word folktales. Forty-four students from 22 classrooms between third and ninth grade participated in the study. All students had current IEP goals related to reading and were identified as performing at least 1 year below grade level. Performance on the task was determined by number of matched words (i.e., every word in the student's retell that was the same as in the original passage) and total number of words included in the student retell. Classroom teachers monitored performance on the story retell tasks, with support from the research team, and used student performance to inform IEP goals and implement instructional modifications. Results indicated that story retell was a reliable measure of reading comprehension and can be used to inform academic goals and instructional recommendations.

Fuchs et al. (1985) explored the validity of story retell measures with 70 boys in fourth to eighth grade with mild to moderate learning disabilities. The same 400-word folktales utilized in Fuchs et al. (1989) were used as stimuli in this study. Performance on the story retell task was measured in three ways: total number of words retold, percentage of content words retold (calculated by dividing the number of matched content words and synonyms included in the retell by the total number of content words in the original passage), and percentage of ideas recalled. All three measures were comparably related to results of an open-ended question-answering task (*mean r* = .69; *range* = .64 to .75). Results suggested that attending to the content words that were included in the student retell provided a more informative measure regarding the quality of the retell than the total number of words.

While Fuchs and colleagues (1985) examined the number of content words in students' story retells that were the same as, or synonyms of, content words in the original passage as a

measure of comprehension, there is a paucity of literature utilizing this type of measure to indicate vocabulary acquisition. The present study used retell both as a measure of reading comprehension and as an indication that students acquired at least Level 4 semantic knowledge of the target words (Beck et al., 2002), allowing them to produce the words during a self-generated, spoken story retell.

Speech-Language Pathologists in the Classroom

School-based speech-language pathologists (SLPs) are facing increasingly challenging caseloads in elementary schools and are often no longer able to meet the needs of all of their students through a traditional one-on-one pull-out service delivery model. Boyle, McCartney, Forbes, and O'Hare (2007) compared four service delivery models for elementary children receiving speech and language services targeting vocabulary goals in the United Kingdom including: individual treatment with an SLP (direct-individual), group treatment with an SLP (direct-group), individual treatment with a speech-language pathology assistant (SLPA; indirect–individual), and group treatment with an SLPA (indirect–group). Results suggested minimal differences in standardized test scores following direct services provided by an SLP and indirect services provided by an SLPA who was given an instructional plan from an SLP ($d < .15$, which indicates a small effect size). Test scores were slightly lower following group therapy when compared to individual therapy provided by either an SLP or an SLPA, but the effect diminished after 1 year (at the immediate post-test, $d = -0.10$; at the 12-month follow-up, $d = 0.01$). The authors concluded that SLPs and SLPAs were able to administer effective vocabulary intervention, and both individual and group service delivery models were effective.

Throneburg, Calvert, Sturm, Paramboukas, and Paul (2000) compared the effects of curricular vocabulary instruction, in kindergarten through third grade, provided via three service

delivery models: a collaborative classroom model utilizing team-teaching by the SLP and the classroom teacher; a push-in model in which the SLP provided whole-class instruction to a class of students without collaborating with the classroom teacher; and a traditional pull-out model in which the SLP worked with students individually outside the classroom. Results suggested that students in the collaborative classroom condition exhibited the greatest gains in vocabulary. The effect size was largest when the collaborative classroom model was compared to the push-in model without collaboration ($d = 1.65$). There also was a moderate effect size when the push-in model without collaboration was compared to pull-out services ($d = 0.3$). The present study further investigated the push-in model, modified for small-group instruction.

The collaborative-consultation approach is supported in the literature as a way for SLPs and teachers to provide more cohesive instruction to students with language and literacy needs (Cabell, McGinty, Breit, & Justice, 2008; O'Connor, Notari-Syverson, & Vadasy, 1996). Girolametto, Weitzman, and Greenberg, (2012), investigated the efficacy of a professional development model in which SLPs trained classroom teachers in language and literacy intervention methods targeting concepts about print, orthography, and phonological awareness. Twenty early childhood educators in separate classrooms participated in the study, and were randomly assigned to either the treatment or control group. Participants in the treatment group participated in an 18-hour professional development program that included three classroom visits for individualized coaching. Participants in the control group were placed on a waiting list throughout the study and participated in the training program after completion of the study. Pre- and post-intervention measures were taken to assess teacher behaviors during shared storybook reading and writing activities. Measures included the rate of occurrence for three types of referencing during instruction: print reference, grapheme reference, and sound reference.

Results showed that teachers in the treatment group incorporated significantly more references to letters and sounds during their instruction than teachers in the control group. The authors concluded that SLPs could positively influence literacy instruction in early childhood classrooms through professional development programs.

Additionally, research has shown evidence in favor of co-teaching models partnering SLPs and classroom teachers during instruction within preschool and kindergarten classrooms Justice and Kaderavek (2004a) proposed an embedded–explicit intervention model for early literacy emphasizing an integrated approach to provide students with meaningful opportunities to learn, both embedded within naturalistic literacy experiences and more structured intervention activities that explicitly target critical emergent literacy goals. The role of the SLP in this model comprises elements from collaborative consultation models (indirect service delivery) and direct service delivery through intervention for specific literacy targets. The embedded–explicit model discussed by Justice and Kaderavek (2004a; 2004b) frames early literacy intervention as a responsibility that is shared by the classroom teacher and the SLP. However, there is a need for further research regarding classroom-based service delivery models for language intervention in older elementary classrooms.

Current vocabulary research focused on students with language and learning disabilities suggests that explicit instructional approaches may be integral to new vocabulary learning. Students with language and learning difficulties may require more instructional time than other students in the classroom, in order to improve and maintain vocabulary knowledge (Jitendra, Edwards, Sacks, & Jacobson, 2004). Pany, Jenkins, and Schreck (1982) found that providing repeated opportunities to practice using new words within meaningful and supportive contexts was essential to successful vocabulary instruction in this population. Because of their knowledge

of language and literacy strategies, SLPs may be able to supplement the classroom instruction for students who require more intensive intervention using a small-group push-in service delivery model to provide intervention that utilizes Graves' (2006) framework, emphasizing direct instruction and opportunities for practice, in order to support vocabulary development and ultimately, improve reading comprehension.

Single-Case Design

Recent publications have emphasized the importance and strengths of single-case research as an experimental methodology (e.g., Horner et al., 2005; Kratochwill et al., 2010). Single-case design is particularly well suited for literacy research (Cihak, 2011). Cihak explained that single-case studies allow researchers to collect multiple measurements before, during, and after an intervention in order to clearly establish the effect of the intervention on each individual's performance. Cihak continued to elucidate the origin of single-case designs as stemming from dissatisfaction among researchers because the results of traditional group studies were often inconsistent with the performance of individuals. This is relevant to literacy research because the multiple factors influencing reading ability can cause immense individual differences within a group of students.

Valencia (2011) suggested that one way of thinking about the complex nature of literacy development is through reading profiles. Valencia described a reading profile as a compilation of reading related skills and strategies for an individual student that characterizes his or her strengths and weaknesses and highlights patterns that may emerge to help explain student performance. The need for this more in-depth record of student abilities grew out of a concern that some researchers oversimplify the prerequisites to skilled reading (Spear-Swerling, 2004)

and that this oversimplification, along with data-averaging in group designs, may have contributed to ambiguous results in many intervention studies (Lipson & Wixson, 1986).

The present study developed an assessment profile (including language and reading measures) for each participant, and utilized research design and data-analysis procedures derived from single-case experimental research methodology (Kazdin, 2011). Single-case designs are characterized by three primary features: an individual "case" is the unit of data analysis, the case serves as its own control within the experiment, and the dependent variables are measured repeatedly within and across conditions (Kratochwill et al., 2010). Repeated measurements allow for any potential effects to be replicated in order to control threats to internal validity; control is assumed when the data show three demonstrations of the experimental effect at three different points in time (Horner et al., 2005). Experimental effects are determined through visual analysis of the level, trend, and variability of the data.

Visual analysis of graphic displays is the hallmark for interpreting the effects of an intervention in single-case experimental design (Kazdin, 2011). The intervention is typically designed to reduce an undesired behavior or increase a preferred behavior (Horner et al., 2005). The rate, or total number, of occurrence for those behaviors is then graphed for each participant across time, for all conditions within the study. Single-case design participants serve as their own controls when evaluating change across time or condition (Kratochwill et al., 2010). Line graphs are the most commonly used graphic display for presenting data in single-case research. Performance at a given moment in time is plotted as a single data point and connected to subsequent data points within the same condition as the study progresses (Kazdin, 2011).

As previously stated, visual analysis involves evaluation of trend, level, and variability of the data. *Level* refers to the mean score of the data within a phase; *trend* refers to the slope of the

best-fitting line through the data within a phase; and *variability* refers to the range, or standard deviation, of the data within a phase (Kratochwill et al., 2010). It is also important to attend to any immediate or abrupt change in level or trend. According to Gast and Spriggs (2010), trend is considered the most important feature for researchers conducting visual analysis of data. Variability in the data can provide additional information regarding the amount of overlap of data points across phases.

Evidence standards have been created to specify design criteria for single-case research (Kratochwill et al., 2010). First, in order to reach valid results, the independent variable must be systematically manipulated. The timing of manipulation should allow for stable performance to be achieved within each phase and discourage repeated patterns that may co-occur with or link to other events in the participant's environment (i.e., history effects). Each dependent variable must be systematically assessed over time by more than one observer, and interobserver agreement must be monitored for at least 20% of the data in each condition. Additionally, the design must include at least five data points within each condition to meet evidence standards (three data points would meet evidence standards with reservations).

Rationale and Research Questions

Vocabulary knowledge and reading comprehension share underlying language processes (McGregor, 2004; Perfetti, 2007). Children's proficiency in multiple aspects of linguistic awareness, including phonological awareness, morphological awareness, and vocabulary knowledge, plays an influential role in literacy development. Additionally, previous studies have emphasized the importance of a multimodal approach to intervention to support generalizable skills (Ukrainetz, 1998; McGregor et al., 2002), and research has demonstrated the benefits of social interaction on language and learning for many years (Vygotsky, 1978; Rogoff, 1990).

28

Direct vocabulary instruction cannot function as the only source of new word learning (Nagy & Herman, 1987), but may have the potential to bolster skill development for students with language and literacy differences. Building upon the foundational principles outlined by Nagy and Herman, Graves (2006) developed an instructional model in which four elements of effective vocabulary intervention were highlighted: providing rich and varied language experiences, teaching individual words, teaching word learning strategies, and fostering word consciousness.

Current vocabulary research focused on students with language and learning disabilities suggests that explicit instructional approaches may be integral to new vocabulary learning, and students with language and learning difficulties may require more direct instructional time than other students in the classroom, in order to improve and maintain vocabulary knowledge (Jitendra et al., 2004). Push-in service delivery by SLPs has been shown to be an effective method for vocabulary intervention (Boyle et al., 2007). While collaborative-consultation and whole class co-teaching models for teachers and SLPs have been explored in the literature, further research is needed to determine the effectiveness of small-group intervention delivered by an SLP within the classroom.

Recent research in reading comprehension has emphasized the importance of comprehensive programs that address multiple components of reading development (Moats, 2004). Additionally, research on comprehensive programs has begun to incorporate direct instruction provided by a specialist or tutor separate from the classroom teacher (Almasi et al., 2011). These trends open the door for SLPs to come into the classroom to provide language-based interventions to support literacy development.

The purpose of the present study was to compare the effectiveness of an interactive vocabulary intervention, which incorporated multiple language levels and modalities and was administered by an SLP within classroom literacy instruction, to traditional vocabulary instruction. The existing literature on vocabulary instruction is richly informative, but there are few resources regarding a push-in service delivery model in which an SLP works with small groups of students on vocabulary development within the classroom environment.

Given the demonstrated need for additional support in vocabulary development for many students, the following question was asked:

Question 1: What is the effectiveness of an interactive, multi-linguistic, and multimodal intervention compared to traditional instruction in improving vocabulary acquisition for second and third grade struggling readers?

It was hypothesized that the interactive, multi-linguistic, and multimodal intervention, administered by an SLP in the classroom, would positively affect word acquisition, as demonstrated by an increase in the accuracy of written definitions, when compared to traditional instruction. This change was expected to occur as a result of the multiple exposures to, and meaningful opportunities to interact with, the target words in spoken language, printed text, and pictorial representations. The interactive intervention was also hypothesized to result in a greater number of target words read correctly during oral reading and used during spoken story retelling, because instruction provided the students with a strong mapping system to link the orthographic and semantic knowledge of the target words with their phonological representations. Additionally, the students were given opportunities to practice using the target words in discourse-level expressive language during collaboration with their peers.

The correlation between reading comprehension and vocabulary knowledge has been reported in reading research for many years (Davis, 1944). Individuals with greater vocabulary skills tend to perform better on measures of reading comprehension (Nation, 2009), and the strength of the correlation between reading comprehension and vocabulary knowledge increases as individuals get older (Torgesen et al., 1997). Scarborough (2001) showed that performance on vocabulary assessments in kindergarten and first grade are strongly predictive of reading comprehension in third and fourth grade, and Cunningham and Stanovich (1997) suggested that results of early vocabulary assessments can be predictive of reading performance as late as 11th grade. Given the link between vocabulary and reading comprehension, the following question was asked:

> Question 2: What is the effectiveness of an interactive, multi-linguistic, and multimodal intervention compared to traditional instruction in improving reading comprehension for second and third grade struggling readers?

As a result of the expected increase in word acquisition during the interactive, multi-linguistic, and multimodal intervention, it also was hypothesized that the interactive intervention would be more effective than traditional instruction in improving reading comprehension. Because the target words contributed content information to the short stories, it was posited that students would be better able to understand, retain, and retell the narratives.

The present study contributes to the literature in four important ways. First, although studies have examined phonological, morphological, and semantic aspects of vocabulary learning, few studies have examined all three dimensions in combination. Second, there has been little research on multimodal vocabulary instruction in the general education classroom.

The present study adds to the literature by exploring the potential benefits of an integrated multimodal and multi-linguistic approach to vocabulary instruction with struggling readers.

Third, it has been well established in the literature that language is learned in a social-interactive way (Vygotsky, 1978; Rogoff 1990), yet classroom instruction, particularly beyond kindergarten, continues to organize instruction and assessment around tasks that require students to complete language-based literacy work independently. Fleming and Alexander (2001) provided a promising lead regarding the potential benefits of peer-peer collaborative learning in regular education, fourth grade classrooms. The present study extended this concept to younger students who face challenges with reading comprehension. Lastly, although SLPs have been interested in classroom-based language intervention for decades, implementing collaborative methods between teachers and SLPs has not yet become common practice. Furthermore, when teacher-SLP collaboration occurs in research, to this date it is primarily through an indirect service delivery model, in which SLPs consult with classroom teachers, or through a whole-class direct teaching or co-teaching model, with very young children.

The present study incorporated two additional innovative components: the use of student-friendly definitions as an instructional tool during intervention, and the use of a story retell task as a measure of vocabulary acquisition as well as reading comprehension. The closest study in the extant literature to the present one is the work of Throneburg et al. (2000). In that study, the push-in vocabulary intervention was offered to the entire classroom through a lesson taught by the SLP. In contrast, the present study utilized small-group instruction to provide more intensive intervention, and enhance the impact of the SLP, for students struggling to meet classroom literacy goals. Because vocabulary is one of the most predictive factors in reading proficiency, and skilled reading is essential for academic success, this study proposed a service delivery

model and intervention plan that could supplement classroom instruction, to better support

literacy development for students with language and learning differences.

CHAPTER 3

METHOD

A single-case alternating treatments design, replicated across four students, was employed to measure the effectiveness of an interactive, multi-linguistic, and multimodal vocabulary intervention (Treatment A) administered by an SLP within the classroom environment, to support vocabulary development and reading comprehension. Treatment A was compared to a more traditional method of classroom literacy instruction (Treatment B). The latter focused on independent work with written text. The design of the present study meets all of Horner et al.'s (2005) quality indicator criteria.

Participants

Four 7- to 8-year-old students attending the same elementary classroom in an English speaking, Midwestern laboratory school, participated in the present study. This was an inclusive classroom, in which students with diagnosed disabilities were enrolled. Students were selected through collaboration between the classroom teacher and the author based on low performance on reading comprehension tasks during classroom instruction and on the Developmental Reading Assessment (DRA; Beaver & Carter, 2003). They participated in the study as a small group during reading instruction. Speech, language, classroom behavior, nonverbal intelligence, and reading ability were assessed prior to the start of intervention to gain an understanding of the individual profile of each student. (See Table 1 for a complete list of assessments.) These assessments were administered by the author and three graduate students in speech-language pathology, who will be referred to as examiners. Each assessment was assigned to one examiner to ensure consistency of administration across students.

The Goldman-Fristoe Test of Articulation—Second Edition (GFTA-2; Goldman &

Fristoe, 2000) was administered in order to assess each student's ability to articulate sounds in

spoken language. Results were taken into consideration when determining if students produced

80% of a word's phonemes correctly during a spoken story retell task. Phonemes that were not

produced correctly on the GFTA-2 were excluded from calculations of word production accuracy

during data analysis.

Six subtests of the Clinical Evaluation of Language Fundamentals—Fourth Edition

(CELF-4; Semel, Wiig, & Secord, 2003) were administered to obtain a Core Language Score to

evaluate each student's general language ability, and a Working Memory Index. The subtests

that contribute to the Core Language Score include: Concepts and Following Directions, Word

Structure, Recalling Sentences, and Formulated Sentences. The subtests that contribute to the

Working Memory Index include: Number Repetition and Familiar Sequences. Additionally, the

Peabody Picture Vocabulary Test—Fourth Edition (PPVT-4; Dunn & Dunn, 2007) was

administered to assess receptive vocabulary, and the Expressive Vocabulary Test—Second

Edition (EVT-2; Williams, 2007) was administered to assess expressive vocabulary.

Because the present study incorporated small-group cooperative instruction, the

classroom teacher completed the Teacher Rating Scales from the Behavior Assessment System

for Children—Second Edition (BASC-2; Reynolds & Kamphaus, 2004) to evaluate the behavior

and emotions of the students at school.

The Test of Nonverbal Intelligence—Second Edition (TONI-2; Brown, Sherbenou, &

Johnsen, 1990) was used to estimate general cognition and nonverbal abilities. Finally, reading

ability was measured using the Gray Oral Reading Test—Fifth Edition (GORT-5; Wiederholt &

Bryant, 2012). The GORT-5 measures three aspects of reading: fluency, calculated based on the

amount of time required to read passages aloud; accuracy, calculated based on the number of errors during oral reading; and comprehension, calculated based on the number of correct responses to open-ended questions after reading. All assessments were administered prior to initiation of vocabulary instruction. No assessments were repeated following treatment because changes were not expected. Standardized tests such as these are designed to measure large, stable, general behaviors that are not likely to be substantially altered by relatively brief interventions (Salvia & Ysseldyke, 1978). These assessments were intended to create a profile for each student to better inform instructional needs and data interpretation.

Table 1 is a display of the scores for each of the four participants. Pseudonyms were assigned to maintain anonymity and confidentiality. What follows is a description of their particular strengths and weaknesses in language, reading, and classroom behavior. Other than teacher identified reading difficulty, the only area in which all students demonstrated diagnostically significant assessment results was classroom behavior. The specific type of behavior that presented the most difficulty for each student varied somewhat, although all students were identified on the *critical item* (i.e., specific area that acts as a red flag for larger behavioral concerns) of "easily annoyed by others." It should be noted that classroom behavior was assessed using the teacher rating scale of the BASC-2. Therefore, the two areas (classroom reading and classroom behavior) that were identified as below grade level expectations for all students were both based on teacher report.

Calvin. Calvin, an African American third grade boy, was 8 years, 7 months old at the beginning of the study (see Table 1). He had no known clinical or educational diagnoses and scored within the expected range for students his age on all assessments except for the BASC-2. The area in which Calvin's behavior deviated from the norms was *School Related Problems.*

36

This area is related to a student's ability to complete tasks in the classroom. Additionally, on assessments that included a confrontational naming task (e.g., GFTA-2 and EVT-2), Calvin demonstrated noticeable word retrieval difficulties. He would often initially respond with a related word, self-identify his error, and attempt multiple additional words until he reached the correct response. He did not receive any prompting from the examiner during this process, but was allowed to self-correct his errors. Calvin scored in the 50th percentile on the GORT-5, and performed within grade level expectations across areas of reading rate, accuracy, and comprehension.

Cliff. Cliff, a European American third grade boy, was 8 years, 7 months old at the beginning of the study (see Table 1). He had been previously diagnosed with a visual processing disorder. He scored within the expected range for students his age on most assessments except for the *Word Structure* (5th percentile) and *Formulated Sentences* (16th percentile; borderline) subtests of the CELF-4, and the BASC-2. The area in which Cliff's behavior deviated from the norms on the BASC-2 was *Internalizing Problems.* This area is related to attention. Additionally, across assessments, Cliff demonstrated a longer response time than his peers and benefitted from being given extra time to complete tasks. Cliff scored in the 5th percentile on the GORT-5, and his deficits were consistent across areas of reading rate, accuracy, and comprehension. Consistent application of letter-sound correspondence when decoding appeared to be very difficult for Cliff, resulting in a slow reading rate, a higher number of inaccurate word productions when reading aloud, and low reading comprehension.

Molly. Molly, a European American second grade girl, was 7 years old at the beginning of the study (see Table 1). She had been previously diagnosed with attention-deficit/ hyperactivity disorder (ADHD) and took prescribed medication for the disorder for the duration

of the study. She scored within the expected range for students her age on most assessments except for the *Working Memory Index* (3rd percentile) of the CELF-4, and the BASC-2. The area in which Molly's behavior deviated from the norms on the BASC-2 was *Internalizing Problems.* As mentioned previously, this area is related to attention. Additionally, across assessments, Molly demonstrated hyperactivity and difficulty with impulse control, often needing to get up and move her body or make nonspeech vocalizations. She also displayed several instances of topic perseveration, in which she continuously returned to a topic of interest (e.g., the video game Minecraft) despite multiple attempts at redirection by the examiner. Molly scored in the 50th percentile on the GORT-5, but this score may have been inflated due to her rapid reading pace. She read very quickly, with limited prosodic features, often skipping the words that she did not recognize. Her reading style could be characterized by a quick rate, moderate to high accuracy, and low comprehension.

Elliot. Elliot, a European American second grade boy, was 7 years, 4 months old at the beginning of the study (see Table 1). He had been previously diagnosed with ADHD but was not taking any prescribed medications during the study. He scored within the expected range for students his age on most assessments except for the *Working Memory Index* (16th percentile; borderline) of the CELF-4, and the BASC-2. The areas in which Elliot's behavior deviated from the norms on the BASC-2 were *Internalizing Problems* and *Adaptive Skills.* These areas are related, respectively, to attention and the student's ability to make transitions and adjust to changes in the classroom. Additionally, Elliot was identified on two other *critical items*: "threatening to harm other students" and "physically hitting other students." Across assessments, Elliot demonstrated reluctance to respond to directions, often expressed by using a nearly inaudibly soft speaking voice. He also displayed several instances of aggression and

socially inappropriate language towards other students. Elliot scored in the 50th percentile on the GORT-5, but his reading style differed greatly from both Calvin and Molly. He read with a moderate pace and very high accuracy. He demonstrated strong decoding skills, which allowed him to correctly produce most words, but also resulted in a staccato style in which he read each word as an individual unit. Consequently, his comprehension was lower than his overall score would suggest.

Classroom Setting and Materials

The classroom contained 18 students between the ages of 7 and 9 years old. One lead teacher and one assistant teacher were present in the classroom each day. Intervention occurred during daily literacy instruction for approximately 20-30 minutes in the morning, between 9:30am and 10:30am. The author, a certified SLP, led small-group vocabulary instruction at a table in the classroom, while the remainder of the class participated in other literacy activities led by the teacher and assistant teacher in other parts of the classroom. One graduate assistant (GA) was present to observe during 40% of sessions (two of the five school days each week), (a) to assist with on-site fidelity measures, and (b) to become familiar with intervention materials and procedures, in preparation for a generalization phase at the end of data collection. During the generalization phase, the GA led the small-group instruction.

Intervention materials included six target words per lesson (90 total words; see Appendix A) that were chosen based on word frequency and orthographic complexity. Word frequency was controlled to ensure that the target words highlighted words that second and third grade readers might encounter in printed text, but did not use often in their conversational language (similar to Beck, McKeown, & Kucan's concept of Tier 2 and Tier 3 words, 2002). Orthographic complexity was considered in order to emphasize words that do not adhere to

common spelling patterns (Gillingham & Stillman, 1997) and therefore, may be difficult to read

based on phonics alone. Words were selected from a list of sight words developed by Dr. Holly

Shapiro (personal communication, February 7, 2014). For the selection process, see "Pilot

Study" in this document. The list was composed of 1,680 words that contained irregular spelling

patterns, gathered from *The American Heritage Word Frequency Book* (Carroll, Davies, &

Richman, 1971). The words were organized from most frequent (*i.e., the*) to least frequent (i.e.,

peristalsis). This word list represents a subset of advanced sight vocabulary, a concept proposed

by Marilyn Nippold (1988) as an important factor in acquiring a literate lexicon.

The grammatical class of target words was controlled by selecting 30 nouns, 30 verbs,

and 30 adjectives, such that two words from each part of speech were combined to create each

individual lesson's word set. The words were distributed across lessons in order of decreasing

word frequency (Carroll et al., 1971), i.e., Lesson 1 contained the most frequent words (*M*

frequency = 64.38 per million; *SD* = 40.04) and Lesson 15 contained the least frequent words (*M*

frequency = 1.76 per million; *SD* = 1.48). (For the word frequencies of target words and the

average word frequency per lesson, see Appendix A.)

Two stories were composed for each word set (i.e., lesson; see Appendix B) following a

simple 3-part plot template: (1) introduction of character and setting; (2) main event/problem;

and (3) resolution. Each story was controlled for number of words (90 to110 words) and number

of propositions (18 to 22 ideas; Kintsch & Keenan, 1973). Stories were segmented into

propositions independently, by two undergraduate research assistants (RAs). What constituted a

proposition tended to be a phrase or sentence with a noun, verb, or adjective that carried new

information. When the RAs disagreed, the author was consulted to reach consensus on the

disputed proposition. The difference in total number of words between the two stories within a

lesson ranged from 0 to 5; Lesson 1 demonstrated the largest difference in length (Story 1

contained 103 words and Story 2 contained 108 words). The difference in total number of

propositions between the two stories within a lesson ranged from 0 to 2; Lesson 5 demonstrated

the largest difference in number of propositions (Story 1 contained 19 propositions and Story 2

contained 21 propositions). Each target word was presented only once within each story.

Individual paper copies of the stories were provided to students during the sessions.

One digital audio recorder was placed at the center of the table to record each session.

One additional digital audio recorder was used to record individual students orally reading and

retelling each story. Paper and pencils were used when working with word lists and definitions

during both treatment conditions. A white board and markers were used to facilitate group work

during Treatment A, and iPads were used to access an online dictionary (wordsmyth.net) during

Treatment B. Pencils were used to create illustrations and compose captions for target words on

standard, white copy paper during Treatment A.

Procedures

Pilot study. Two phases of piloting procedures were used to develop the materials and

methods that would be implemented during data collection. During the first phase, a group of

students in first to third grade read words from the complete list of 1,680 potential target words

that were arranged according to decreasing word frequency (Carroll et al., 1971). These students

were then asked if they knew what the words meant. Students were asked to define or generate a

spoken sentence using 30% of the words for which they indicated knowledge. Accuracy was

documented using a simple plus/minus system to indicate if each student read the words

correctly (a plus symbol represented a correct response and a minus symbol represented an

incorrect response). The criterion for correctness was that 80% of the phonemes produced by the

student matched standard adult production of the target word. Data were also recorded to document if each student indicated knowledge of the word, and if each student was able to provide a definition of or a sentence using the target word. This step was intended to inform the selection of the target words that would be used during intervention. No data were taken about the individual students; the goal was solely to determine which words typical students had not yet mastered at this age. Words that were unfamiliar to the pilot group were selected as target words for vocabulary instruction in the main study.

During the second piloting phase, a selection of students from the previously mentioned group of first to third graders participated in three 20- to 30-minute lessons targeting six words. The lessons adhered to the procedures outlined in the following section ("Main study"). Again, no data about the individual students were recorded; this step was solely to verify that the planned treatments could be completed within the 20- to 30-minute time frame and to ensure that all materials and procedures were sufficiently prepared before data collection began.

Main study. For the main study, fifteen lessons were completed by a new group of second and third grade students (from a different school that did not participate during the pilot study), across three phases (see Figure 1). Nine lessons were conducted using the experimental intervention (Treatment A). This treatment is abbreviated DRAW because these lessons were informally referred to as "drawing days" by the students and teachers. Six lessons were conducted using the traditional instruction (Treatment B). This treatment is abbreviated DICT because these lessons were informally referred to as "dictionary days" by the students and teachers. The lessons were distributed across three phases. The first was a training phase (Lessons 1 and 2), during which one lesson was conducted using each of the treatments to familiarize students with expectations and routines. These lessons were not included in the data

42

analyses because students needed varying levels of support to complete the tasks. The second was the treatment phase (Lessons 3 through 12), in which the treatments alternated across 10 lessons, five in each treatment condition. The third was a generalization phase (Lessons 13, 14, and 15), during which the GA led three DRAW lessons to explore the consistency of student performance across interventionists. The GA attended 40% of the days during Lessons 1 through 12. When present, she took data on treatment fidelity and assisted with pre- and post-treatment measures on Days 1 and 3 of each lesson.

Each "lesson" required three days to complete (e.g., Lesson 1 included sessions on Day 1, Day 2, and Day 3; see Table 2). All participants maintained the same lesson schedule because intervention was provided to all students simultaneously. During both DRAW and DICT treatment conditions, students met individually with the author to complete pre-treatment tasks on Day 1 and post-treatment tasks on Day 3. Students met with the author as a small group for treatment on Day 2 of each lesson.

Day 1 of each lesson began with all four students independently reading a short narrative, containing six target words, aloud into an audio recorder. Because there were two stories that corresponded to each word set, the story chosen for Day 1 of the lesson was balanced across students such that two students received Story 1 and two students received Story 2. The students were then asked to retell the story orally into an audio recorder. "Tell me that story again" was used to prompt the students' story retell, and "Do you remember anything else from that story" was used to confirm the completion of the retell. Finally, the students were given a list of the six target words. The SLP read each word aloud and asked the students to write anything they knew about the words on a sheet of paper. Prompts such as "What can you tell me about that word" and "What definition could you write for the word _____" were used. No semantic or syntactic

43

cues were provided. These tasks were used to gather the pre-test measures for the given set of target words.

On Day 2 of each lesson, small-group instruction differed between treatments. During DRAW lessons, students began Day 2 by participating in a discussion about the words, led by the SLP, to highlight how each word looked (e.g., identifying different word parts; morphology) and sounded (phonology), and to make personal connections to the students' background knowledge (semantics). Additionally, semantic information was provided through student-friendly definitions printed on a sheet of paper, consistent with the procedure used by Graves (2006). Students were given several minutes to think about the words and draw a small visual representation of each word on a sheet of paper divided into six sections, three on the front and three on the back of the paper. Finally, students were given an opportunity to explain their pictures to the group and collaboratively write a caption using the target word, for at least one image per word. The student-friendly definitions acted as a resource to support student contributions during discussion and as a reference when students were composing captions. This intervention used the multi-linguistic approach by incorporating direct instruction related to morphology, phonology, and semantics; incorporated multimodality by providing opportunities for students to read, write, speak, listen, and draw, to engage with new words; and employed a social-interactive model, by allowing students to share their drawings with the group and collaboratively compose captions using the written target words in a meaningful context.

During DICT lessons, students began Day 2 of each lesson by looking up definitions for the six target words on an iPad, using an online dictionary (www.wordsmyth.net). This dictionary included beginner, intermediate, and advanced settings. For all DICT lessons, students used the intermediate level of the dictionary which provided informative, but less

44

complex and elaborated definitions than the advanced setting. The Wordsmyth dictionary also included a pronunciation feature, in which the website would provide an auditory model of the word, but students were not permitted to use this function. The students copied the definitions from the website onto a sheet of paper. After completing all six definitions, students were instructed to write a sentence using each word. Each student worked individually on these tasks. These vocabulary-learning activities represent a "business as usual" or traditional classroom instruction model.

Similar to Day 1, Day 3 of each lesson was the same during both treatments. The third day began with all four students independently reading a short narrative, containing six target words, aloud into an audio recorder. Because the narratives were balanced across students on Day 1, on Day 3 each student read the narrative that they had not previously read. The students then retold the story orally into an audio recorder and again, they were given a list of the six target words. The SLP read each word aloud and asked the students to write anything they knew about the words on a sheet of paper. These tasks were used to gather the post-test measures for a given set of target words.

Procedural Fidelity

Fidelity checklists (see Appendix C) were completed during each session by the author and by the GA during or after 50% of the Day 2 sessions. The checklist addressed items related to the consistency of the time and place of (i.e., Setting), duration of (i.e., Dosage), and number of students present during (i.e., Population) the treatment sessions. The number of students present was considered an important component of procedural fidelity because the treatments were intended to be administered with a small group of four students. A change in the number of students present, could have substantially altered the experience of the group. Additionally,

checklist items included procedural elements such as: "Did the interventionist highlight what the word looks like," "Did the interventionist highlight personal connections," "Did students complete all six definitions," etc., that specifically addressed the components of each treatment. All sessions were audio recorded so that fidelity measures could be completed for sessions in which the GA was not present. Treatment fidelity was effectively 100%. The only item that was missed was "Population" (see Appendix C) because Calvin was absent for the final two generalization lessons (Lessons 14 and 15). Reliability between the author and the GA was calculated by dividing the percentage of items in agreement per session by the total number of items scored per session and then averaging the resulting values across sessions. Interobserver agreement between the author and the GA, was 100% on the fidelity checklist.

Measures

In order to answer Question 1, data were collected for three dependent variables (DVs) to measure word acquisition: (1) accuracy of written definitions, (2) word recognition during oral reading, and (3) inclusion of target words during a spoken story retell. In order to answer Question 2, data were collected for a fourth DV to measure reading comprehension: (4) number of propositions included in story retell. Data for all DVs were collected on Days 1 and 3 of each lesson. Day 1 of each lesson was used to gather baseline data for each student's performance for all target words on the word definition, oral reading, and story retell tasks. All DV data were calculated as gain scores (i.e., number of words gained and percent of propositions gained) by subtracting Day 1 performance from Day 3 performance.

Written word definitions were scored as either correct or incorrect. Credit for a word definition was awarded if the definition met three criteria: It (1) contained at least one element that indicated a possible meaning of the word, (2) contained at least one element that indicated

46

the appropriate grammatical class of the word, and (3) was written in a definitional form, as opposed to an example sentence. A single word could fulfill all criteria. For example, for the target word *gallop*, the definition *run* would be scored as correct because it has a similar meaning to the target word, it is a verb, and a synonym is an acceptable definitional form. However, *like a horse* would not be scored as correct because, although it is related in meaning, it does not indicate an appropriate grammatical class. Similarly, *The horse galloped* would not be scored as correct because, although it indicates appropriate semantic and syntactic features, it is not written in a definitional form.

Word recognition during oral reading and the inclusion of target words during a spoken story retell were also scored as correct or incorrect. The students were given credit for a target word during oral reading if their pronunciation was consistent with an intelligible adult production of the word (i.e., virtually 100% of the phonemes were correct). Initially, this scoring was completed using an "80% of phonemes correct" criterion. However, this cutoff allowed for errors in recognition of the irregular spelling pattern to be identified as correct for words longer than four phonemes (e.g., *ravine*; target production = /rəvin/ or "ruh veen"; common error = /rəvain/ or "ruh vīne", meets 80% criterion but fails to recognize and pronounce the irregular vowel correctly). Consequently, student gains in recognition of the irregular spelling pattern from Day 1 to Day 3 were not visible in the gain scores. Therefore, the criterion was increased to more accurately represent fully correct word recognition of the target words. The 80% criterion was maintained in the story retell task, such that a correct production of the target word needed to match the standard pronunciation of the word for 80% of its phonemes (excluding dialectal variations and documented articulatory deviations). Additionally, to be scored as correct, the target word needed to be used by the student in its appropriate grammatical class

(e.g., a noun used as a noun). For example, the target word *promising* was presented as an adjective (e.g., he had a promising career ahead) and therefore the student had to use the word as an adjective, rather than a verb, in his or her retell.

Total gain in the number of propositions included in each story retell was calculated as a gain in percent correct from Day 1 to Day 3, because the number of propositions in each story varied slightly. To determine scoring, the students' spoken story retells were compared to a line-by-line list of all propositions included in the original narrative. For example, the original sentence *Sue ran across the playground* contains three propositions: (1) *Sue*, (2) *ran*, and (3) *across the playground*. Student responses were scored as correct for an idea if the content was present in the story retell. The words chosen to represent the idea did not need to match the original story verbatim. For example, the story retell sentence *Sue went through the playground* would be scored as correct for all three ideas, while *Sue went through the parking lot* would be scored as correct for only the first two propositions. Each proposition was scored as correct or incorrect and a total percent correct for each story was calculated by dividing the number of propositions correct in the story retell by the total number of propositions in the original story.

Interobserver Agreement

Two undergraduate research assistants (RAs) independently scored all definition, word recognition, and story retell tasks, using the written copies of the definitions and the digital audio recordings of students reading aloud and retelling the stories. RAs were undergraduate seniors, majoring in Speech and Hearing Science. They were blind to the specific instructional strategies that were used during intervention and to how treatment conditions were distributed across lessons. Both RAs reached 80% accuracy on three consecutive practice sets before scoring student data. Reliability across scorers was calculated by dividing the number of items in

agreement for each measure per session by the total number of items scored for each measure per session. For items in which the two RAs disagreed, the author acted as a third scorer and resolved all disagreements.

Reliability across scorers was calculated for each measure per session. The average reliability across all 15 lessons for each measure was as follows: written definitions = 93% interobserver agreement (*range* = 83% to 100% per child, per day), word recognition = 99% interobserver agreement (*range* = 96% to 100%), vocabulary retell = 100% interobserver agreement, and proposition retell = 88% interobserver agreement (*range* = 84% to 95%).

Data Analysis

Data are displayed using line graphs for each measure across the four students (see Figures 2 through 6). Figures 2 and 3 are displays of data gathered from the written word definition task, Figure 4 is a display of data from the word recognition task, Figure 5 is a display of data related to the inclusion of target words during spoken story retell, and Figure 6 is a display of the data related to the number of propositions included in story retell. Each data point in the figures represents the amount of change calculated for one dependent variable between Day 1 and Day 3 of each lesson for a particular student. All data were de-identified and converted to digital files that were stored on a secure university network.

Vocabulary acquisition. Recall Question 1: What is the effectiveness of an interactive, multi-linguistic, and multimodal intervention compared to traditional instruction in improving vocabulary acquisition for second and third grade struggling readers? In order to answer Question 1, quantitative data were collected across three dependent variables to measure word acquisition hereafter referred to as: written definitions, word recognition, and vocabulary retell.

Written word definitions were copied, verbatim, from student-produced paper copies into Excel documents by the two RAs. One of the RAs then verified each document as an accurate copy of the original student-produced text, such that each transcript was created by one RA and confirmed by the other. When a discrepancy was found between RAs, the author independently reviewed the document to resolve the difference. Each definition was given either a 0 or 1 score in an adjacent column of the Excel spreadsheet, and gain scores for target words within each lesson were tabulated automatically through programmed equations in the spreadsheet.

Audio recordings of students reading each story aloud were transcribed and coded by the two trained RAs, using the Systematic Analysis of Language Transcripts software (SALT; Miller, 2008). One RA transcribed and a second RA confirmed the accuracy of all transcripts. All attempts at target words were scored using the following codes: C = correct (intelligible production), A = attempted (incorrect pronunciation), and S = skipped (omitted). When a discrepancy was found between RAs, the author listened to the recording to resolve the difference. When multiple productions for a target word were present during oral reading, successfully meeting criteria for any one production resulted in a correct score for that word. All target words and codes were copied into an Excel spreadsheet. Each target word was given either a 0 (if coded as A or S) or 1 (if coded as C) score in an adjacent column of the Excel spreadsheet, and gain scores for target words within each lesson were tabulated automatically through programmed equations in the spreadsheet.

Audio recordings of all spoken story retells were transcribed by the trained RAs using the SALT software. Again, one RA transcribed, the second RA confirmed the accuracy of all transcripts, and the author resolved all discrepancies. All attempts at target words were phonetically transcribed independently by the two trained RAs, to determine if at least 80% of

phonemes in the child-produced target word matched a standard adult production of the word. When a discrepancy was found between RAs, the author listened to the recording to resolve the difference. All phonetic transcriptions of target words were copied into Excel documents. When multiple productions for a target word were present in a retell, successfully meeting criteria for any one production resulted in a correct score for that word. Each target word was given either a 0 or 1 score in an adjacent column of the Excel spreadsheet, and gain scores for target words within the story retells for each lesson were tabulated automatically, through programmed equations in the spreadsheet.

Reading comprehension. Recall Question 2: What is the effectiveness of an interactive, multi-linguistic, and multimodal intervention compared to traditional instruction in improving reading comprehension for second and third grade struggling readers? In order to answer Question 2, quantitative data were collected on the number of propositions included in story retell (hereafter referred to as proposition retell).

The ideas included in each original narrative were used to perform a secondary analysis on the story retell transcripts. Each idea entered in an Excel spreadsheet was given either a 0 or 1 score in an adjacent column. The transcripts were scored and confirmed by the RAs, and the author resolved all discrepancies. Percent correct for each narrative and percent change between Day 1 and Day 3 for each lesson were tabulated automatically.

Consistent with Kratochwill et al. (2010), visual analysis was used to: evaluate level, trend for any systematic change, and variability within each phase; document the presence of an intervention effect, and the amount of overlap between treatments; and examine sudden changes within the data for possible confounding factors.

Visual analysis was then used to determine whether the data showed sufficient evidence of a treatment effect for each of the measured variables. If an effect was found, visual analysis was also used to draw conclusions about the magnitude of the effect. A causal relation was interpreted when the data revealed at least three demonstrations of an effect (Horner et al., 2005). To reach consensus about visual inspection of the data, the author conferred with a recognized expert in the field of Special Education, where single-case research is frequently employed.

Generalization

The final three lessons were led by the GA, a graduate clinician enrolled in the second year of an M.A. program in speech-language pathology at the University of Illinois, to determine generalization of the students' performance across interventionists. The GA observed 40% of the sessions throughout the semester and attended one, 1-hour meeting with the author to discuss questions and concerns regarding implementation of the DRAW treatment prior to the first generalization session. Treatment fidelity was calculated for all of the generalization lessons.

Social Validity

The classroom teacher and student participants were interviewed using a semi-structured format following completion of the generalization lessons (6 weeks later), by the two undergraduate RAs, who did not participate in data collection during the treatment phase and were blind to the particular details of the two treatment methods. The interviews (see Appendix D) requested information from both teachers and students about their general attitudes towards the intervention, opinions about practical applications, judgments about the significance of changes in the students' learning behaviors related to vocabulary and reading comprehension, and evaluations of students' overall interest in language and words. All interviews were

transcribed, and data from the questionnaire and interviews were compiled into a summary report by the author.

CHAPTER 4

RESULTS

Data for vocabulary acquisition and reading comprehension are presented in Figures 2 to 6 and Table 3. Visual analysis was used to compare individual students' performance on vocabulary acquisition and reading comprehension measures across treatment conditions. Trends in and comparisons of student performance across treatments will be reported.

Vocabulary Acquisition

Data were collected across three dependent variables to measure vocabulary acquisition: (1) student performance on a written definition task, (2) word recognition during oral reading, and (3) inclusion of target words during spoken story retell. Total gain scores were calculated for each measure by adding the scores from all of the DRAW lessons within the treatment phase and from all of the DICT lessons within the treatment phase separately for each student, such that each student was given a total gain score for each measure in each treatment condition (see Appendix E). The maximum total gain score for all vocabulary acquisition measures within each treatment was 30 words (i.e., 6 words per lesson, and 5 lessons per treatment). If a student's gain score was negative for a given lesson, indicating higher performance on Day 1 than Day 3 of the lesson, that number was subtracted from the overall score to represent total net gain. Visual inspection was used to analyze main treatment effects from Lessons 3 to 12. Visual inspection was also used to analyze generalization of effects across interventionists from Lessons 13 to 15.

Written definitions. Figure 2 is a comparison of the two treatments with respect to the number of target words the four students gained during each lesson, as measured by the accuracy of their written definitions. Recall that written definitions were scored as correct according to

three criteria (which all had to be met): accurate meaning, appropriate grammatical class, and definitional form.

Calvin. From Lessons 3 to 12, Calvin gained 18 words following the experimental intervention and 3 words following the traditional instructional approach. His gain scores ranged from 2 to 5 words per lesson during the DRAW treatment, and 0 to 2 words per lesson during the DICT treatment. Although his performance was variable within both treatments, visual analysis revealed that a moderate gap in Calvin's performance between the two treatments remained present throughout the study. The high level of performance on the written definition task for the DRAW treatment was maintained through the first generalization lesson. However, due to illness, Calvin was not able to complete the final two generalization lessons.

Cliff. From Lessons 3 to 12, Cliff gained 11 words following the DRAW treatment and 0 words following the DICT treatment. His gain scores ranged from 1 to 5 words per lesson during the DRAW treatment, but he never gained a word during the DICT treatment. Visual analysis revealed a small gap for Lessons 3 to 10 in Cliff's performance between the two treatments. A large gap emerged at Lesson 11. While his performance was somewhat variable during the DRAW treatment, the trend in his DICT performance was flat and consistently 0 across all lessons. Cliff's performance on the written definition task was not consistently maintained throughout the generalization phase. For Lessons 13 to 15 he gained 0, 3, and 0 words, respectively, with the DRAW treatment.

Molly. From Lessons 3 to 12, Molly gained 10 words following the DRAW treatment, but had a net loss of 2 words following the DICT treatment. Her gain scores ranged from 1 to 4 words per lesson during the DRAW treatment, and -2 to 1 during the DICT treatment. Negative scores indicate that she demonstrated greater accuracy in her written definitions on Day 1 of

those lessons. This occurred twice for Molly; she provided fewer definitions on Day 3 than Day 1 of both Lessons 8 and 12. As seen in Figure 2, visual analysis revealed that her data for the DRAW treatment demonstrate an upward trend, while her data for the DICT treatment demonstrate a downward trend. As a result, the gap between treatment conditions started off small, but became larger as Molly progressed through the lessons. A gain score of 1 to 2 words per lesson on the written definition task was maintained for the DRAW treatment throughout the generalization phase.

Elliot. From Lessons 3 to 12, Elliot gained 21 words following the DRAW treatment, but only 1 word following the DICT treatment. His gain scores ranged from 4 to 5 words per lesson during the DRAW treatment, and 0 to 1 during the DICT treatment. Visual analysis revealed stable and parallel trends in his performance across treatments. His scores during the DRAW treatment were consistently 4 words greater than his scores during the DICT treatment, resulting in a gap between treatments that was relatively large in magnitude. Elliot maintained a gain score of 3 to 5 words per lesson on the written definition task with the DRAW treatment throughout the generalization phase.

Grammatical class of words gained. Figure 3 is a comparison of the number of words the students gained for written definitions across the three grammatical classes, during each of the treatment conditions and cumulatively. Across the treatment and generalization lessons, the total number of words acquired by Calvin and Elliot did not differ greatly for grammatical class. Calvin gained 8 nouns, 10 verbs, and 8 adjectives; Elliot gained 11 nouns, 12 verbs, and 11 adjectives. In contrast, Cliff and Molly acquired a greater number of nouns than either verbs or adjectives. Cliff gained 7 nouns, 3 verbs, and 4 adjectives; Molly gained 6 nouns, 3 verbs, and 3

adjectives. Although Cliff and Molly demonstrated a preference towards nouns overall, none of the four students gained any nouns during the DICT treatment.

Word recognition in oral reading. Figure 4 is a comparison of the two treatments with respect to word recognition during oral reading, for the four students during each lesson. Recall that a word was scored as recognized (i.e., correct) as long as it was read aloud as an intelligible adult-like production of the target word.

Calvin. From Lessons 3 to 12, Calvin gained 12 words following the DRAW treatment, but only 2 words following the DICT treatment. His gain scores ranged from 2 to 3 words per lesson during the DRAW treatment, and -1 to 1 word per lesson during the DICT treatment. Visual analysis revealed that his performance was relatively stable within both treatments: His data maintained a moderate gap between the two treatments and did not overlap. Stable performance on the word recognition task with the DRAW treatment was maintained through the one generalization lesson that he was able to attend.

Cliff. From Lessons 3 to 12, Cliff gained 14 words following the DRAW treatment and 11 words following the DICT treatment. His gain scores ranged from 0 to 5 words per lesson during the DRAW treatment, and 1 to 3 words during the DICT treatment. Although the total number of words gained during the DRAW treatment represents a 27% increase over the DICT treatment, the data demonstrate a high degree of overlap and do not show a clear treatment effect. Visual analysis revealed a downward trend in performance for both treatment conditions. In contrast to the treatment phase, Cliff's performance on the word recognition task suggests an upward trend during the generalization phase. For Lessons 13 through 15 he gained 2, 4, and 4 words, respectively, with the DRAW treatment.

Molly. From Lessons 3 to 12, Molly gained 14 words following the DRAW treatment and 5 words following the DICT treatment. Her gain scores ranged from 2 to 4 words per lesson during the DRAW treatment, and 0 to 2 during the DICT treatment. Visual analysis revealed a small to moderate gap between her performance during the DRAW and DICT treatments. The DRAW treatment is consistently more successful (by 2 words, for four of the five lessons, and 1 word for one lesson). A gain score of 3 to 4 words per lesson on the word recognition task with the DRAW treatment was maintained throughout the generalization phase.

Elliot. From Lessons 3 to 12, Elliot gained 15 words following the DRAW treatment, but only 3 words following the DICT treatment. His gain scores ranged from 2 to 5 words per lesson during the DRAW treatment, and 0 to 1 during the DICT treatment. Visual analysis revealed that his performance was consistently higher following the DRAW treatment, with a gap that ranged in magnitude from 1 to 4 words. Elliot maintained a gain score of 3 to 4 words per lesson on the word recognition task with the DRAW treatment throughout the generalization phase.

Expressive vocabulary in story retell. Figure 5 is a comparison of the two treatments with respect to the number of words the four students used during story retell. Recall that vocabulary retell was scored as correct as long as it closely resembled standard pronunciation and matched the target grammatical class. Visual analysis of each student's data revealed a large amount of overlap between treatments.

Calvin. From Lessons 3 to 12, Calvin gained 9 words following the DRAW treatment and 7 words following the DICT treatment. His gain scores ranged from 0 to 4 words per lesson during the DRAW treatment, and 0 to 3 during the DICT treatment. Calvin gained 2 words during the one generalization lesson that he was able to attend.

Cliff. From Lessons 3 to 12, Cliff gained 1 word following the DRAW treatment and 1 word following the DICT treatment. His gain scores ranged from -1 to 1 word per lesson during the DRAW treatment, and 0 to 1 during the DICT treatment. Cliff gained 1, 2, and 0 words, respectively, during the generalization lessons.

Molly. From Lessons 3 to 12, Molly gained 3 words following the DRAW treatment and 0 words following the DICT treatment. Her gain scores ranged from 0 to 1 words per lesson during the DRAW treatment and were consistently 0 across all of the DICT lessons. Molly gained 0 words during the generalization lessons.

Elliot. From Lessons 3 to 12, Elliot gained 11 words following the DRAW treatment and 4 words following the DICT treatment. His gain scores ranged from 0 to 3 words per lesson during the DRAW treatment, and 0 to 2 during the DICT treatment. Elliot gained 2, 1, and 2 words, respectively, during the generalization lessons.

Summary of vocabulary acquisition measures. A summary of the preceding findings for written definitions, word recognition, and vocabulary retell can be found in Appendix E. The table compares each student's overall gains in the DRAW and DICT treatments for each measure. In all but one case, the gains were greater for the DRAW treatment than for the DICT treatment: Cliff's overall scores for vocabulary retell were equally low across treatments (i.e., only a gain of 1 word per treatment). A treatment effect is apparent for the written definition and word recognition measures. Appendix F is a display of the magnitude of the effect for these two measures (i.e., the gap between performance during the DRAW treatment versus the DICT treatment) for each student across the 10 treatment lessons. For the purposes of the scores displayed in Appendix F, the lessons have been arranged in pairs containing one DRAW lesson and one DICT lessons. The lessons were well suited to this type of comparison because the

alternating treatments design allowed a direct comparison of word sets that were comparable in word frequency across the DRAW and DICT treatments within each pair. Because of the step-down decrease in word frequency every two lessons (i.e., one DRAW and one DICT), each lesson pair was closely matched on average frequency across the six target words.

The gap between treatments for performance on the written definition task ranged from 0 to 5 across students: Half of the lesson pairs (10 of 20 lesson pairs; 20 lesson pairs = 5 pairs for each of the four students) favored the DRAW treatment by a large magnitude of 4 or 5 words (of a possible 6 words), six of the lesson pairs favored the DRAW treatment by a moderate magnitude of 2 or 3 words, three of the lesson pairs favored the DRAW treatment by a small magnitude of 1 word, and one lesson pair, for Molly, was equivalent across treatments.

The gap between treatments for performance on the word recognition task ranged from 1 to 4 across students: Nearly three fourths of the lesson pairs (14 of 20 lesson pairs) favored the DRAW treatment by a moderate to moderately large magnitude of 2 to 4 words (of a possible 6 words), four of the lesson pairs favored the DRAW treatment by a small magnitude of 1 word, and Cliff demonstrated two instances in which the DICT treatment favored the DRAW treatment by 1 word.

Table 3 is a comparison of performance on the three vocabulary acquisition measures within each student, with respect to the total number of words gained across the five lessons for each treatment. The following patterns were evident. In the DRAW treatment, vocabulary retell showed the least gain of the three measures, for every student. For Calvin and Elliot, written definitions yielded greater gains than word recognition: For Cliff and Molly the reverse was true. In the DICT treatment, written definitions showed the least gain of the three measures, for

every student. For Calvin and Elliot, vocabulary retell yielded greater gains than word recognition: For Cliff and Molly, again the reverse was true.

Reading Comprehension

Data were collected on the percent of propositions (i.e., ideas) included in students' spoken story retells, to measure reading comprehension. Figure 6 is a comparison of the two treatments with respect to the percent of propositions that the four students recalled. This figure displays changes in proposition retell from the Day 1 to the Day 3 story. (Recall that the Day 3 story was a different narrative than the Day 1 story.) Percent of propositions is reported rather than number of propositions because the latter varied across stories, from 18 to 22 total propositions. The majority of student scores fell within 20% above or below 0% (i.e., no gain), as highlighted in Figure 6. In the following paragraphs, change (gains and losses) will be compared for each student across the two treatments. Visual analysis of each student's data revealed complete overlap between treatment conditions.

Calvin. During Lessons 3 to 12, Calvin included an average of 11% fewer propositions on Day 3 of the DRAW treatment compared to Day 1, and 2% more propositions on Day 3 of the DICT treatment compared to Day 1. His gain scores in percent of propositions ranged from a loss of 43% to a gain of 18% of propositions per lesson (demonstrating a loss for three lessons, no change for one lesson, and a gain for one lesson) during the DRAW treatment, and a loss of 13% to a gain of 37% per lesson (demonstrating a loss for three lessons and a gain for two lessons) during the DICT treatment. Calvin's scores included two outliers: a loss of 43% of propositions in Lesson 5 (DRAW) and a gain of 37% of propositions in Lesson 6 (DICT). These outlying scores contributed to the average DRAW score being much lower than the DICT score for Calvin.

Cliff. During Lessons 3 to 12, Cliff included an average of 12% more propositions on Day 3 of the DRAW treatment compared to Day 1, and 1% more propositions on Day 3 of the DICT treatment compared to Day 1. His gain scores in percent of propositions ranged from a loss of 11% to a gain of 33% of propositions per lesson (demonstrating a loss for one lesson and a gain for four lessons) during the DRAW treatment, and a loss of 16% to a gain of 20% per lesson (demonstrating a loss for two lessons and a gain for three lessons) during the DICT treatment. Cliff's gain of 33% of propositions during Lesson 5 (DRAW) is a mild outlier, and did not greatly influence his overall average score for the DRAW treatment.

Molly. During Lessons 3 to 12, Molly included an average of 6% more propositions on Day 3 of the DRAW treatment compared to Day 1, and 4% more propositions on Day 3 of the DICT treatment compared to Day 1. Her gain scores in percent of propositions ranged from a gain of 1% to 11% propositions per lesson (demonstrating a gain for all five lessons) during the DRAW treatment, and a loss of 10% to a gain of 11% per lesson (demonstrating a loss for one lesson and a gain for four lessons) during the DICT treatment.

Elliot. During Lessons 3 to 12, Elliot included an average of 6% more propositions on Day 3 of the DRAW treatment compared to Day 1, and 16% more propositions on Day 3 of the DICT treatment compared to Day 1. His gain scores in percent of propositions ranged from a loss of 16% to a gain of 24% of propositions per lesson (demonstrating a loss for one lesson, no change for one lesson, and a gain for three lessons) during the DRAW treatment, and a loss of 11% to a gain of 63% per lesson (demonstrating a loss for two lessons and a gain for three lessons) during the DICT treatment. The sudden change in his performance during Lesson 6 was a primary reason for his increased average score during the DICT treatment.

Social Validity

Recall that the two undergraduate RAs returned to the classroom 6 weeks after the completion of data collection to talk with the teacher and students about the study. At this time, the teacher completed a 10-item questionnaire, related to her experiences as a collaborating partner, in which she indicated the degree to which she agreed with given statements on a 4-point scale (1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree; see Appendix D). On this questionnaire, the items referred only to "intervention", which from the teacher's point of view would include both treatments. These items requested information within three broad areas related to social validity: (1) the potential benefit to students, (2) the effectiveness of the treatment, and (3) the feasibility of the service delivery model.

Within the first area (i.e., potential benefits), the teacher's responses indicated that she strongly agreed that vocabulary was an important skill for reading comprehension and overall academic success (Items 1 and 2), that it was beneficial to have the SLP in the classroom (Item 9), and that the students enjoyed participating (Item 10). She also agreed that the intervention enriched her classroom curriculum (Item 7). Within the second area (i.e., effectiveness), her responses indicated that she strongly agreed that the intervention produced effective results (Item 3) and agreed that the students' overall classroom performance improved (Item 8). Within the third area (i.e., feasibility), the teacher's responses indicated that she strongly agreed that the time requirements of the intervention were reasonable (Item 4), and she agreed that the intervention was easily incorporated into her existing classroom schedule and could be maintained long-term (Items 5 and 6).

The teacher and students also met with the RAs individually to discuss their experiences in the study using a semi-structured interview format (see Appendix D). The teacher interview

revealed that she felt positive about her collaboration with the author and that all of the students in her class looked forward to having the author in the classroom. When speaking about her students, she said, "they definitely formed a relationship with her" and "looked forward to her being there, beyond just the students who were participating with her, all the students." She also noted that she did not have any preconceived notions about having an SLP in her classroom prior to the present study, but that she felt "very positively" about it at the time of the interview. She characterized the participating students as eager and excited to participate, saying "They would jump up when she came in the classroom, and when she was doing things one-on-one, everybody wanted to go first." She also reiterated how important vocabulary was to success in her classroom and in helping students become strong communicators, both with oral and written language. While she did not note a specific impact on the students' day-to-day tasks in the classroom, "because they were looking forward to an academic activity," she saw a more general improvement in classroom performance.

Student interviews revealed that all of the students enjoyed working in the small-group format, and three of the four students preferred the DRAW treatment to the DICT treatment. Although Elliot preferred the DICT treatment because he was enthusiastic about typing on the iPad, he did not enjoy copying the definitions onto his paper during this treatment. Elliot's interview revealed that he enjoyed drawing and he felt good about learning new words because "We're getting older." He also reported that the work he did as part of the present study might help him with his classroom work because his class was "learning about patterns, and words include patterns."

Calvin's interview revealed that he preferred the DRAW treatment because he enjoyed drawing and getting to see the other students' drawings, but he didn't like that the drawing days

64

could get loud. He said that the part of the DICT treatment that he enjoyed was the "hard work" and he described these days as a "challenge." He also said that he liked learning new words because "There's different words you can use for different things." Furthermore, he thought the work he did as part of the present study would help him in his classroom because if someone said a word, he would already know what it meant, if it was one of the words that was discussed in the small group.

During Cliff's interview, he reported that he enjoyed drawing, both on drawing days and in his life outside of the present study. He couldn't recall anything that he didn't like about the DRAW treatment and said, "I only liked it." He also couldn't recall anything about the DICT treatment, positive or negative. He felt good about learning new words because he enjoyed "being able to know what they mean", and while he reported thinking that the work he did as part of the present study would help him with his classroom work, he wasn't sure how.

Finally, similar to the other three students, Molly enjoyed drawing and preferred the DRAW treatment because she could draw whatever she wanted. There was nothing that she did not like about the DRAW treatment, but reported, "I didn't like typing or writing down" on the dictionary days. She said, "It was like a test." Overall, she felt good about learning new words "because they're words that I never heard before" and "she [the author] teaches me what they mean." During the semi-structured interviews, all students expressed positive feelings about their ability to learn new words and the importance of vocabulary in their classroom tasks. Additionally, the four students identified the drawing component as potentially the most memorable and most enjoyable part of either treatment.

CHAPTER 5

DISCUSSION

The present study focused on four students in one classroom at a private elementary school, who were identified by their teacher as struggling with classroom literacy tasks. The purpose of the study was to compare the effects of an interactive, multi-linguistic, and multimodal intervention (DRAW) to the effects of traditional instruction (DICT) on vocabulary acquisition and reading comprehension. Overall, the data suggest a treatment effect that favors the DRAW treatment when compared to the DICT treatment, in relation to vocabulary acquisition; all four students gained a larger number of words following the DRAW treatment. No clear treatment effect for reading comprehension was apparent. The validity of these findings is strengthened by the rigor of the present design, which met the quality indicator criteria outlined in Horner et al. (2005). Moreover, experimental control was demonstrated by the magnitude and consistency of the gap between the DRAW and DICT treatments for written definitions and word recognition.

Vocabulary Acquisition

Written definitions. All four students demonstrated greater gains in vocabulary knowledge with the DRAW treatment than with the DICT treatment, as measured by the accuracy of written definitions. The magnitude of the overall word gain with the DRAW treatment was moderate to large, ranging from about a third of the potential words acquired (10 and 11 of the 30 possible words for Molly and Cliff, respectively) to about two thirds of the words acquired (18 and 21 of the 30 possible words for Calvin and Elliot, respectively). Additionally, all four students demonstrated the least amount of vocabulary gain on the written definition measure (which was worse than word recognition or vocabulary retell) during the

66

DICT treatment, even though the treatment directly addressed writing definitions. This finding is particularly meaningful because written definitions are the most direct measure of word knowledge. Even though the students did not practice writing definitions during the DRAW treatment, the interactive, multi-linguistic, and multimodal activities appear to have resulted in a greater retention of semantic and syntactic knowledge of the words for all students. This improved level of acquisition was maintained, even as the words became less frequent (i.e., rarer and likely more difficult for students) in the DRAW treatment. Three of the four students maintained a stable level of low, near zero, acquisition throughout the DICT treatment, but Molly's downward trend during this treatment suggests that as the words became more difficult, the DICT treatment became even less effective. This might, in part, be due to the complexity of the definitions available in the online dictionary for these rare words.

Most notably, the drawing and conversational components of the DRAW treatment appeared to help students construct and retain semantic features of the target words, similar to Mathes et al. (2007). While composing their written definitions on Day 3, students would often verbally recall their pictures from Day 2 to help them remember what the words meant (Myers, 1983). For example, when thinking about what to write for a definition of the word *courteous* on Day 3 of Lesson 5, Calvin recalled how he had drawn a picture of Cliff not being courteous, by being rude and cutting in line, on Day 2 of the lesson. This is consistent with the Theoretical Model of Lexical Storage and Processing described by McGregor et al. (2002), and suggests that these students benefitted from nonlinguistic visual processing to construct and recall conceptual information related to new words (Ukrainetz, 1998).

Incorporating social stories into their pictures and talking about the words with other students as a small group (Vygotsky, 1978; Rogoff, 1990) also may have contributed to the

students gaining a better understanding of the words during the DRAW treatment (Fleming & Alexander, 2001). For example, when discussing the word *prestigious* on Day 2 of lesson 9, the students began to list professions that they thought were important and admirable. Molly began by talking about doctors and nurses, like her mom, and Calvin contributed that the President has a very prestigious job. Students then extended their ideas by drawing a picture of someone with a prestigious job (e.g., Calvin drew a picture of the President). Conversely, the lack of social-interaction and the decontextualized nature of the dictionary definitions during the DICT treatment may have hindered the students' ability to develop meaningful semantic constructs (Bereiter & Scardamalia, 1982).

   Grammatical class of words gained. With respect to the written definitions task, the DRAW treatment was more effective at teaching new vocabulary than the DICT treatment for target words in all grammatical classes (i.e., nouns, verbs, and adjectives), for every student. Gains were comparable for each grammatical class for Calvin and Elliot, for all lessons considered together. In contrast, Cliff and Molly showed an effect for the grammatical class of the words gained on the written definitions task, acquiring more nouns than verbs or adjectives. Semantic information related to nouns might have been easier to depict visually for these students, and therefore resulted in greater word learning from the drawing activity for noun targets. Verbs and adjectives are less tangible than nouns (Marinellie & Chan, 2006), because they describe an action or an attribute in relation to an object, and therefore may be more difficult to represent in simple drawings.

   Word recognition in oral reading. A treatment effect for word recognition in oral reading was demonstrated by three of the four students. Calvin, Molly, and Elliot were all able to recognize, and accurately pronounce, more target words during oral reading of short narratives

following the DRAW treatment than the DICT treatment. The magnitude of the overall word gain with the DRAW treatment was moderate to moderately large, ranging from a little over a third of the potential words acquired (12 of 30 words for Calvin) to about half of the words (14 of 30 words for Molly and Cliff, and 15 of 30 words for Elliot). This effect may have been a result of the increased number of opportunities for the students to hear and say the words during the DRAW treatment (Pany et al., 1982), which encouraged a strong mapping between the phonological and orthographic representations of each word (Perfetti, 2007; Plaut et al., 1996). Cliff did not demonstrate an effect, although his total gain scores for word recognition were slightly higher in the DRAW treatment than the DICT treatment (i.e., 14 and 11 words gained, respectively). Additionally, visual analysis revealed a downward trend in Cliff's performance in both treatments, as the words became less frequent. Recall that Cliff had a diagnosed visual processing disorder. This difference may have made visual recognition of the target words more difficult or more variable for him, regardless of the treatment condition. This effect appears particularly important for rare words, that Cliff likely had limited exposure to in previous reading experiences.

Expressive vocabulary in story retell. No lesson-by-lesson treatment effects were observed for students' use of the target words during spoken story retells. However, three of the four students demonstrated greater total gains for vocabulary retell in the DRAW treatment than the DICT treatment. In particular, Elliot had a total gain of nearly three times as many words used during story retell in the DRAW treatment as in the DICT treatment (11 and 4 words, respectively). Cliff's gains were comparable across treatments. Students were not instructed to attempt to use the target words during this task, so it may have been too open-ended to accurately measure the students' ability to use the words productively. As a result, this measure may have

been confounded by word preference or other factors related to vocabulary use in spontaneous language samples (Nagy & Herman, 1987). In other words, when students did not use the target words in the story retell, it may not indicate that they were unable to use the words, but rather that they chose not to. Word preference could be related to a variety of cultural and cognitive-linguistic factors as well as an individual's familiarity with the vocabulary and therefore, new words would be less likely to be chosen in either condition (White et al., 1990). Additionally, struggling readers have been shown to have difficulty on measures of expressive vocabulary (Chiappe et al., 2004). The relatively low intensity of the treatments in this study may not have enabled a high enough level of performance on this task to reveal a treatment effect.

Summary of vocabulary acquisition measures. All students achieved greater total word gains for the written definition and word recognition measures, and greater or comparable gains for the vocabulary retell measure, during the DRAW treatment compared the DICT treatment. Furthermore, students successfully gained approximately one third to two thirds (i.e., 10 to 21 words) of the 30 possible target words during the five DRAW treatment lessons on the written definition and word recognition tasks, a proportion that appears more substantial than predictions from previous theoretical accounts (Nagy & Herman, 1987). These gains were accomplished even as the vocabulary sets became progressively lower in word frequency. Moreover, these were not clinician-specific effects: In the present study, although a new SLP led the final three DRAW lessons, for the most part, each child continued to learn new vocabulary, for both the written definition and word recognition measures. Therefore, it appears that the DRAW treatment could be effectively generalized to clinical practice.

Consistent with Perfetti (2007) and the connectionist model of reading, these results suggest that the DRAW treatment was better able to support word learning in the three linguistic

domains that are essential for successful reading: semantics (i.e., written definitions), orthography (i.e., word recognition), and phonology (i.e., word recognition and vocabulary retell; Plaut et al., 1996). This is likely a result of the more comprehensive instructional program provided during the DRAW treatment. Similar to the improved assessment results of struggling readers following Moats' (2004) integrated intervention, the collaborative, multi-linguistic, and multimodal features of the DRAW treatment may have led to greater gains in student performance in the present study.

In contrast, the DICT treatment did not include explicit instruction on the orthographic or phonological representations of the target words. Perhaps as a consequence of the omission of these two linguistic domains, within the connectionist model (Plaut et al., 1996), this treatment was not able to effectively address the semantic component, as written definition gains were consistently lower than word recognition and vocabulary retell gains during the DICT treatment.

When examining individual student's gains across the three measures of vocabulary acquisition, it appeared that Calvin and Elliot demonstrated one pattern of performance, and Cliff and Molly demonstrated another. In the DRAW treatment, vocabulary retell showed the least gain of the three measures for every student, but for Calvin and Elliot, written definitions yielded greater gains than word recognition. This might have been because, of the four students, Calvin and Elliot were the strongest writers and were, therefore, better able to express the new word knowledge in writing. However, in the DICT treatment, they may not have been able to retain as much information about the new words and consequently, demonstrated fewer gains on the written definitions task. In contrast, Cliff and Molly demonstrated the greatest gains in word recognition, compared to the written definition and vocabulary retell tasks, in both treatment conditions. During the initial assessment battery, both students struggled with decoding

71

unfamiliar words: Cliff appeared to have difficulty with letter-sound correspondence and Molly often skipped unknown words. As a result of this history of decoding difficulties, these students might have placed a larger emphasis on learning how to read the target words than on understanding their definitions or incorporating them into spoken story retells.

Reading Comprehension

No clear effect of the treatment condition on reading comprehension was apparent; all four students demonstrated overlapping performance across treatment conditions. Although story retelling is a common measure of reading comprehension in assessment protocols (Leslie & Caldwell, 2006; Stieglitz, 2002) and reading research (Kintsch & Van Dijk, 1978; Stein & Glenn, 1979; McKeown et al., 1985; Idol & Croll, 1987; Fuchs et al., 1989; etc.), as used in the present study, spoken story retell was an open-ended task that allowed for great variability in student responses, and the inclusion of propositions was a complex measure. Both the task and the measure probably included many potential confounds such as working memory, executive functioning, and expressive language skills.

The free-form structure of responses likely contributed to the large amount of performance variability for each student across the lessons and did not control for the length of each retell (Fuchs et al., 1989) or student engagement during the retell activity. Additionally, this task was linguistically and cognitively complex, requiring students to understand the text as they read, retain the information, and generate a spoken retelling (Francis et al., 2005). The first component, understanding the text, was the intended measure, but interaction from the other two components was not controlled. For example, Molly and Elliot scored in the delayed and borderline ranges, respectively, on the Working Memory Index of the CELF-4. This may have resulted in difficulty with the second component, retaining the information. Cliff also scored in

the delayed and borderline ranges on expressive subtests of the CELF-4, and Calvin

demonstrated difficulty with confrontational naming tasks during assessment. These language

differences may have interfered with the students' ability to retell a story using spoken language

(Francis et al., 2005).

Perhaps due to both the high linguistic demands and the lack of structure in the task, story

retell appeared to be a difficult task for all of the students. This resulted in generally low-volume

responses that varied from day to day. With the exception of four outlying data points, all

students scored approximately 20% above or below 0% gain on the proposition retell measure.

All four outliers occurred during Lessons 5 and 6 (DRAW and DICT, respectively), but no

external factors were identified to account for student performance during these lessons. This

suggests that student performance was not effectively influenced by either treatment and that the

variability in the number of propositions recalled was likely due to chance.

Social Validity

The questionnaire revealed that the classroom teacher believed, overall, that the

vocabulary intervention provided in the present study benefitted students in several ways. (She

did not have knowledge that the two treatments were being compared. Consequently, her

responses on the questionnaire evaluated the intervention as a whole, i.e., considering the

DRAW and DICT treatments together.) First, consistent with the Report of the National Reading

Panel (NICHD, 2000), she believed that vocabulary was important to reading comprehension and

general academic success. Second, she corroborated the value of having an SLP in the

classroom for a push-in service delivery model, as investigated by Throneburg et al. (2000).

Third, the teacher believed that an SLP providing supplemental vocabulary instruction for

struggling readers could enrich the established classroom curriculum, and that in the present

study, this was effective in improving classroom performance. The latter findings are similar to those of Almasi et al. (2011), which showed that additional skilled professionals providing direct instruction within the classroom resulted in greater academic gains.

Additionally, the teacher's responses on the questionnaire indicated that, in her view, the push-in service delivery program was feasible and could be incorporated into her classroom routine in the long-term. While previous research has shown that push-in service delivery by SLPs is an effective method for vocabulary intervention (Boyle et al., 2007), the present study adds to this research by including the classroom teacher's perspective on the time commitment required, the benefits experienced by having the SLP in the classroom, and the overall feasibility of a push-in intervention.

The semi-structured interview conducted with the classroom teacher emphasized the positive role that the author played in the classroom environment. The teacher mentioned that all of the students in the class formed relationships with the author and the four participants looked forward to interacting with her. These relationships might have contributed to the students' eagerness to participate in the intervention. The teacher was enthusiastic about seeing this type of excitement directed towards an academic activity and felt that it benefitted the students because it improved their general attitudes towards, and performance during, other classroom activities. This might be particularly important for students with language and learning differences who commonly feel discouraged by academic tasks.

During Elliot's interview, he reported a preference for the DICT treatment because he was able to type the words on the iPad, but he did not enjoy copying the definitions. Typing single words might have been enjoyable for him because of the technological component (i.e., typing on the virtual keyboard using the iPad's touchscreen) and the low cognitive demands

of copying individual words. In contrast, copying the definitions was done with paper and pencil, which appeared less engaging than the iPad, and required a higher level of cognitive and linguistic processing in order to read the definitions, determine what to copy, and transcribe multiple words or phrases. Elliot also indicated that he felt good about learning new words because "We're getting older." Although he did not elaborate during this interview, one interpretation of this statement is that Elliot associates growth in word knowledge and vocabulary with older children and adults. This suggests a potentially positive attitude towards word learning because it makes him feel more grown up and mature.

Calvin echoed Elliot's idea related to learning more about words as you get older, and elaborated to include the value of using a variety of words for different purposes. Indeed, elaborating word meanings might help alleviate some of his word finding difficulties and allow him to retrieve appropriate words for specific contexts. His interview emphasized other potential aspects of life in which vocabulary knowledge might be important. He mentioned that one benefit of the work that he did as part of the present study might be that he would be better able to understand people, because he would have already learned what the words meant. This could suggest that Calvin has experienced difficulty understanding people in the past, possibly both inside and outside of the classroom, and extends the potential benefits of this type of intervention beyond academic success to other professional and social communication scenarios.

Cliff's interview revealed his interest in drawing outside of the classroom. Recall that he was diagnosed with a visual processing disorder and reading was very difficult for him. Therefore, incorporating a more creative visual task not only made the DRAW treatment more enjoyable, but might also have allowed him to utilize his strength as an artist to support word learning. Cliff struggled to remember specific elements of the DICT treatment, but easily spoke

about his experiences drawing. This might have been the result of a lower level of engagement

during the DICT treatment and greater saliency of the DRAW treatment.

Molly's interview focused on the contrast between the DRAW and DICT treatments with

regard to student contributions. She enjoyed the DRAW treatment because she had the freedom

to choose what she drew, whereas in the DICT treatment, the content of the definition was

determined by the online dictionary. Recall that Molly demonstrated some characteristics that

were consistent with autism spectrum disorder. One such characteristic was an emphasis on

specific topics of interest (e.g., Minecraft). The DRAW treatment allowed her to connect her

personal interests to the target words, both through discussion and drawing. In this way, the

DRAW treatment represented an educational model in which Molly could contribute to her own

learning and build her knowledge throughout the activities, rather than simply practice for tests,

which was how she seemed to view the DICT treatment. Another potential benefit of learning

new words is that Molly might become more willing to attempt to decode and interpret

unfamiliar words when reading, instead of skipping over them.

The interviews revealed that the students enjoyed working with an SLP and the specific

activities of the treatments. The teacher and each of the four students reported enjoying their

participation in the present study. This finding is promising, considering research emphasizing

the important influence that motivation can have on reading comprehension (Duke et al., 2004).

Three of the four students reported a preference for the DRAW treatment because they liked the

drawing activity and getting to talk with their peers about their drawings. This is consistent with

multimodal (Ukrainetz, 1998) and social-interactive (Vygotsky, 1978; Rogoff, 1990) approaches

to intervention. All students also reported being enthusiastic about learning new words. This

suggests that participating in the present study helped students develop an interest in, and self-

motivation related to, learning new words (i.e., word consciousness; Graves, 2006; Cunningham, 2014). For such students, learning strategies to become a "word gatherer" or "logophile" may ultimately enhance vocabulary growth over and above the gains made from explicit instruction.

Conclusions

A treatment effect was demonstrated for vocabulary acquisition. All four students displayed greater gains in the accuracy of written definitions following the DRAW treatment, compared to the DICT treatment. Nouns were more accessible to two of the four students. Three students displayed greater gains in word recognition in oral reading following the DRAW treatment, compared to the DICT treatment. No clear effects were observed in the data collected from the story retell task, which proved challenging for all of the students. Therefore, a treatment effect for reading comprehension was not demonstrated.

These findings are significant because they provide evidence of the value of SLPs implementing vocabulary intervention through a small-group push-in service delivery model. The results also support the relatively new concept of social-interactive vocabulary instruction that incorporates several linguistic aspects of a word (i.e., how it sounds, how it looks in print, and what it means), and opportunities to engage with words through multiple modalities (i.e., listening, speaking, reading, writing, and drawing). Student performance on vocabulary acquisition measures, bolstered by social validity measures, suggest that an interactive, multi-linguistic, and multimodal approach is an effective and feasible model for intervention targeting the acquisition of advanced sight vocabulary.

Clinical Implications

It is clear that the current method of dictionary-based vocabulary instruction that is used in the United States public schools, which relies almost exclusively on independent work in the

written modality, is not meeting the needs of all students (Riedel, 2007). Struggling readers may be able to learn better through interactive, multi-linguistic, and multimodal activities. The results from this study indicate that it is not necessary to shape instruction around the assessment tasks for vocabulary acquisition; this contradicts the idea of "teaching to the test." Therefore, even if a particular vocabulary assessment is required by the teacher, school, or district to be an individual written exam (e.g., writing word definitions); collaborative, integrated, and multimodal instructional strategies can be used to teach the knowledge and skills required, and help prepare students, for traditional assessments.

Pushing-in to classrooms for small-group intervention appears to be an effective and feasible model to address the vocabulary needs of many students in a time-efficient manner. This is increasingly important as caseloads continue to grow in the public school setting and clinicians work to meet the needs of a large number of students. The intervention used in the present study did not require an unreasonable amount of time from the classroom teacher, was easily incorporated into the classroom schedule, and could be maintained over a long period of time.

This type of intervention provides many opportunities for students to actively engage in word learning, successfully participate in activities, and build positive learning experiences. Although the students in this study were selected based on reading performance in the classroom, the only clinical characteristic that emerged consistently across all four students during the assessment battery was difficulty with classroom behavior. By second and third grade, students who struggle with literacy often have begun to demonstrate challenging behaviors. This may, in part, be due to repeated failures and frustration with academic tasks. Designing interventions

that utilize student contributions and actively engage learners might not only allow students to learn better, but also reduce their challenging behavior and improve their attitudes about school.

Limitations

The findings of the present study characterize the performance of four individual students on vocabulary and reading measures, but because of the small number of participants, caution should be exercised in generalizing these findings to a larger population. Additionally, the private school classroom may not have presented all of the challenges that would be faced in a public school setting. For example, the private school allowed a more flexible curriculum and did not adhere to Common Core State Standards (NGACBP, 2010) or other district-mandated curriculum guidelines.

Another limitation of the present study is that the results also only characterize short-term gains. Because the data are limited to pre- and post-treatment performance for each lesson, broader conclusions about vocabulary retention are not addressed. The treatment phase was also limited to 10 lessons (i.e., 5 DRAW and 5 DICT lessons), which may not have been sufficient for more global and lasting gains in vocabulary (Almasi et al., 2011). To this end, a maintenance phase would be a valuable addition to the study design.

Finally, the spoken story retell task may not have been an accurate measure of reading comprehension. Research has suggested that reading comprehension is an elusive construct to measure and that children can perform differently across reading comprehension tasks (Berninger et al., 2003). Student performance on this measure may have been confounded by larger processes such as working memory, executive functioning (i.e., organization), and expressive language skills. In other words, measuring the proportion of propositions recalled may not have reflected all aspects of students' reading comprehension.

79

Future Research

Future research is needed to develop an alternative comprehension task that will tap reading comprehension more effectively, while accounting for other cognitive, linguistic, and cultural factors that potentially impact comprehension ability. Possible areas to explore include: (1) combinations of explicit and implicit questions (Leslie & Caldwell, 2006), (2) structured story retells that provide more explicit instructions to prompt students to follow a basic story grammar and can be analyzed in larger thematic units rather than specific propositions (Newby et al., 1989; Stein & Glenn, 1979) and (3) written or spoken prompts that require students to recall how the words were used in the story.

Additionally, it is important to extend this research to a public school setting. This will allow access to a larger population of students and inclusion of students with diagnosed language impairments. Due to the strict curricular guidelines in many public schools, an extension to this setting will also provide opportunities to investigate how the DRAW treatment can be used to support student success with Common Core State Standards (NGACBP, 2010). Such a move will likely require an extended treatment phase that is integrated into the classroom curriculum and includes both short-term (benchmark) progress monitoring and long-term (cumulative) measures.

Finally, further research is needed to explore the effects of the individual components of the DRAW treatment. The intervention was developed as a combination of three instructional frameworks: interaction (or collaboration), multiple linguistic domains, and multimodality. The present study demonstrated the combined effect of all three components. Future research is needed to determine if all components are equally effective, and if certain students may benefit from a refined intervention that emphasizes one or more component.

Table 1

Student Profiles

	Calvin	Cliff	Molly	Elliot
Grade	3	3	2	2
Age (start)	8;7	8;7	7;0	7;4
Gender	Male	Male	Female	Male
Race	African Am.	European Am.	European Am.	European Am.
Diagnoses	none	Vis. Proc. Dis.	ADHD (w/Rx)	ADHD
GFTA-2	within norms	within norms	within norms	within norms
CELF-4 Core Language Scale	108 (70%ile)	96 (39%ile)	94 (34%ile)	123 (94%ile)
CELF-4 Working Memory Index	115 (84%ile)	100 (50%ile)	72 (3%ile)	85 (16%ile)
PPVT-4	101 (53%ile)	118 (88%ile)	130 (98%ile)	119 (90%ile)
EVT-2	100 (50%ile)	114 (82%ile)	120 (91%ile)	108 (70%ile)
BASC-2	Sch. Prob.	Int. Prob.	Int. Prob.	Int/Adapt. Sks
TONI-2	93 (32%ile)	105 (63%ile)	108 (70%ile)	103 (58%ile)
GORT-5	50%ile	5%ile	50%ile w/ low comprehension	50%ile w/ low comprehension

Note. Assessment results that indicated potential areas of difficulty for each student are highlighted. GFTA-2 = Goldman-Fristoe Test of Articulation – 2nd edition. CELF-4 = Clinical Evaluation of Language Fundamentals – 4th edition. PPVT-4 = Peabody Picture Vocabulary Test – 4th edition. Expressive Vocabulary Test – 2md edition. BASC-2 = Behavior Assessment System for Children – 2nd edition. TONI-2 = Test of Nonverbal Intelligence – 2nd edition. GORT-5 = Gray Oral Reading Test – 5th edition. Am. = American. Vis. Proc. Dis. = Visual Procession Disorder. ADHD = Attention-Deficit/Hyperactivity Disorder. Rx = Prescription Medication. Sch. Prob. = School Problems. Int. Prob. = Internalizing Problems. Int/Adapt. Sks = Internalizing Problems and Adaptive Skills.

Table 2

Lesson Schedule

| Lesson	Session	Day
	1-1 [pretest Lesson 1]	1
1	1-2	2
	1-3 [posttest Lesson 1]	3
	2-1 [pretest Lesson 2]	4
2	2-2	5
	2-3 [posttest Lesson 2]	6
	3-1 [pretest Lesson 3]	7
3	3-2	8
	3-3 [posttest Lesson 3]	9
etc.		

Note. The information in this table demonstrates the relation between "lessons", "sessions", and "days" within the experimental design. Fifteen lessons were completed in total.

Table 3

Total Vocabulary Acquisition

| | DRAW | | | DICT | | |
Student	Written Definitions	Word Recognition	Vocabulary Retell	Written Definitions	Word Recognition	Vocabulary Retell
Calvin	18	12	9	3	3	7
Elliot	21	15	11	1	3	4
Cliff	11	14	1	0	11	1
Molly	10	14	3	-2	5	0

Note. DRAW = experimental intervention. DICT = traditional instruction. The totals in this table were calculated by adding the individual student's scores for each measure across the five DRAW lessons and the five DICT lessons of the treatment phase (Lessons 3 to 12), separately. Cells highlighted in green indicate each student's highest gain score within each treatment condition. Cells highlighted in red indicate each student's lowest gain score within each treatment condition. The maximum score for each measure = 30 words (6 words per lesson, 5 lessons per treatment).
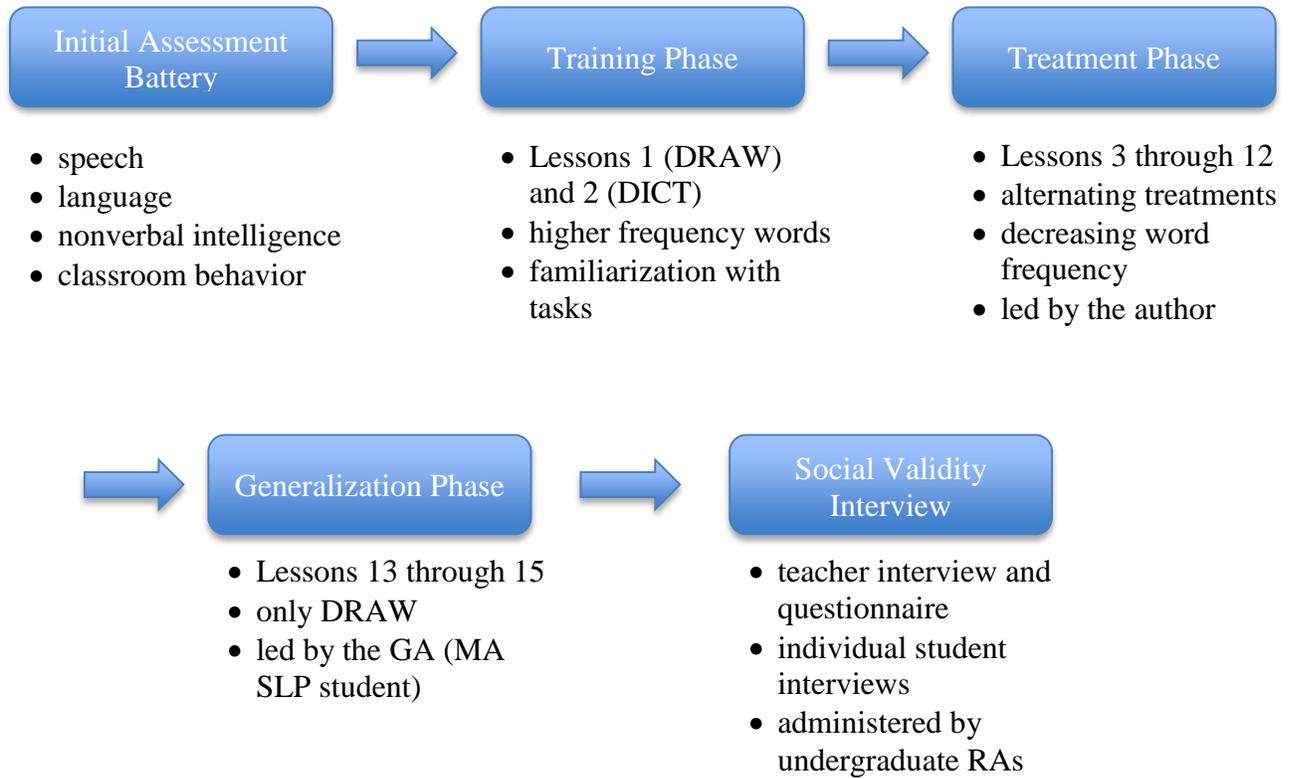
Figure 1. Main study design.

Figure 2. Written definitions. DRAW = experimental intervention. DICT = traditional instruction. Lessons 1 and 2 = training, Lessons 3 to 12 = treatment, Lessons 13 to 15 = generalization. Maximum gain possible for each lesson = 6 words (i.e., 10 lessons = 60 words, at 30 words per treatment).

Figure 3. Written definitions: Grammatical class of target words gained. Data are summed across all lessons within each treatment condition. GEN = generalization (three DRAW lessons). DICT = traditional instruction (five lessons). DRAW = experimental intervention (five lessons).

Figure 4. Word recognition in oral reading. DRAW = experimental intervention. DICT = traditional instruction. Lessons 1 and 2 = training, Lessons 3 to 12 = treatment, Lessons 13 to 15 = generalization. Maximum gain possible for each lesson = 6 words (i.e., 10 lessons = 60 words, at 30 words per treatment).
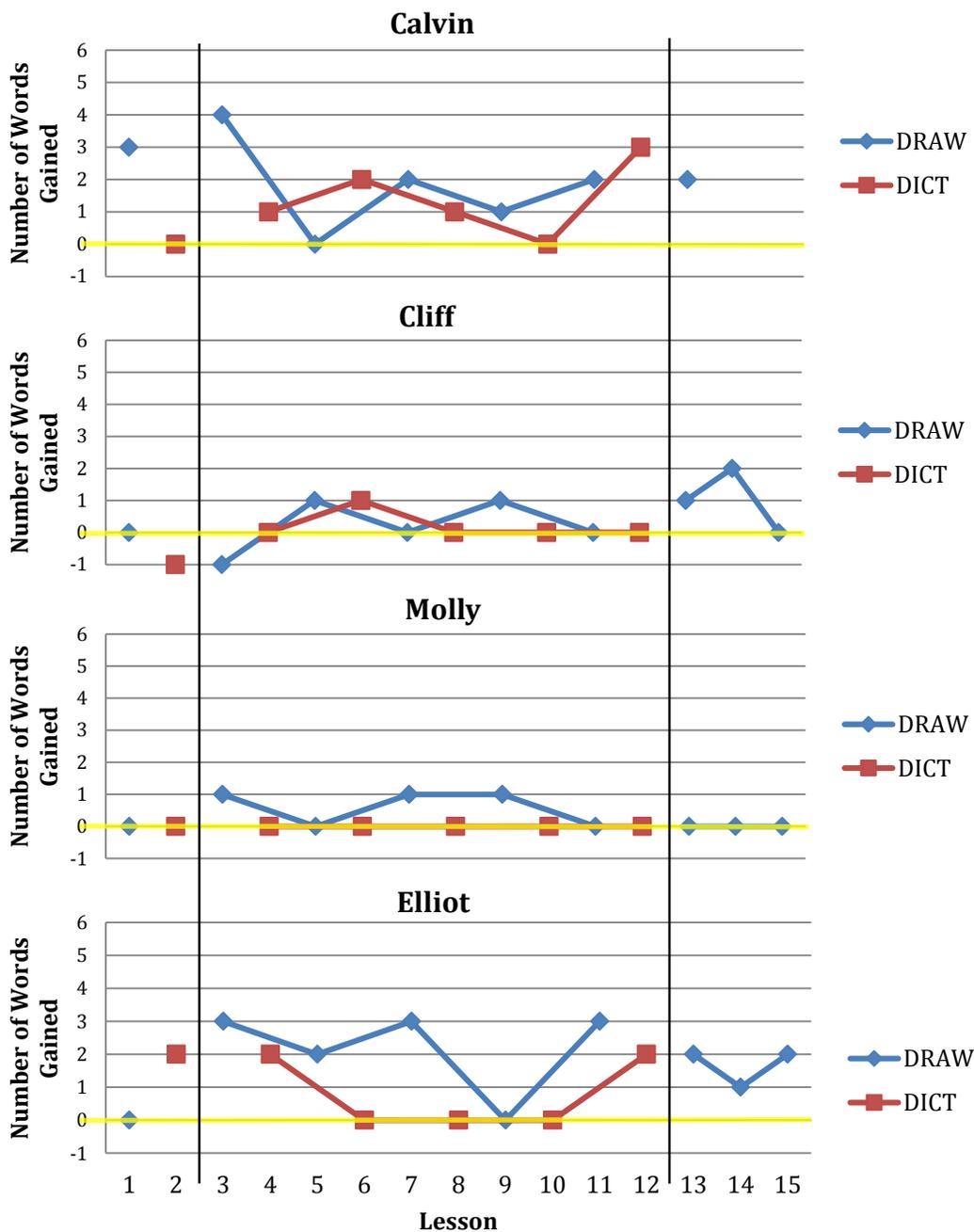
Figure 5. Expressive vocabulary in story retell (vocabulary retell). DRAW = experimental intervention. DICT = traditional instruction. Lessons 1 and 2 = training, Lessons 3 to 12 = treatment, Lessons 13 to 15 = generalization. Maximum gain possible for each lesson = 6 words (i.e., 10 lessons = 60 words, at 30 words per treatment).
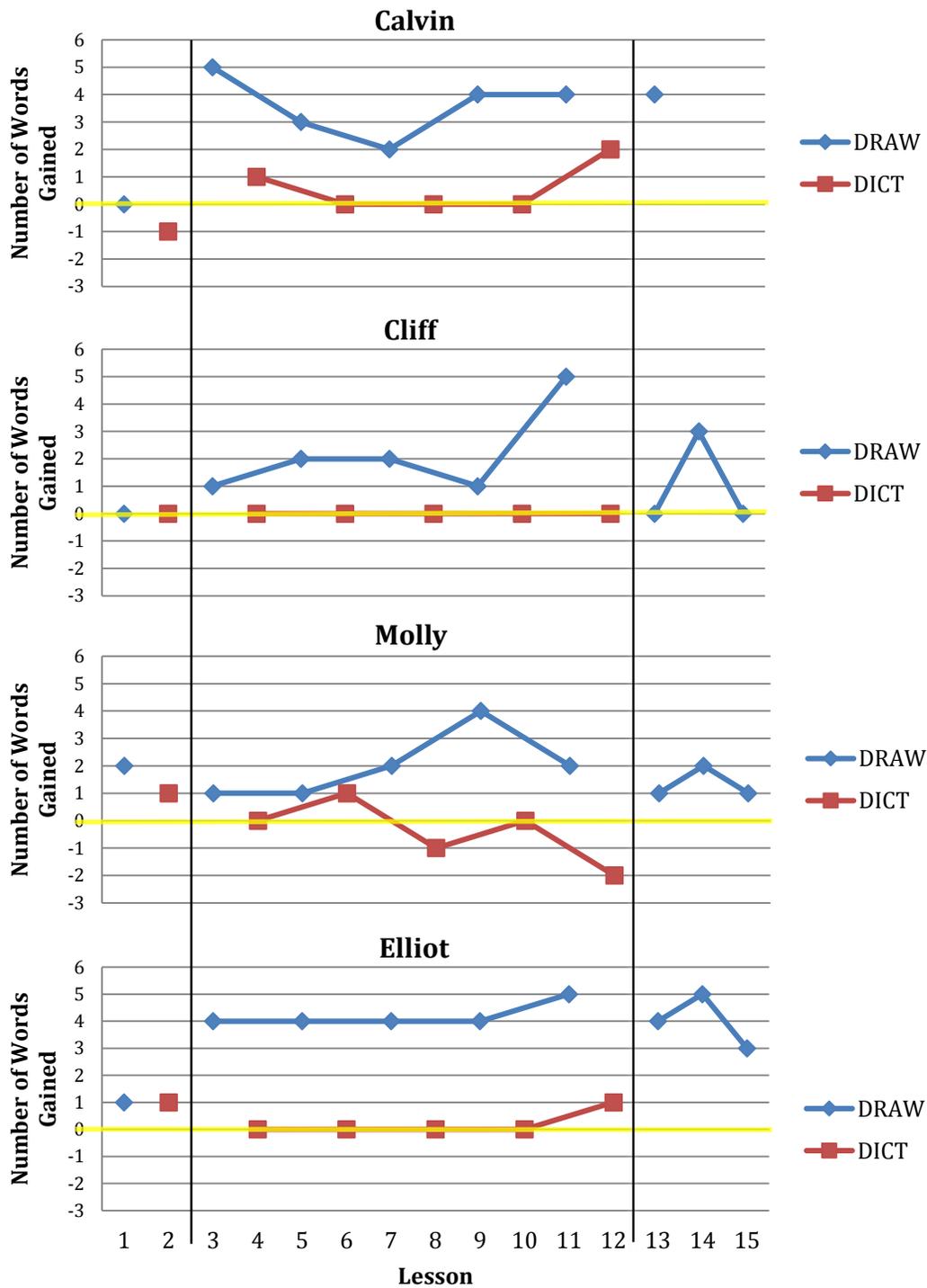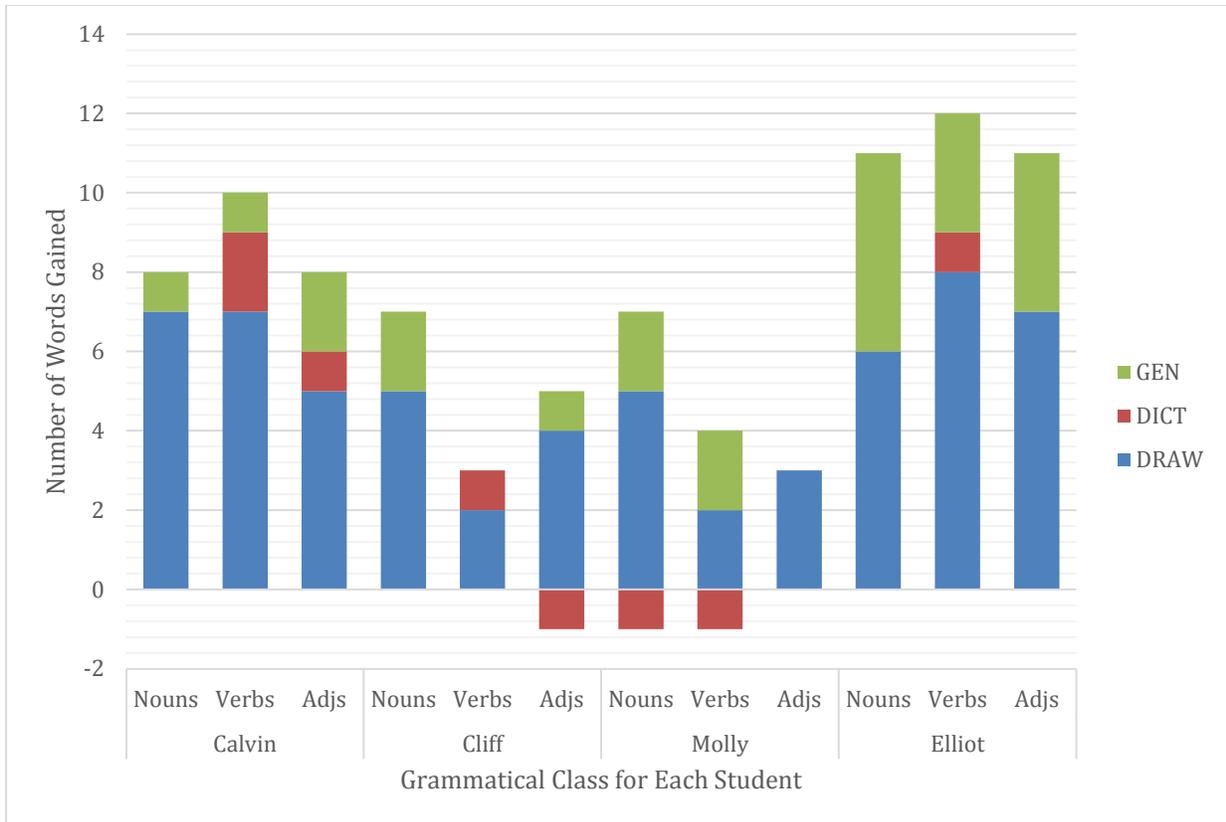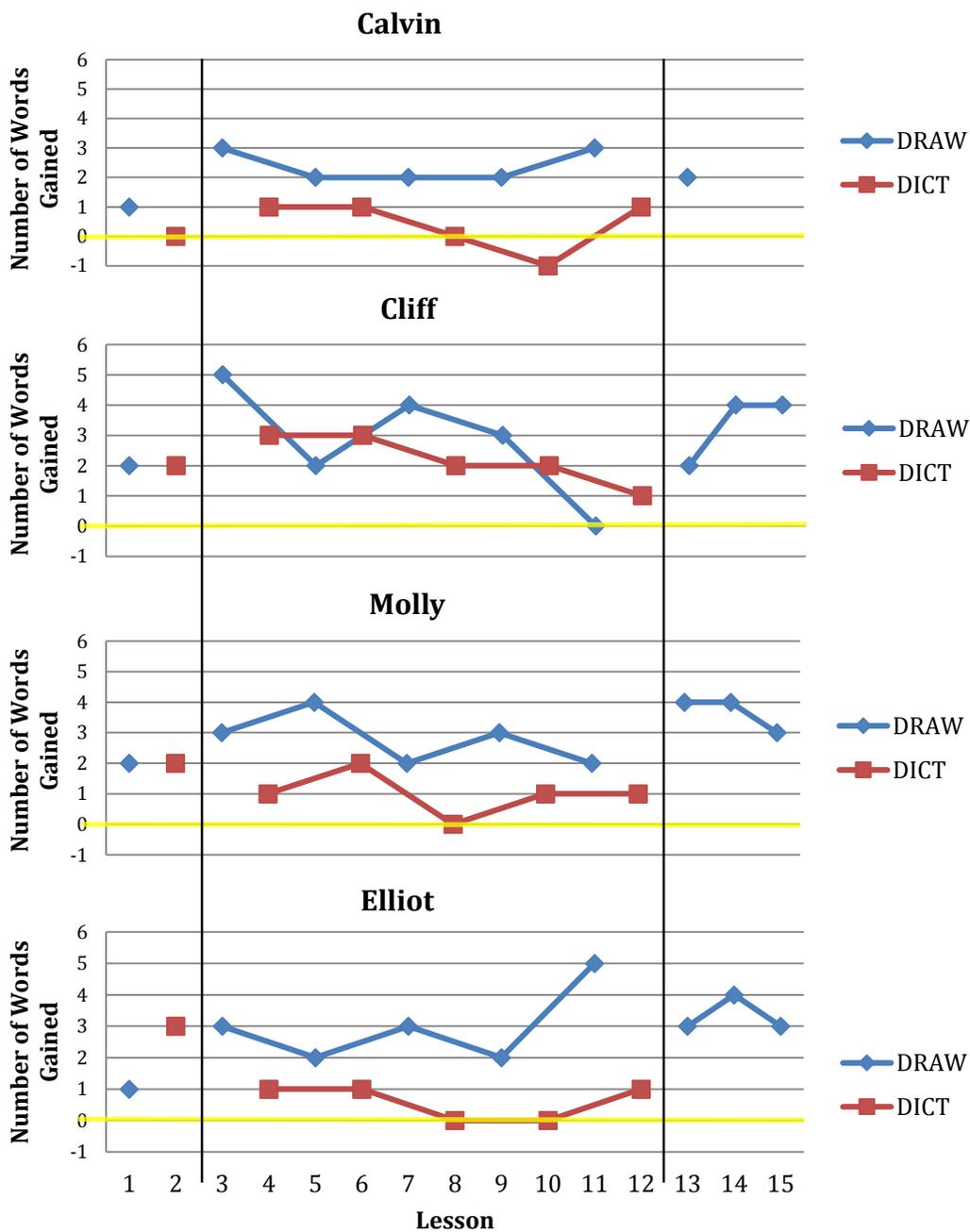
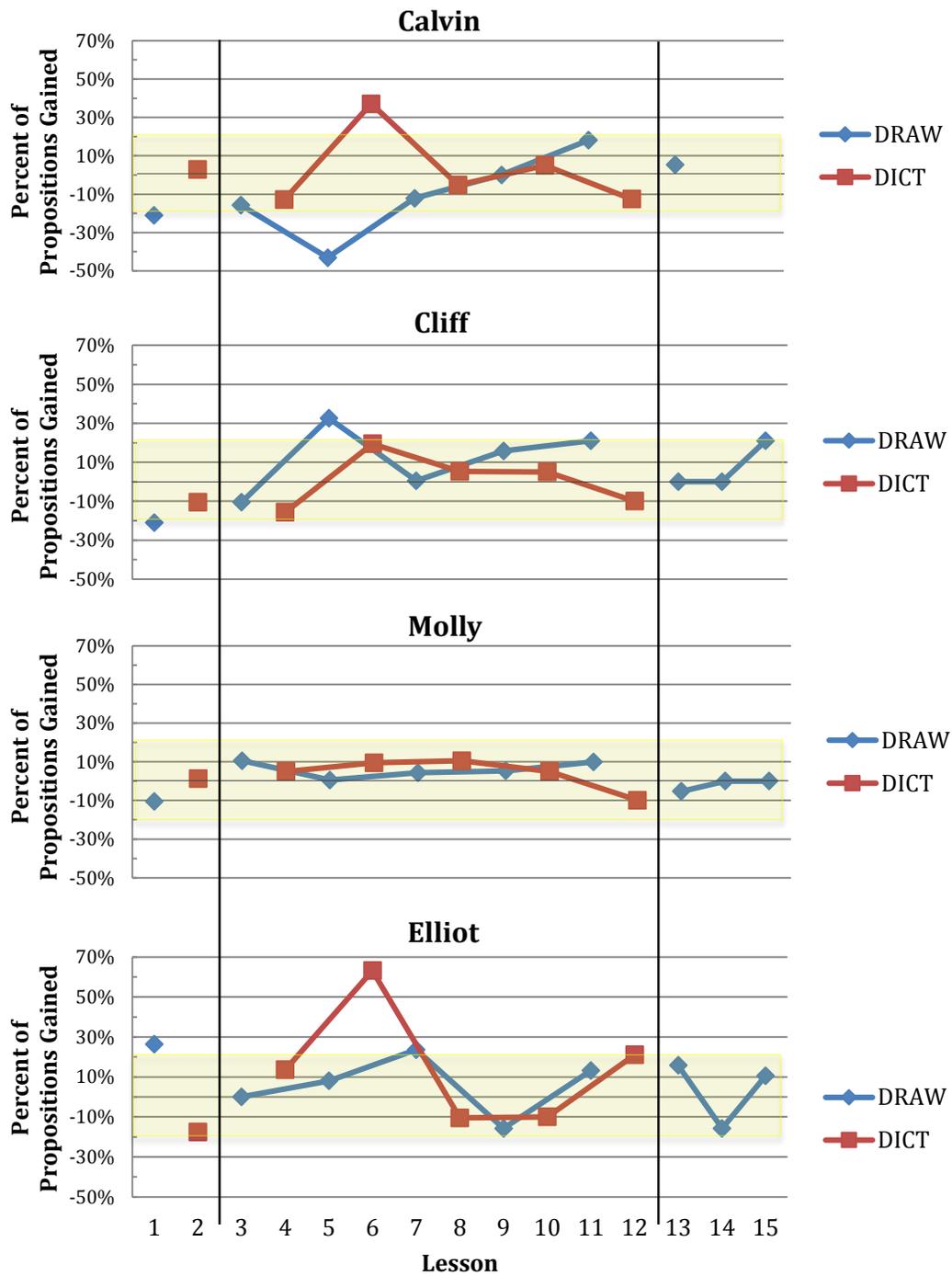*Figure 6.* Proposition retell. DRAW = experimental intervention. DICT = traditional instruction. Lessons 1 and 2 = training, Lessons 3 to 12 = treatment, Lessons 13 to 15 = generalization. Highlighted region = 20% above and below 0% (i.e., no gain). Total number of propositions possible in each lesson ranged from 18 to 22.

REFERENCES

Allinder, R. M., Dunse, L., Brunken, C. D., & Obermiller-Krolikowski, H. J. (2001). Improving fluency in at-risk readers and students with learning disabilities. *Remedial and Special Education, 22*(1), 48-54.

Almasi, J. F., Palmer, B. M., Madden, A., & Hart, S. (2011). Interventions to enhance narrative comprehension. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research* (pp. 315-328). New York, NY: Routledge.

Anderson, R. C., & Nagy, W. E. (1992). The vocabulary conundrum. *American Educator, 14,* 44-47.

Apel, K., & Masterson, J. J. (2001). Theory-guided spelling assessment and intervention: A case study. *Language, Speech, and Hearing Services in Schools*, *32*, 182–195.

Beaver, J., & Carter, M. (2003). *Developmental Reading Assessment.* Parsippany, NJ: Celebration Press.

Beck, I. L., & McKeown, M.G. (2001). Text talk: Capturing the benefits of read-aloud experiences for young children. *The Reading Teacher, 55*(1), 10-20.

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

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*(4), 506-521.

Bereiter, C., & Scardamalia, M. (1982). From conversation to composition: The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 1–64). Hillsdale, NJ: Erlbaum.

Berninger, V. W., Nagy, W. E., Carlisle, J., Thomson, J., Hoffer, D., Abbott, S., Abbot, R. D., Richards, T., & Aylward, E. (2003). Effective treatment for children with dyslexia in grades 4-6: Behavioral and brain evidence. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 381–417). Baltimore, MD: York Press.

Blachowicz, C., & Obrochta, C. (2005). Vocabulary visits: Virtual field trips for content vocabulary development. *Reading Teacher, 59,* 262-268.

Boyle, J., McCartney, E., Forbes, J., & O'Hare, A. (2007). A randomised controlled trial and economic evaluation of direct versus indirect and individual versus group modes of speech and language therapy for children with primary language impairment. *Health Technology Assessment, 11*(25), 1–158.

Brinton, B., & Fujiki, M. (2006). Improving peer interaction and learning in cooperative learning groups. In T. A. Ukrainetz (Ed.), *Contextualized language intervention: Scaffolding preK-12 literacy achievement.* Eau Claire, WI: Thinking Publications.

Brown, L., Sherbenou, R. J., & Johnsen, S. (1990). *Test of Nonverbal Intelligence—Second Edition.* Austin, TX: Pro-Ed.

Cabell, S. Q., McGinty, A. S., Breit, A., & Justice, L. (2008). Designing quality tier one learning environments for emergent and early readers. *Perspectives on Language Learning and Education, 15,* 4-12.

Cain, K., Oakhill, J. V., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology, 96*(1), 31-42.

Carroll, J. B., Davies, P., and Richman, B. (1971). *The American Heritage Word Frequency Book*. Boston, MA: Houghton Mifflin.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition, 91,* 77–111.

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

Chiappe, P., Chiappe, D. L., & Gottardo, A. (2004). Vocabulary, context, and speech perception among good and poor readers. Educational Psychology, 24, 825–843.

Cihak, D. (2011). Single-subject and case-study designs. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research* (pp. 419-433). New York, NY: Routledge.

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

Cunningham, P. M. (2014). *What really matters in vocabulary: research-based practices across the curriculum.* Upper Saddle River, NJ: Pearson Education, Inc.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology, 33*(6), 934-945.

Davis, F. B. (1944). Fundamental factors in reading comprehension. *Psychometrika, 9,* 185-197.

Duke, N., Pressley, M., & Hilden, K. (2004). Difficulties with reading comprehension. In C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and literacy: Development and disorders* (pp. 501-520). New York, NY: Guilford Press.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test—Fourth Edition.* Minneapolis, MN: Pearson, Inc.

Dyson, A. H. (2013). *Rewriting the basics: Literacy learning in children's cultures.* New York, NY: Teachers College Press.

Fleming, V. M., & Alexander, J. M. (2001). The benefits of peer collaboration: A replications with a delayed posttest. *Contemporary Educational Psychology, 26*, 588-601.

Francis, D. J., Fletcher, J. M., Catts, H. W., & Tomblin, J. B. (2005). Dimensions affecting the assessment of reading comprehension. In S. G. Paris & S. A. Stahl (Eds.), *Children's reading comprehension and assessment* (pp. 369-394). Mahwah, NJ: Erlbaum.

Fuchs, L. S. (1992). Identifying a measure for monitoring student reading progress. *School Psychology Review, 21,* 45-58.

Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1989). Monitoring reading growth using student recalls: Effects of two teacher feedback systems. *Journal of Educational Research, 83,* 103-110.

Fuchs, L.S., Fuchs, D., & Maxwell, L. (1985). The validity of informal comprehension measures. *Remedial and Special Education, 9*, 20-27.

Gillingham, A., & Stillman, B. W. (1997). *The Gillingham manual: Remedial training for children with specific disability in reading, spelling, and penmanship* (8th ed.). Cambridge, MA: Educators Publishing Service.

Girolametto, L., Weitzman, E., & Greenberg, J. (2012). Facilitating emergent literacy: efficacy of a model that partners speech-language pathologists and educators. *American Journal of Speech-Language Pathology, 21,* 47-63.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation—Second Edition*. Circle Pines, MN: American Guidance Service, Inc.

Gough, P. B., & Turner, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7,* 6-10.

Graves, M. F. (2006). *The vocabulary book.* Newark, DE: International Reading Association.

Graves, M. F., & Silverman, R. (2011). Interventions to enhance vocabulary development. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research* (pp. 315-328). New York, NY: Routledge.

Graves, M. F., & Watts-Taffe, S. M. (2002) The place of word consciousness in a research-based vocabulary program. In S. J. Samuels & A. E. Farstrup (Eds.), *What research has to say about reading instruction* (pp. 140-165). Newark, DE: International Reading Association.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2,* 127-160.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165-179.

Idol, L., & Croll, V. J. (1987). Story-mapping training as a means of improving reading comprehension. *Learning Disabilities Quarterly, 10*(3), 214-229.

Jitendra, A. K., Edwards, L. L., Sacks, G., & Jacobson, L. A. (2004). What research says about vocabulary instruction for students with learning disabilities. *Exceptional Children, 70*(3), 299-322.

Justice, L. M., & Kaderavek, J. N. (2004a). Embedded-explicit emergent literacy intervention I: Background and description approach. *Language, Speech, and Hearing Services in Schools, 35,* 201-211.

Justice, L. M., & Kaderavek, J. N. (2004b). Embedded-explicit emergent literacy intervention II:

      Goal selection and implementation in the early childhood classroom. *Language, Speech,*

      and Hearing Services in Schools, 35, 212-228.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*

      (2nd ed.). New York, NY: Oxford University Press.

Kim, Y., Apel, K., & Al Otaiba, S. (2013). The relation of linguistic awareness and vocabulary

      to work reading and spelling for first-grade students participating in Response to

      Intervention. *Language, Speech, and Hearing Services in Schools, 44,* 337-347.

Kintsch, W., & Keenan, J. M. (1973). Reading rate and retention as a function of the number of

      propositions in the base structure of sentences. *Cognitive Psychology, 5*, 257–274.

Kintsch, W., & Van Dijk, T, A. (1978). Towards a model of text comprehension and production.

      Psychological Review, 85(5), 363-394.

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for

      improving literacy. *Language, Speech, and Hearing Services in Schools, 40,* 341-351.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., &

      Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from

      What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*.

      Providence, RI: Brown University Press.

Lanter, E., & Watson, L. R. (2008). Promoting literacy in students with ASD: The basics for the

      SLP. *Language, Speech, and Hearing Services in Schools, 39,* 33-43.

Leslie, L., & Caldwell, J. (2006). *Qualitative reading inventory—Fourth edition.* Boston, MA:

      Allyn and Bacon.

Lindsey, K. A., Manis, F. R., & Bailey, C. E. (2003). Prediction of first-grade reading in

Spanish-speaking English- language learners. *Journal of Educational Psychology, 95,*

482–494.

Lipson, M. Y., & Wixson, K. K. (1986). Reading disability research: An interactionist

perspective. *Review of Educational Research, 56*(1), 111-136.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure as

recall. *Cognitive Psychology, 9*, 111–151.

Manion, V., & Alexander, J. M. (1997). The benefits of peer collaboration on strategy use,

metacognitive causal attribution, and recall. *Journal of Experimental Child Psychology,

67,* 268–289.

Marinellie, S. A., & Chan, Y. (2006). The effect of word frequency on noun and verb definitions:

A developmental study. *Journal of Speech, Language, and Hearing Research, 49,* 1001-

1021.

Marinellie, S. A., & Johnson, C. J. (2003). Adjective definitions and the influence of word

frequency. *Journal of Speech, Language, and Hearing Research, 46*, 1061–1076.

Mathes, P. G., Pollard-Durodola, S. D., Cárdenas-Hagan, E., Linan-Thompson, S., & Vaughn, S.

(2007). Teaching struggling readers who are native Spanish speakers: What do we know?

Language, Speech, and Hearing Services in Schools, 38, 260-271.

McGregor, K. K. (2004). Developmental dependencies between lexical semantics and reading. In

C. A. Stone, E. R. Silliman, B. J. Ehren, & K. Apel (Eds.), *Handbook of language and

literacy: Development and disorders* (pp. 302-317). New York, NY: Guilford Press.

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic

representation and naming in young children. *Journal of Speech, Language, and Hearing Research, 45*, 332-346.

McKeown, M. G. (1993). Creating effective definitions for young word learners. *Reading Research Quarterly, 28,* 16-31.

McKeown, M. G., Beck, I. L., Omanson, R. C., & Pople, M. T. (1985). Some effects of the

nature and frequency of vocabulary instruction on the knowledge and use of words. *Reading Research Quarterly, 20*, 522-535.

Miller, J. F. (2008). Systematic Analysis of Language Transcripts, English Version [Computer

software]. Madison, WI: SALT Software, LLC.

Moats, L. C. (2004). Efficacy of a structured, systematic language curriculum for adolescent

poor readers. *Reading and Writing Quarterly, 20*, 145-159.

Myers, C. (1983). Drawing as prewriting in preschool. In M. Myers & J. Gray (Eds.), *Theory and practice in the teaching of composition* (pp. 75-85). Urbana, IL: NCTE.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English?

Reading Research Quarterly, 19, 304-330.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications

for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Erlbaum.

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. Mosenthal, P. D.

Pearson, & R. Barr (Eds.), *Handbook of reading research* (pp. 269-284). Mahwah, NJ: Erlbaum.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language

    acquisition. *Reading Research Quarterly, 47*(1), 91-108.

Nation, K. (2009). Reading comprehension and vocabulary: What's the connection? In R. K.

    Wagner, C. Schatschneider, & C. Phythian-Sence (Eds.), *Beyond decoding: The*

    behavioral and biological foundations of reading comprehension (pp.176-194). New

    York, NY: Guilford Press.

Nation, K., & Snowling, M. (2004). Beyond phonological skills: Broader language skills

    contribute to the development of reading. *Journal of Research in Reading, 27*, 342–356.

National Governors Association Center for Best Practices, Council of Chief State School

    Officers. (2010). Common Core State Standards. Washington D.C.: National Governors

    Association Center for Best Practices, Council of Chief State School Officers.

National Institute of Child Health and Human Development. (2000). Report of the National

    Reading Panel. Teaching children to read: An evidence-based assessment of the

    scientific research literature on reading and its implications for reading instruction (NIH

    Publication No. 00-4769). Washington, DC: U.S. Government Printing Office.

Newby, R. F., Caldwell, J., & Recht, D. R. (1989). Improving the reading comprehension of

    children with dysphonetic and dyseidetic dyslexia using story grammar. *Journal of*

    Learning Disabilities, 22(6), 373-380.

Nippold, M. A. (1988). The literate lexicon. In M. A. Nippold (Ed.), *Later language*

    development: Ages nine through nineteen (pp.29-47). Boston, MA: College-Hill.

O'Connor, R. E., Notari-Syverson, A., & Vadasy, P. (1996). Ladders to literacy: The effects of

    teacher-led phonological activities for kindergarten children with and without disabilities.

    Exceptional Children, 63, 117–130.

Oullette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98,* 554–566.

Pany, D., Jenkins, J. R., & Schreck J. (1982). Vocabulary instruction: Effects on word knowledge and reading comprehension. *Learning Disability Quarterly, 5,* 202-215.

Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Plaut, D. C., McClelland, J. L., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103,* 56–115.

Reynolds, C. R., & Kamphaus, R. W. (2004). *Behavior Assessment System for Children—Second Edition*. Circle Pines, MN: American Guidance Service, Inc.

Riedel, B. W. (2007). The relation between DIBELS, reading comprehension, and vocabulary in urban first-grade students. *Reading Research Quarterly, 42*(4), 546-567.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Roth, F. P., & Baden, B. (2001). Investing in emergent literacy intervention: A key role for speech-language pathologists. *Seminars in Speech and Language, 22*(3), 163–173.

Salvia, J., & Ysseldyke, J. E. (1978). *Assessment in remedial and special education.* Boston, MA: Houghton Mifflin.

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

Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals—Fourth Edition*. Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company.

Sénéchal, M., LeFevre, J., Smith-Chant, B. L., & Colton, K. V. (2001). On refining theoretical models of emergent literacy: The role of empirical evidence. *Journal of School Psychology, 39*(5), 439–460.

Spear-Swerling, L. (2004). A road map for understanding reading disability and other reading problems: Origins, prevention, and intervention. In R. B. Ruddell & N. Unrau (Eds.), *Theoretical model and processes of reading* (pp. 517-573). Newark, DE: International Reading Association.

Stahl, S. A., & Nagy, W. E. (2006). *Teaching word meanings.* Mahwah, NJ: Lawrence Erlbaum Associates.

Stieglitz, F. (2002). *Stieglitz informal reading inventory.* Boston, MA: Allyn & Bacon.

Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38,* 934-947.

Sticht, T. G., Beck, L. J., Hauke, R. N., Kleiman, G. M., & James, J. H. (1974). *Auding and Reading: A Developmental Model*. Alexandria, VA: Human Resources Research Organization.

Target words. (2014). In *Wordsmyth: The premier educational dictionary-thesaurus*. Retrieved September 12 through November 10, 2014, from http://www.wordsmyth.net

Throneburg, R., Calvert, L., Sturm, J., Paramboukas, A., & Paul, P. (2000). A comparison of service delivery models: Effects on curricular vocabulary skills in the school setting. *American Journal of Speech-Language Pathology, 9*(1), 10–20.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Burgess, S. R., & Hecht, S. A. (1997). Contributions of phonological awareness and rapid automatic naming ability to the growth of word-reading skills in second- to fifth-grade children. *Scientific Studies of Reading, 1,* 161-185.

Ukrainetz, T. A. (1998). Stickwriting stories: A quick and easy narrative representation strategy. *Language, Speech, and Hearing Services in Schools, 29,* 197-206.

Vadasy, P. F., & Nelson, J. R. (2012). *Vocabulary instruction for struggling students*. New York, NY: Guilford Press.

Valencia, S. W. (2011). Reader profiles and reading disabilities. In A. McGill-Franzen & R. L. Allington (Eds.), *Handbook of reading disability research* (pp. 419-433). New York, NY: Routledge.

Vygotsky, L. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.

Wagner, R. K., Muse, A. E., & Tannenbaum, K. R. (Eds.) (2007). *Vocabulary acquisition: Implications for reading comprehension.* New York, NY: Guilford Press.

White, T. G., Graves, M. F., & Slater, W. H. (1990). Growth of reading vocabulary in diverse elementary schools: Decoding and word meaning. *Journal of Educational Psychology, 82,* 281-290.

Wiederholt, J. L., & Bryant, B. R. (2012). *Gray Oral Reading Test—Fifth Edition*. Austin, TX: Pro-Ed.

Williams, K. T. (2007). *Expressive Vocabulary Test—Second Edition.* Circle Pines, MN: American Guidance Service, Inc.

Wolter, J. A., Wood, K. A., & D'zatko, K. W. (2009). The influence of morphological awareness on the literacy development of first-grade children. *Language, Speech, and Hearing Services in Schools, 40,* 286–298.

APPENDIX A: TARGET WORDS

Lesson	Target Word	Word Frequency	Mean Frequency
Training Phase			
1	purpose	63.76	M = 64.38 SD = 40.04
1	journey	39.57	M = 64.38 SD = 40.04
1	caught	144.22	M = 64.38 SD = 40.04
1	prove	45.07	M = 64.38 SD = 40.04
1	broad	51.04	M = 64.38 SD = 40.04
1	unusual	42.59	M = 64.38 SD = 40.04
2	canoe	26.06	M = 26.66 SD = 29.99
2	soldiers	86.47	M = 26.66 SD = 29.99
2	sought	15.28	M = 26.66 SD = 29.99
2	patrolling	7.28	M = 26.66 SD = 29.99
2	scarce	10.42	M = 26.66 SD = 29.99
2	anxious	14.45	M = 26.66 SD = 29.99
Treatment Phase			
3	debris	7.3	M = 8.22 SD = 2.36
3	plateau	12.79	M = 8.22 SD = 2.36
3	pity	7.97	M = 8.22 SD = 2.36
3	gauge	7.57	M = 8.22 SD = 2.36
3	hasty	7.82	M = 8.22 SD = 2.36
3	amateur	5.87	M = 8.22 SD = 2.36
4	debt	6.55	M = 7.85 SD = 2.47
4	leisure	6.37	M = 7.85 SD = 2.47
4	conquer	7.34	M = 7.85 SD = 2.47
4	campaigned	12.46	M = 7.85 SD = 2.47
4	hearty	5.72	M = 7.85 SD = 2.47
4	fiery	8.63	M = 7.85 SD = 2.47
5	guild	6.46	M = 5.32 SD = 1.65
5	bureau	3.40	M = 5.32 SD = 1.65
5	halt	6.47	M = 5.32 SD = 1.65
5	honored	7.09	M = 5.32 SD = 1.65
5	courteous	3.26	M = 5.32 SD = 1.65
5	promising	5.25	M = 5.32 SD = 1.65
6	wharf	5.38	M = 4.94 SD = 2.65
6	abroad	3.13	M = 4.94 SD = 2.65
6	corral	7.75	M = 4.94 SD = 2.65
6	freighted	2.49	M = 4.94 SD = 2.65
6	doughy	2.51	M = 4.94 SD = 2.65
6	conscious	8.39	M = 4.94 SD = 2.65

7	colt	3.23	*M* = 3.73 *SD* = 1.02
	handkerchief	2.65	
	stalk	2.62	
	entombed	4.45	
	fatigued	4.80	
	vague	4.65	
8	heritage	5.88	*M* = 3.69 *SD* = 1.58
	hearth	3.78	
	reign	1.92	
	inherited	1.81	
	subtle	4.53	
	earnest	4.22	
9	ravine	3.36	*M* = 2.99 *SD* = 1.38
	colonel	5.55	
	merited	1.99	
	resign	2.19	
	prestigious	2.92	
	aerial	1.93	
10	depot	2.5	*M* = 2.83 *SD* = 1.67
	dialogue	5.31	
	wrought	4.23	
	mourning	1.90	
	buoyant	0.65	
	quaint	2.4	
11	league	1.97	*M* = 2.20 *SD* = 0.91
	reservoir	3.95	
	rendezvous	1.96	
	rationing	2.24	
	mischievous	1.61	
	perilous	1.44	
12	drought	3.24	*M* = 2.17 *SD* = 1.11
	exposure	1.05	
	hovering	1.98	
	ensure	2.26	
	perishable	0.87	
	guaranteed	3.6	
Generalization Phase			
13	heir	2.69	*M* = 2.02 *SD* = 1.18
	complexion	2.01	
	adjourned	0.58	
	fastened	1.75	
	hysterical	1.18	
	privileged	3.92	

14	knoll	1.37	*M* = 1.80 *SD* = 0.91
	trough	2.94	
	plagued	2.70	
	sieved	1.38	
	benign	0.50	
	foreign	1.91	
15	toll	4.44	*M* = 1.76 *SD* = 1.48
	coupe	1.14	
	forfeit	0.74	
	sabotaged	0.36	
	celestial	2.32	
	inherent	1.58	

Note. Word Frequency = U value from Carroll, Davies, and Richman (1971), which represents the frequency-per-million words, adjusted for the dispersion over 17 subject categories. Within each lesson, the target words are ordered in the following way: two nouns, then two verbs, then two adjectives. Across the 15 lessons, the mean frequency for each set of six target words decreases each lesson from Lessons 1 to 15.

APPENDIX B: STORIES

Lesson 1

Story 1:
John woke up early in the morning and rushed to get dressed. He had one goal today. His purpose was to find his twin sister an unusual gift. Every year before this, he had always gotten her the same thing, a new pair of gloves. This year, John wanted to prove that he could get her something different. He caught the bus and began his journey to the city. He looked in ten different stores and didn't find anything. He was about to give up and then he saw it, a gold necklace with a broad purple stone in it. It was perfect!

Story 2:
John woke up very early. He had an unusual day planned. His purpose was to try something that he had never done before. He was going to go fishing with his brother. When John was little, his brother teased him because he was scared of touching fish. John wanted to prove that he wasn't scared anymore. He met his brother outside and they started their journey to the lake. Soon, John saw a big, broad pier with several people already fishing. John and his brother got their lines ready and started fishing. John caught the first fish and quickly pulled it off of the hook. He did it!

Lesson 2

Story 1:
Last week, Sam watched a movie about a group of soldiers who discovered a hidden island. They were patrolling the waters around their military base when they saw smoke rising in the distance. They paddled their canoe towards the smoke for a long time and finally arrived at the shore of the new land just before dark. At first, they sought food and shelter because they were hungry and cold, but resources were scarce. They started to feel anxious as it got darker outside. They woke up early the next day and rushed home.

Story 2:
Yesterday, Sam read a book about a boy who had toy soldiers that came alive at night. The boy felt anxious when his parents turned his light off at bedtime. Every night when the boy went to sleep, the toys would start patrolling his room to help him feel safe. The boy sought out more toys every time he went to the store with his parents, but his opportunities were scarce. He planned to buy a toy car, a toy canoe, and a toy airplane so that his toys could get anywhere in his room.

Lesson 3

Story 1:
Alex felt so excited when he woke up this morning because the weatherman was predicting the first big summer storm and Alex was an amateur storm chaser. When the storm started to roll in, he tried to gauge how fast it was moving and where it was headed. His actions were hasty. He got in his car and chased after it! It was headed right for a town that was built on a plateau at the base of the mountain. Alex started to pity the people that lived in the town as he saw debris flying through the air. He hoped everyone was safe.

Story 2:
Alex felt nervous this morning. He was an amateur mountain biker, and today was his first big race of the summer. He was hasty as he got ready. He arrived at the race just before it started. He saw the other racers look back and pity him as he struggled to take his jacket off. Then, the race started! He tried to gauge how far behind he was, but there was debris on the track, so he needed to focus on steering. After a steep hill, he reached a plateau and could see all the racers ahead of him. He was determined to catch up.

Lesson 4

Story 1:
Tim was the youngest in his family, and like many youngest children, he was spoiled. So far, he had lived an easy life. His time was filled with hearty meals and leisure. One night, about a week ago, he had a scary dream that his school owed a huge debt to the bank and was going to close down if he didn't help. He woke up with a fiery resolve to work hard and do something important with his time. He decided to conquer his fear and run for student body president. He campaigned all week, and the election was finally here…

Story 2:
Tim grew up in a small, relaxed town where people preferred leisure and rarely worked. He noticed that many of his neighbors had a lot of debt because they only worked a few hours a week and didn't earn enough money to pay their bills. He campaigned to convince them to work more. He gave fiery speeches at town hall meetings to inspire people to conquer their laziness, but most people didn't come to the meetings. He decided he could get more people to listen if he invited them to his house for dinner. He cooked a hearty meal and invited everyone over…

Lesson 5

Story 1:
Billy was a talented artist. His family always told him how beautiful his pictures were, but he was nervous about showing his artwork to anyone else. One day, his teacher told him about a contest being held by the local artists' guild and convinced Billy to enter. He won the award for best student sketch and was honored by the artists at a special dinner. They told him that he had a promising career ahead of him. When Billy accepted his award, he asked the audience to halt their applause and gave a courteous speech to thank them. He proudly displayed the award on top of his bureau.

Story 2:
Billy loved to build things. He studied how to make birdhouses, and then tables, and now was determined to build a bureau for his mom. He worked hard, but was having a lot of trouble. He decided to halt the project and start again the next day. When he finished, his mom was so proud of him that she sent a picture of what he had made to the carpenters' guild. They were impressed, and honored Billy by publishing the picture in the newspaper. The article predicted that Billy would have a promising career as a carpenter. Billy was grateful, and wrote a courteous letter to thank them.

Lesson 6

Story 1:
Mike grew up in Maine and worked around boats his whole life. When he was a teenager, he used to go down to the harbor and help corral the smaller boats as they came in to dock at the wharf. His first paying job was to load boxes of goods onto big ships that were going to be freighted abroad. One summer, Mike was working on a very hot day. He was conscious of the temperature, but he didn't realize the problem. He was loading the goods, just like every other day, when he started to notice a doughy mess leaking out of the boxes…

Story 2:
Mike lived in Florida and loved the oranges that grew near his house. When he got older, he became a chef and made cookies flavored with the local fruit. His recipes became famous and he freighted his baked goods abroad to many countries. He would freeze them the night before, and then drive down to the wharf early in the morning. One summer, it was so hot in his car that everything melted into a doughy mess. He was conscious of his deadline, so he needed to corral all of his friends to help him remake the cookies before the ship was set to leave…

Lesson 7

Story 1:
Andy's family owned a stable that cared for 10 horses. Many of the horses were old and were quickly fatigued, but Andy's favorite horse was a 2-year-old colt that could run for hours. He could sprint faster than the old horses, but could also slow down to help Andy quietly stalk ducks when he was hunting. Andy's dad told him about a lake with the best ducks, but his directions were vague. Andy left at sunrise and within an hour, realized he was already lost. He felt entombed in the dense forest. He cleaned his compass with his handkerchief and headed towards home. He would have to try again tomorrow.

Story 2:
Andy lived on a farm with horses, chickens, pigs, and one old dog named Scotty. Scotty loved to stalk the animals and tackle them when they weren't looking. This worked great with the chickens and even the pigs, but was a little more difficult with the horses. One day, Scotty snuck up on a young colt at the farm. The horse was surprised and kicked over a huge wall of hay bales that entombed Scotty. By the time Scotty escaped, he was so fatigued he could barely stand. Andy soaked his handkerchief in cold water and wiped Scotty's dusty fur. For a month, Scotty had a vague smell of hay.

Lesson 8

Story 1:
Danny was home alone one night last winter, the day before his birthday. It was very cold so he decided to build a fire in the hearth. As he sat by the fire, he looked up at the painting above the mantle that his dad had inherited from his grandfather. It was a portrait of his great great grandfather during the years when Alfonso the Peacemaker reigned. The man in the portrait looked very earnest with a subtle glint in his eyes that made him seem kind. The painting reminded Danny of his Spanish heritage and of the tradition in his family of using your strength to promote peace.

Story 2:
Danny got home from school and saw a package waiting for him in his room. It was a birthday gift from his grandfather. When he opened it, he saw the woven blanket that his grandfather inherited from the Chief. The blanket was made when Chief Pontiac reigned, and told the story of the rich cultural heritage of the Ottawa Tribe. Danny wrapped himself in the blanket and knelt on the hearth so he could feel the heat from the fire. The blanket had a subtle smell of smoke, and Danny looked earnest as he thought about all the men in his family who had worn the blanket before him.

Lesson 9

Story 1:
Nick was a new pilot in the US Air Force. He had trained in aerial photography and his great work as a student had merited him praise from all of his teachers. Now, he had the prestigious honor of becoming the youngest pilot in the Air Force. He was getting ready for his first mission when he became extremely scared of heights. Just thinking about the deep ravine that he would have to fly over, made him feel sick. He knew he couldn't risk failing the mission, so he told his colonel the problem. They decided that Nick should resign his position until they could find out what was wrong.

Story 2:
Nick's dad had been a Marine for 50 years. He slowly made his way up the ranks and had become a colonel by the time Nick was in high school. Now, at the age of 72, Nick's dad was a full general. His work over the years had merited him many prestigious awards, and although he wouldn't have changed his decision to join the Marines, he couldn't help but feel that he had missed out on other things. He planned to resign next month and go bungee jumping at the ravine near the base. He heard about the beautiful aerial views there, and was excited to see it for himself!

Lesson 10

Story 1:
Zack just moved into a house on the beach, but a hurricane was coming and everyone was told to evacuate. When Zack returned, he saw the destruction wrought by the storm. He was mourning the loss of his house when his sister arrived. She had such a buoyant attitude even though it was a hard time. She was just happy that no one was hurt. They began a dialogue about what to do next, and she suggested they stay in the quaint inn she had seen in town. When Zack got a new house, they would go to the furniture depot to replace the things he had lost.

Story 2:
Zack went to school in Virginia. His class went on a field trip to see the town where the Battle of Bull Run took place. It was hard to imagine what the quaint town looked like with the destruction wrought by the war. After they saw the battlefield the class visited the cemetery, and the typically buoyant students became very serious while they were mourning the soldiers who lost their lives. After the field trip, the class returned to school and began a dialogue about the service of soldiers. After school, Zack went to the school supplies depot and decided to send care packages to children of veterans.

Lesson 11

Story 1:
Charlie was grounded because he had been very mischievous lately, so he wasn't allowed to go out with his friends. He had to sneak out his window onto the perilous ledge, carefully walk over to the fire escape, and climb down so that his parents wouldn't see him leave. He was going to rendezvous with his team for one last practice before the league final tomorrow. He arrived just as the captain was rationing the baseballs. Each player got three balls to practice hitting across the reservoir. This motivated the boys because if they didn't hit the ball hard enough, it would be lost in the deep water.

Story 2:
Charlie was on his way to rendezvous with Dan and Tom early in the morning. They had just joined a downhill skiing league and were practicing for their first race. On his way to the slopes, he drove past the frozen reservoir and saw some mischievous kids running and sliding across the ice. He started to worry that the rain from the night before might have caused perilous conditions if the top layer of snow was frozen solid. He got to the hill and saw his friends rationing what was left of the hot chocolate. They each finished a small cup and headed up the hill to practice.

Lesson 12

Story 1:
Chris lived in California, and everyone in his town was trying to use less water because they were in a drought. His mom stopped watering her garden and had to cover all of her rose bushes in the backyard to ensure that exposure to the strong sunlight and dry heat wouldn't kill the flowers. People don't always think of flowers as perishable, but if they do not get enough water, they will die. Thankfully, the forecast for tomorrow was calling for rain! It wasn't guaranteed, but when Chris woke up the next morning he saw clouds hovering in the sky.

Story 2:
Chris was at school, and his class was learning about a great drought in Africa that was killing all the crops. The people who lived there needed to find ways to preserve perishable foods to ensure that their families would have something to eat the next year. They developed ways to dry meat with direct exposure to the hot sun, but they had to be careful because if they left the meat outside too long, they would see vultures hovering overhead. To solve this problem, they built cages around the meat so it was guaranteed the vultures would stay away.

Lesson 13

Story 1:
Joe was only 18 years old, but he was the heir to the French throne and might become king very soon. His father's illness was getting worse. Joe went to visit him this morning and was worried by his pale complexion. Before Joe left his bedside, his father fastened a pin with the royal crest to his jacket and said he would be a great king. Joe was meeting with the nobles that afternoon and before the meeting adjourned, he needed to convince them he was ready. They thought it was hysterical to call him king and laughed at the idea of such an irresponsible, privileged kid ruling the country.

Story 2:
Joe was the heir to his father's insurance company, but he knew he wanted to be a comedy writer. His jokes were hysterical and writing made him happy. Joe was headed to his final staff meeting today and before the meeting adjourned, he planned to quit his job at the company. He stepped out of the car, fastened his tie for the last time, and walked into the meeting. He sat down at the table and his complexion turned bright red. He couldn't wait any longer. He burst out and said that he felt privileged to have worked at the company, but he was quitting to become a writer.

Lesson 14

Story 1:
Ben lived on a farm that grew corn and raised chickens. Last year, after the seeds began to sprout, he noticed a few small red bugs crawling on the plants. They were some sort of foreign beetle that he didn't recognize, but they looked a lot like ladybugs so he thought they were benign. They plagued the farm for a whole year and ate every plant in sight. The harvest that year was very small. Ben's family didn't even have enough corn to feed the chickens, so they sieved the dirt from the knoll behind their house and filled the trough with worms until they could buy more chicken feed.

Story 2:
Ben lived in Montana with his family. Lately, their town had been plagued by bad luck. There were terrible storms and many people were getting sick. Ben's dad heard that there was enough gold in California for everyone to fill a trough, and many foreign people were traveling there to get rich. Ben's parents decided to move. When they arrived, they joined a group that sieved the dirt from the American River. One day, they noticed everyone gathered on a knoll near the river, crowded around something covered in green slime. It looked gross, but it was benign. They cleaned off the slime and revealed a huge piece of gold.

Lesson 15

Story 1:
Ryan was the basketball team manager and today was the championship game. At 5am, he left in his small coupe to pick up two of the team's star players. He stopped to pay a toll on the highway and realized all his money was gone! His inherent distrust made him suspect the other team had sabotaged him. If he couldn't find a way to pay, he wouldn't be able to pick up his players, and his team would have to forfeit. He pulled to the side of the road and saw the sun starting to rise. He was inspired by the celestial beauty and felt determined to find a solution.

Story 2:
Ryan decided to stay home last night and watch a movie. The movie was about a spy on a secret mission to find the person who had sabotaged the election. The first scene showed the spy speeding down the highway in his silver coupe, not even stopping to pay the toll. He lost control of the car and crashed off the side of the road. He woke up to a view of the celestial bodies in the clear night sky. He was running out of time! He needed to convince the candidates to forfeit. He was hoping their inherent honesty would help them see that the election was not fair.

APPENDIX C: FIDELITY CHECKLIST

Category	Question	Yes	No
Setting	Did the lesson begin between 9:30-10:00am?		
	Did the lesson take place in the idea room?		
Dosage	Did the lesson take between 20-30 minutes?		
Population	Were all students present?		
Intervention A: Dictionary	Were all students provided with an online dictionary?		
	Were all students provided with a writing sheet?		
	Were all students provided with a pencil?		
	Did students work independently?		
	Did students complete all six definitions?		
Intervention B: Drawing	Were all students provided with written definitions?		
	Were all students provided with a drawing sheet?		
	Were all students provided with a pencil?		
	Did the interventionist highlight what the word looks like?		
	Did the interventionist highlight what the word sounds like?		
	Did the interventionist highlight personal connections?		
	Were students given the opportunity to read the words?		
	Were students given the opportunity to write the words?		
	Were students given the opportunity to say the words?		
	Were students given the opportunity to hear the words?		
	Were students given the opportunity to draw related pictures?		
	Did students complete all six pictures?		
	Did students collaboratively write a sentence for each picture?		

APPENDIX D: SOCIAL VALIDITY

TEACHER QUESTIONNAIRE: (completed by the teacher 6 weeks after treatment)

Please indicate the degree to which you agree or disagree with the following statements.

		Strongly Disagree	Disagree	Agree	Strongly Agree
1	The intervention focused on an important skill for reading comprehension.	1	2	3	(4)
2	The intervention focused on an important skill for overall academic success.	1	2	3	(4)
3	I believe that this intervention will produce effective results.	1	2	3	(4)
4	The time requirements of this intervention were reasonable.	1	2	3	(4)
5	The intervention was easily incorporated into my classroom routine.	1	2	(3)	4
6	The intervention could be integrated into my classroom schedule long-term.	1	2	(3)	4
7	The intervention enriched the existing literacy curriculum.	1	2	(3)	4
8	The students' classroom performance improved during intervention.	1	2	(3)	4
9	It was beneficial to have the speech-language pathologist in the classroom.	1	2	3	(4)
10	Students enjoyed the intervention.	1	2	3	(4)

TEACHER INTERVIEW:

1. Could you tell us a little about your overall experience during the study? (e.g., was collaboration with the researcher difficult? Did you or your students benefit from having the researchers in the classroom? Is there anything that would have improved the experience?)

2. From your perspective, how would you characterize the students' experiences during intervention? (e.g., did you see reluctance or excitement to participate?)

-Was there anything you noticed about our work that particularly engaged the children or were difficult... aspects of either intervention that you noticed influence the students...

-Thoughts on small group format – children's response to group format

3. Did you notice any changes in classroom performance for the student participants?

-Conclude by thanking the teacher for her time and asking if there is anything else she would like to share.

STUDENT INTERVIEW:

-Provide a brief introduction to tell each student that you are there to talk about the word study group that they were in last semester with Mary. Do not indicate that you have a relationship with me so that students feel free to express their true feelings.

-All of these questions can be followed up with "why?" or "tell me more about that."

1. What was something that you enjoyed about working with the small group?

2. What was something that you did not enjoy about working with the small group? What would you like to add or change?

3. Did you like the drawing days or the dictionary days better?

4. How do you feel about learning new words?

5. Do you think the work you did with Mary might help with other work that you do in your classroom?

-Conclude by thanking the student for his/her time and asking if there is anything else he/she would like to share.

APPENDIX E: TOTAL WORDS GAINED FOR EACH TREATMENT

	DRAW Total Gain (Lessons 3, 5, 7, 9, 11)		DICT Total Gain (Lessons 4, 6, 8, 10, 12)
Written Definitions			
Calvin	18	>	3
Cliff	11	>	0
Molly	10	>	-2
Elliot	21	>	1
Word Recognition			
Calvin	12	>	3
Cliff	14	>	11
Molly	14	>	5
Elliot	15	>	3
Vocabulary Retell			
Calvin	9	>	7
Cliff	1	=	1
Molly	3	>	0
Elliot	11	>	4

Note. The maximum possible number of total words gained for each student within either the DRAW or DICT condition was 30 words (5 lessons in each condition with 6 words per lesson).

APPENDIX F: LESSON-BY-LESSON DIFFERENCES IN VOCABULARY GAINS

Student	Lesson 3 - Lesson 4	Lesson 5 - Lesson 6	Lesson 7 - Lesson 8	Lesson 9 - Lesson 10	Lesson 11 - Lesson 12
Written Definitions					
Calvin	4	3	2	4	2
Cliff	1	2	2	1	5
Molly	1	0	3	4	4
Elliot	4	4	4	4	4
Word Recognition					
Calvin	2	1	2	3	2
Cliff	2	-1	2	1	-1
Molly	2	2	2	2	1
Elliot	2	1	3	2	4

Note. The odd numbered lessons are from the DRAW treatment. The even numbered lessons are from the DICT treatment. All values represent the difference in the number of words gained in the two treatment conditions (DRAW minus DICT) for the pair of lessons identified in the column header. Positive numbers indicate greater gains in the DRAW treatment. Negative numbers indicate greater gains in the DICT treatment.